Some Forms of Collectively Bringing About or ‘Seeing to it that’

Marek Sergot¹

Received: 23 October 2017 / Accepted: 13 March 2020 / Published online: 22 January 2021
© The Author(s) 2021

Abstract

One of the best known approaches to the logic of agency are the ‘stit’ (‘seeing to it that’) logics. Often, it is not the actions of an individual agent that bring about a certain outcome but the joint actions of a set of agents, collectively. Collective agency has received comparatively little attention in ‘stit’. The paper maps out several different forms, several different senses in which a particular set of agents, collectively, can be said to bring about a certain outcome, and examines how these forms can be expressed in ‘stit’ and stit-like logics. The outcome that is brought about may be unintentional, and perhaps even accidental; the account deliberately ignores aspects such as joint intention, communication between agents, awareness of other agents’ intentions and capabilities, even the awareness of another agent’s existence. The aim is to investigate what can be said about collective agency when all such considerations are ignored, besides mere consequences of joint actions. The account will be related to the ‘strictly stit’ of Belnap and Perloff (*Annals of Mathematics and Artificial Intelligence* 9(1–2), 25–48 1993) and their suggestions concerning ‘inessential members’ and ‘mere bystanders’. We will adjust some of those conjectures and distinguish further between ‘potentially contributing bystanders’ and ‘impotent bystanders’.

Keywords Logic of agency · Stit logics · Group agency · Joint action

1 Introduction

Much work on the logic of action focusses on the conditions under which it is the actions of a particular agent, or group of agents, that can be said to be the cause of, or responsible for, a certain outcome or state of affairs. With some exceptions, notably Pörn’s logic of ‘brings it about’ [20], the semantics is usually based on a

✉ Marek Sergot
m.sergot@imperial.co.uk

¹ Department of Computing, Imperial College London, London SW7 2AZ, UK

branching-time structure of some kind. The best known examples are the ‘*stit*’ (‘seeing to it that’) logics associated with Belnap and colleagues (see e.g. [2, 3, 5, 16, 17]). There are many variants, differing primarily in their treatment of temporal features. There are also connections to formalisms in game theory and theoretical computer science: the relationship between *stit* and Coalition Logic [19] for example is well established [9, 10], and there is a range of *stit* and *stit*-like logics interpreted on semantical structures such as transition systems [21, 22], concurrent game structures [6], models of distributed processes [8], and others. A recent paper by Ciuni and Lorini [12] compares various semantics for a temporal *stit* logic and also provides a brief classification and references to other forms of temporal *stit*.

Often, it is not the actions of an individual agent that bring about a certain outcome but the joint actions of a group of agents. Although *stit* logics typically provide operators for group as well as individual agency, to allow expressions of the form ‘group (set of agents) G sees to it that φ ’, group agency in *stit* has received comparatively little attention, beyond the observation that group *stit* has the property of superadditivity: if a group (set of agents) G sees to it that φ then every superset of G also sees to it that φ . There are exceptions.¹ Herzig and Schwarzenruber [15] investigate properties of group agency in an atemporal *stit* logic (the ‘deliberative *stit*’) including satisfiability and axiomatisability. Belnap and Perloff [4] discuss a form of *stit* they call ‘strictly *stit*’ and make a number of suggestions for further investigation. Belnap and Perloff’s suggestions will be examined in the second half of the paper.

The aim of this paper is to map out several different forms, several different senses in which the joint actions of a particular set of agents can be said to bring about a certain outcome, and to examine how these forms can be expressed in *stit* logics. The outcome that is brought about may be unintentional, and perhaps even accidental; the account deliberately ignores aspects of joint action such as joint intention, communication between agents, awareness of other agents’ intentions and capabilities, even the awareness of another agent’s existence. The aim is to investigate what can be said about collective agency when all such considerations are ignored, besides mere consequences of joint action.

Suppose, for example, that two agents a and b are positioned at either end of a table. Each can lift or lower its end. On the table stands a vase. If one agent lifts and the other does not, or if one agent lowers its end and the other does not, then the table tilts, and if it tilts, then the vase falls and breaks. Suppose that one agent lifts and the other lowers its end and the vase falls and breaks. Which of the two agents, if either, ‘brings about’, or is responsible for, the breaking of the vase? It seems wrong to pick on one or other of them; their actions collectively brought it about. If we add another agent into the picture, an agent c whose actions cannot affect the table or the vase, or interfere in any way with the lifting and lowering by a and b , then it also seems right to say that when a lifts and b lowers, it is the set $\{a, b\}$ of agents that brings about the breaking of the vase and not the set of agents $\{a, b, c\}$. c had nothing to do with it. In this example the set $\{a, b\}$ of contributing agents is unique; that need not always be so.

¹ See also e.g. [11, 13].

In examples like this it is possible that the agents are coordinating their actions. Perhaps it is their intention to dislodge the vase and break it. But it is also possible that they are acting quite independently; one lifts, the other lowers, for whatever reason (or perhaps for no reason at all), and the tilting of the table and the falling of the vase are incidental. Even then it is meaningful to say that agents a and b collectively 'brought it about' that the table tilted and the vase fell. The phrase 'saw to it that' is less appropriate, inasmuch as it hints at some kind of joint intention or purpose. It is important to distinguish between collective action (deliberative or not) and acting collectively with joint/shared intention. This paper considers only the first.

The account will be developed in a general formal framework that is common to many forms of *stit* and *stit*-like logics. Differences in specific kinds of *stit* are often due to their treatment of temporal features, and those features play no role in this paper. The development will then be related to the 'strictly *stit*' of Belnap and Perloff [4] and their further suggestions concerning the distinction between what they call 'inessential members' and 'mere bystanders'. We will adjust some of their conjectures and distinguish further between what we call 'potentially contributing bystanders' and 'impotent bystanders'.

Of course the *stit* framework is not the only way to talk about actions or the responsibility of an agent or group of agents for a certain outcome. This paper is limited to investigating how forms of group agency can be expressed within *stit*.

One final caveat: as in [4] we will not discuss in this paper the treatment of cases where the outcome of collective action is determined by the sequential actions of the members. We discuss only simultaneous joint actions, or actions that can be thought of as simultaneous.

Section 2 of the paper introduces the formal framework and some of the terminology encountered in the *stit* literature. Section 3 investigates how one can express in *stit* that G is a minimal set of agents whose actions jointly bring about a certain outcome and some of the properties of that construction. Section 4 introduces another sense of group agency—the set of agents who are the 'contributors' to the bringing about of the outcome in the sense that they are the members of those minimal sets. The second part of Section 4 relates that notion to Belnap and Perloff's distinction between 'inessential members' and 'mere bystanders'. Section 5 distinguishes further between 'potentially contributing bystanders' and 'impotent bystanders'.

2 Preliminaries

2.1 Syntax and Semantics

Ag is a non-empty, finite set of agents.

The language extends classical propositional logic with operators $\Box$, $[x]$ and $[G]$ for every agent x in Ag and every non-empty subset G of Ag .

$$\varphi ::= \text{any atom } p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi \mid [x]\varphi \mid [G]\varphi$$

We have the usual truth-functional abbreviations. $\Diamond$, $\langle x \rangle$ and $\langle G \rangle$ are the respective duals.

Semantics In abstract terms, models are relational structures of the form

$$\mathcal{M} = \langle W, \sim, \{\sim_x\}_{x \in Ag}, V \rangle \quad (1)$$

W is a non-empty set of possible worlds, $\sim$ and every $\sim_x$ are equivalence relations on W , and V is an evaluation function mapping propositional atoms to subsets of W . We further require that, for every $x \in Ag$:

$$\sim_x \subseteq \sim \quad (2)$$

These structures are similar to the abstract models of the deliberative *stit* (*dstit*) discussed in [1] *except that* the models presented there have a simpler form (the relation $\sim$ is omitted) because they incorporate an extra, very strong ‘independence of agents’ assumption characteristic of *stit*. They can be presented equivalently as models of the form Eq. 1 with additional (rather strong) restrictions on the $\sim_x$ relations over and above Eq. 2.

For a less abstract reading, W can be seen as the set of transitions in a transition system. For any transitions τ, τ' in W , $\tau \sim \tau'$ represents that τ and τ' are transitions from the same initial state, and $\tau \sim_x \tau'$ that τ and τ' are transitions from the same initial state ($\sim_x \subseteq \sim$) in which agent x performs the same action in τ' as it does in τ . This is the reading adopted in [21, 22] for example.

For readers familiar with *stit* logics, and the deliberative *stit* (*dstit*) in particular, W can be thought of as the set of moment-history pairs m/h . $\tau \sim \tau'$ then represents two moment-history pairs $\tau = m/h$ and $\tau' = m/h'$ through the same moment m . The equivalence relations $\sim_x$ determine each agent x ’s choice function.

The formal exposition that follows does not depend on any particular concrete reading however. We will tend to refer to the elements of W simply as ‘situations’.

The truth conditions are

$$\mathcal{M}, \tau \models \Box\varphi \quad \text{iff} \quad \mathcal{M}, \tau' \models \varphi \text{ for every } \tau' \text{ such that } \tau \sim \tau' \quad (3)$$

$$\mathcal{M}, \tau \models [x]\varphi \quad \text{iff} \quad \mathcal{M}, \tau' \models \varphi \text{ for every } \tau' \text{ such that } \tau \sim_x \tau' \quad (4)$$

We write $\models \varphi$ to denote that φ is valid in all models.

It is also convenient to employ a functional notation. Let:

$$alt(\tau) =_{\text{def}} \{\tau' \mid \tau \sim \tau'\} \quad (5)$$

$$alt_x(\tau) =_{\text{def}} \{\tau' \mid \tau \sim_x \tau'\} \quad (6)$$

alt is for ‘alternative’. (*alt*(τ) and *alt* _{x} (τ) are thus the $\sim$ and $\sim_x$ equivalence classes containing τ , respectively.)

Condition Eq. 2 says that, for every $x \in Ag$ and every $\tau \in W$,

$$alt_x(\tau) \subseteq alt(\tau) \quad (7)$$

The truth conditions can be expressed as:

$$\mathcal{M}, \tau \models \Box\varphi \quad \text{iff} \quad alt(\tau) \subseteq \|\varphi\|^{\mathcal{M}} \quad (8)$$

$$\mathcal{M}, \tau \models [x]\varphi \quad \text{iff} \quad alt_x(\tau) \subseteq \|\varphi\|^{\mathcal{M}} \quad (9)$$

where $\|\varphi\|^{\mathcal{M}}$ is the 'truth set' of formula φ in the model $\mathcal{M}$, i.e., the set of worlds in $\mathcal{M}$ in which φ is true.

In terms of transitions, $alt(\tau)$ is the set of transitions from the same initial state as τ , and $alt_x(\tau)$ is the set of transitions from the same initial state as τ in which x performs the same action as it does in τ . $alt_x(\tau)$ is the $\sim_x$ -equivalence class to which τ belongs: it can be thought of as the particular action performed by x (deliberatively, intentionally, or possibly unwittingly) in the transition τ . The set of all possible actions that could be performed by x at τ is the partition $\{alt_x(\tau') \mid \tau' \sim \tau\}$ of $alt(\tau)$.

In terms of *dstit*-models and moment-history pairs, when $\tau = m/h$, $alt(\tau)$ can be thought of as the set of histories passing through m , and $alt_x(\tau)$ is $Choice_x^m(h)$, i.e., the action performed by x at moment m in history h , or equivalently, the subset of histories passing through m in which x performs the same action at moment m as it does at moment m in history h .

$\Box$ and each $[x]$ are normal modal operators of type S5. The schema

$$\Box\varphi \rightarrow [x]\varphi \quad (10)$$

is valid for all agents x in Ag . $[x]$ is what some authors (e.g. Horty [16]) call the 'Chellas *stit*'.

2.2 Group Actions

This account generalises naturally to dealing with the joint actions of groups (sets) of agents. Let G be a non-empty subset of Ag . $alt_x(\tau)$ represents the action performed by x in τ , which is the subset of $alt(\tau)$ in which x performs the same action as it does in τ . $\bigcap_{x \in G} alt_x(\tau)$ is the subset of $alt(\tau)$ in which every agent in G performs the same action as it does in τ , and is thus a representation of the joint action performed by the group G in τ . (Again, readers familiar with *stit* logics will recognise the construction.)

The truth conditions are:

$$\mathcal{M}, \tau \models [G]\varphi \quad \text{iff} \quad alt_G(\tau) \subseteq \|\varphi\|^{\mathcal{M}} \quad (11)$$

where

$$alt_G(\tau) =_{\text{def}} alt(\tau) \cap \bigcap_{x \in G} alt_x(\tau) \quad (12)$$

$$\sim_G =_{\text{def}} \sim \cap \bigcap_{x \in G} \sim_x \quad (13)$$

For technical reasons it is convenient to allow expressions of the form $[\emptyset]\varphi$. $\sim_G$ and $alt_G(\tau)$ are defined above so that $\sim_{\emptyset} = \sim$ and $alt_{\emptyset}(\tau) = alt(\tau)$. Hence $\models [\emptyset]\varphi \leftrightarrow \Box\varphi$.

Expressed in the relational notation:

$$\begin{aligned} \mathcal{M}, \tau \models [G]\varphi & \quad \text{iff} \quad \mathcal{M}, \tau' \models \varphi \text{ for every } \tau' \in alt(\tau) \cap \bigcap_{x \in G} alt_x(\tau) \\ & \quad \text{iff} \quad \mathcal{M}, \tau' \models \varphi \text{ for every } \tau' \in alt_G(\tau) \\ & \quad \text{iff} \quad \mathcal{M}, \tau' \models \varphi \text{ for every } \tau' \text{ such that } \tau \sim_G \tau' \end{aligned}$$

Axiomatisation $\Box$ and every $[x]$ and every $[G]$ are normal modal operators of type S5. We have $\models [\{x\}]\varphi \leftrightarrow [x]\varphi$. The logic is the smallest normal logic containing all instances of the following axiom schemas, for all non-empty subsets G and H of Ag :

$$\begin{aligned} \Box & \quad \text{type S5} \\ [G] & \quad \text{type S5} \\ \Box\varphi & \rightarrow [G]\varphi & \quad \text{('necessity')} \end{aligned} \tag{14}$$

$$\begin{aligned} [G]\varphi & \rightarrow [H]\varphi & (G \subseteq H) & \quad \text{('superadditivity')} \\ [\emptyset]\varphi & \leftrightarrow \Box\varphi \end{aligned} \tag{15}$$

(The structure of the models is exactly the same as for ‘distributed knowledge’ and proofs can be found in any standard text on epistemic logic. See e.g. [14, Ch. 3, Thm. 3.4.1]. The modality $\Box$ causes no additional difficulty.)

$\Box\varphi$ could also be defined as shorthand for $[\emptyset]\varphi$, which simplifies the statement of some formal properties. For example, ‘necessity’ Eq. 14 would then just be a special case of ‘superadditivity’ Eq. 15.

2.3 Two ‘Brings it About’ Modalities

Let $\partial_x\varphi$ represent that agent x brings it about, perhaps deliberately, perhaps unwittingly, that φ . $\partial_x\varphi$ is defined as follows:

$$\partial_x\varphi \stackrel{\text{def}}{=} [x]\varphi \wedge \neg\Box\varphi \tag{16}$$

$\partial_x\varphi$ is satisfied at τ in a model $\mathcal{M}$ when: (1) (necessity) $\mathcal{M}, \tau \models [x]\varphi$, that is, φ is true in all alternatives $alt_x(\tau)$ of τ in which x acts in the same way as it does in τ , or as we also say, φ is necessary for how x acts in τ , or x guarantees that φ at τ ; and (2) (counteraction) had x acted differently than it did in τ then φ might have been false: there is an alternative τ' of τ in which φ is false (and hence in which x acts differently than it does in τ).

This is exactly the construction used in the definition of ‘deliberative *stit*’ [16, 17]

$$[x \text{ dstit}: \varphi] \stackrel{\text{def}}{=} [x]\varphi \wedge \neg\Box\varphi$$

The notation $\partial_x\varphi$ is chosen in preference to *dstit* partly because it is more concise, but more importantly, because *dstit*, in common with other forms of *stit*, incorporates a very strong ‘independence of agents’ assumption.² Nothing in this paper depends on those additional ‘agent independence’ properties. Since *dstit* models are a subclass of models of form Eq. 1 all logical properties of ∂_x are also properties of $[x \text{ dstit}: \cdot]$. $[x]$ is of type S5, and so, amongst other things:

$$\models \partial_x\varphi \rightarrow \varphi \tag{17}$$

²See e.g. Horty [16, p30]: ‘...simultaneous actions by distinct agents must be independent in the sense that the choices of one agent cannot affect the choices available to another; at each moment, each agent must be able to perform any of his available actions, no matter which actions are performed at that moment by the other agents.’

Notice that $\partial_x \varphi \wedge \partial_y \varphi$ is satisfiable even when $x \neq y$. Indeed

$$\models \partial_x \varphi \wedge \partial_y \varphi \leftrightarrow [x]\varphi \wedge [y]\varphi \wedge \neg \Box \varphi \quad (18)$$

Suppose for example that a and b are both pushing against a spring-loaded door (from the same side) and thereby keeping it shut, though either one of them is strong enough by itself to keep it shut. If k represents that the door is shut, then $\partial_a k \wedge \partial_b k$ is true: it is necessary for what a does that the door remains shut (a is strong enough by itself), and there is an alternative counterfactual situation in which the door fails to remain shut (namely, where neither a nor b pushes). So $\partial_a k$ is true. And likewise for b . The same example (an example of what is often referred to as 'overdetermination') works for *dstit*: $[x \text{ dstit}: \varphi] \wedge [y \text{ dstit}: \varphi]$ for $x \neq y$ is satisfiable in *dstit* models.

It is possible to define a stronger kind of 'brings it about' modality. Let:

$$\partial_x^+ \varphi =_{\text{def}} [x]\varphi \wedge \neg [Ag - \{x}]\varphi \quad (19)$$

$\partial_x^+ \varphi$ is satisfied in τ in a model $\mathcal{M}$ when: (1) (necessity) φ is necessary for how x acts in τ , $\mathcal{M}, \tau \models [x]\varphi$; and (2) (counteraction) had x acted differently than it did in τ then φ might have been false *even if all other agents*, besides x , had acted in the same way as they did in τ .

We have the following properties:

$$\models \partial_x^+ \varphi \rightarrow \partial_x \varphi \quad (20)$$

$$\models \partial_x^+ \varphi \wedge \partial_y^+ \varphi \rightarrow \perp \quad (x \neq y) \quad (21)$$

$\partial_x^+ \varphi$ represents a sense in which it is x and x alone who is responsible for bringing about that φ .

Note that:

$$\models \partial_x^+ \varphi \rightarrow \partial_x \varphi \wedge \bigwedge_{y \neq x} \neg \partial_y \varphi \quad (22)$$

$$\models \partial_x^+ \varphi \leftrightarrow \partial_x \varphi \wedge \bigwedge_{y \neq x} \neg \partial_y \varphi \quad (23)$$

The former is because $\models [y]\varphi \rightarrow [Ag - \{x}]\varphi$ for any $x \neq y$ ('superadditivity' Eq. 15), and so $\models \bigvee_{y \neq x} [y]\varphi \rightarrow [Ag - \{x}]\varphi$. But Eq. 22 is not a biconditional. That is because $\not\models [Ag - \{x}]\varphi \rightarrow \bigvee_{y \neq x} [y]\varphi$. If the actions of $Ag - \{x\}$, collectively, guarantee that φ , that does not imply there is some individual y in $Ag - \{x\}$ whose actions guarantee that φ .

The definitions can be generalised.

Definition 1 For every subset G of Ag

$$\partial_G \varphi =_{\text{def}} [G]\varphi \wedge \neg \Box \varphi$$

$$\partial_G^+ \varphi =_{\text{def}} [G]\varphi \wedge \neg [Ag - G]\varphi$$

With $[\emptyset]\varphi \leftrightarrow \Box \varphi$, we have $\models \partial_{\emptyset} \varphi \leftrightarrow \perp$, $\models \partial_{\emptyset}^+ \varphi \leftrightarrow \perp$ and $\models \partial_{Ag}^+ \varphi \leftrightarrow \partial_{Ag} \varphi$.

Since $\models \neg [Ag - G]\varphi \rightarrow \neg \Box \varphi$ ('necessity' Eq. 14), $\partial_G^+ \varphi$ can be defined equivalently, for all subsets G of Ag :

$$\partial_G^+ \varphi \leftrightarrow \partial_G \varphi \wedge \neg \partial_{Ag - G} \varphi \quad (24)$$

Operators corresponding to ∂_G where G is a set of agents are found in the literature on *stit*. The definition takes the form $[G \text{ stit}: \varphi] =_{\text{def}} [G]\varphi \wedge \neg\Box\varphi$. Operators corresponding to ∂_G^+ are not encountered in the literature on *stit* but they could be defined in analogous fashion, as

$$[G \text{ stit}: \varphi] \wedge \neg[(Ag-G) \text{ stit}: \varphi]$$

It is straightforward to re-express the axiomatisation of $[G]$ in terms of ∂_G . We note only the following properties for future reference, for all subsets G of Ag :

$$\text{if } \models \varphi \rightarrow \psi \text{ then } \models \neg\Box\psi \rightarrow (\partial_G\varphi \rightarrow \partial_G\psi) \quad (25)$$

$$\models (\partial_G\varphi \wedge \Box\psi) \rightarrow \partial_G(\varphi \wedge \psi) \quad (26)$$

$$\models \partial_G\varphi \rightarrow \partial_G\partial_G\varphi \quad (27)$$

$$\models \neg\partial_G\varphi \rightarrow [G]\neg\partial_G\varphi \quad (28)$$

The following is standard terminology in the *stit* literature.

Definition 2 An expression ψ is *agentive* in G when $\models \psi \leftrightarrow \partial_G\psi$.

The property Eq. 27 together with $\models \partial_G\varphi \rightarrow \varphi$ implies that $\partial_G\varphi$ is ‘agentive in G ’ for any φ .

$\neg\partial_G\varphi$ on the other hand is not agentive in G : $\not\models \neg\partial_G\varphi \rightarrow \partial_G\neg\partial_G\varphi$; the property Eq. 28 is weaker. These are also all features of *dstit*: they are properties of models of the form Eq. 1 and do not depend on the ‘independence of agents’ constraint incorporated in *dstit* models.

For more general forms of *stit*, ψ is said to be agentive in G when ψ is equivalent to $[G \text{ stit}: \psi]$. Belnap and Perloff [4, p39] say that $[G \text{ stit}: [G \text{ stit}: \varphi]]$ is equivalent in *stit* theory to $[G \text{ stit}: \varphi]$ and that $\neg[G \text{ stit}: \varphi]$ is not in general agentive in G . From the proof of their Theorem 17 it seems that the analogue of property Eq. 28 also holds for the version of *stit* employed in that paper.³

Note that $\partial_x^+\varphi \rightarrow \partial_x^+\partial_x^+\varphi$ and $\partial_G^+\varphi \rightarrow \partial_G^+\partial_G^+\varphi$ are *not* valid.⁴ For future reference it is not difficult to derive:

$$\models \partial_G^+\partial_G^+\varphi \leftrightarrow [G]\partial_G^+\varphi \quad (29)$$

from which follows

$$\models (\partial_G^+\varphi \rightarrow \partial_G^+\partial_G^+\varphi) \leftrightarrow (\neg[Ag-G]\varphi \rightarrow [G]\neg[Ag-G]\varphi) \quad (30)$$

³In presenting the features of *stit* in that paper, Belnap and Perloff refer to earlier works, including [2, 3] amongst others, but say [4, p26] that for convenience of exposition the postulates are summarised in a ‘conceptually different but mathematically equivalent’ variant and limited to ideas ‘that we think work equally well for either *stit* or *dstit*’. It is likely that *stit* refers to the ‘achievement *stit*’ [5] but some details are omitted. It is difficult to check every detail of the formal development without reference to these earlier works.

⁴This may seem slightly surprising at first sight but not if we read the $[G]\varphi$ modalities as expressing ‘distributed knowledge’ in group G . $\partial_x^+\varphi$ would then be read as saying that x is the only one who knows that φ . If x is the only one who knows that φ , we would not infer that x is the only one who knows that x is the only one who knows that φ . Or even that x knows that x is the only one who knows that φ .

To see this, note first that $\models \neg[Ag-G]\partial_G^+\varphi$. This is because both $\models [Ag-G][G]\varphi \rightarrow [Ag-G]\varphi$ and $\models [Ag-G]\neg[Ag-G]\varphi \rightarrow \neg[Ag-G]\varphi$. From the definition, $\models \partial_G^+\partial_G^+\varphi \leftrightarrow [G]\partial_G^+\varphi \wedge \neg[Ag-G]\partial_G^+\varphi$, and so $\models \partial_G^+\partial_G^+\varphi \leftrightarrow [G]\partial_G^+\varphi$. Then for Eq. 30, from the definition (and $[G]$ normal) $\models [G]\partial_G^+\varphi \leftrightarrow ([G][G]\varphi \wedge [G]\neg[Ag-G]\varphi)$, and $\models [G][G]\varphi \leftrightarrow [G]\varphi$ because $[G]$ is of type S5.

Herzig and Schwarzentruher [15] identify a number of further properties of group agency operators in *dstit* (the ‘deliberative *stit*’) as well as investigating satisfiability and axiomatisability. Those properties however depend crucially on the ‘independence of agents’ assumption in *dstit* and will not be employed in this paper.

3 Collective agency: Inessential members

Do $\partial_G\varphi$ or $\partial_G^+\varphi$ provide a plausible representation of collective agency? We have ‘superadditivity’ of both ∂_G and ∂_G^+ :

$$\models \partial_G\varphi \rightarrow \partial_H\varphi \quad \text{if } G \subseteq H \quad (31)$$

$$\models \partial_G^+\varphi \rightarrow \partial_H^+\varphi \quad \text{if } G \subseteq H \quad (32)$$

and therefore also

$$\models \partial_G\varphi \rightarrow \partial_{Ag}\varphi \quad (33)$$

$$\models \partial_G^+\varphi \rightarrow \partial_{Ag}^+\varphi \quad (34)$$

Belnap and Perloff [4, p40] observe that this is also a feature of the version of *stit* they employ. Their ‘Fact 14’ states that, given $G \subseteq H \subseteq Ag$, if $[G \text{ stit}: \varphi]$ then $[H \text{ stit}: \varphi]$.

Clearly neither $\partial_G\varphi$ nor $\partial_G^+\varphi$ expresses that it is the set G of agents that is responsible for bringing it about that φ , except in a very weak sense indeed.

We can show:

$$\models \partial_G^+\varphi \wedge \partial_H^+\varphi \rightarrow \perp \quad \text{if } G \cap H = \emptyset \quad (35)$$

So although G and H need not be *unique* they must have some members in common. Notice the special case: $\partial_x^+\varphi \wedge \partial_y^+\varphi \rightarrow \perp$ if $\{x\} \cap \{y\} = \emptyset$, i.e., if $x \neq y$ then $\models \partial_x^+\varphi \wedge \partial_y^+\varphi \rightarrow \perp$. This is something. But still we have $\models \partial_G^+\varphi \rightarrow \partial_H^+\varphi$ for every $G \subseteq H$.

3.1 Minimal sets of agents

∂_G and ∂_G^+ express a very weak kind of collective agency. If the set G of agents collectively brings it about that φ then so, in a very weak sense, does every superset of G ; indeed the set Ag also collectively brings it about that φ . But if our aim is to ascribe notions of responsibility for φ , this is not what we are aiming at. In the table-vase example when a and b collectively tilt the table and break the vase, c has nothing to do with it.

A natural way of looking at it is that the ‘necessity’ condition $[G]\varphi$ is too weak: G can be ‘too big’—it might contain x who contributes nothing to the bringing about of φ : $[G-\{x\}]\varphi$ might also be true for some $x \in G$. Horty [16, footnote 15, p33]

refers to such an x as a ‘free-rider’. We will follow Belnap and Perloff [4] and refer to such an x as an ‘inessential member’ of G . In later sections we will distinguish further between ‘inessential members’ and ‘mere bystanders’.

It seems inescapable to look at the subsets of G in an expression $[G]\varphi$, and insist that for the necessity condition, G should be *minimal* in some sense. There are several possible ways of expressing this requirement. Let us consider the obvious one first.

Definition 3 For every subset G of Ag :

$$[G]^{\min}\varphi =_{\text{def}} [G]\varphi \wedge \neg \bigvee_{H \subset G} [H]\varphi$$

(Ag is finite so this is well formed)

$$\Delta_G^{\min}\varphi =_{\text{def}} [G]^{\min}\varphi \wedge \neg \Box\varphi$$

that is, in full:

$$\Delta_G^{\min}\varphi =_{\text{def}} [G]\varphi \wedge \neg \bigvee_{H \subset G} [H]\varphi \wedge \neg \Box\varphi$$

$\mathcal{M}, \tau \models \Delta_G^{\min}\varphi$ iff (1) $\mathcal{M}, \tau \models [G]\varphi$, (1') there is no proper subset $H \subset G$ such that $\mathcal{M}, \tau \models [H]\varphi$, and (2) there is an alternative (in which G , collectively, acted differently than it did in τ) in which φ is false. The minimality condition (1') could be regarded as part of the ‘necessity’ condition or as part of the counteraction condition. It makes no difference.

Notice that, as defined above, $\models [\emptyset]^{\min}\varphi \leftrightarrow \Box\varphi$, and $\models \Delta_{\emptyset}^{\min}\varphi \leftrightarrow \perp$. For the case $G \neq \emptyset$, by ‘necessity’ Eq. 14, we have $\models \neg[H]\varphi \rightarrow \neg\Box\varphi$ for every $H \subseteq Ag$ and so:

$$\models \Delta_G^{\min}\varphi \leftrightarrow [G]^{\min}\varphi \quad (G \neq \emptyset) \quad (36)$$

Although we are primarily interested here in $\Delta_G^{\min}\varphi$ it is still useful to refer to $[G]^{\min}\varphi$ from time to time and so we will retain it. For the case $G = \{x\}$: $\models \Delta_{\{x\}}^{\min}\varphi \leftrightarrow \partial_x\varphi$.

We can easily derive:

$$\models \Delta_G^{\min}\varphi \leftrightarrow \partial_G\varphi \wedge \neg \bigvee_{H \subset G} \partial_H\varphi \quad (37)$$

Equation 37 corresponds exactly to Belnap and Perloff’s definition of ‘*strictly stit*’ in [4]. What we are calling the ‘minimality condition’ they call the ‘no inessential members’ condition. We will discuss that definition in more detail in Section 3.2 below.

It is very useful to have an alternative characterisation of $[G]^{\min}\varphi$ and hence of $\Delta_G^{\min}\varphi$. We have:

$$\models [G]^{\min}\varphi \leftrightarrow [G]\varphi \wedge \neg \bigvee_{x \in G} [G - \{x\}]\varphi \quad (38)$$

and also

$$\models \Delta_G^{\min}\varphi \leftrightarrow \partial_G\varphi \wedge \neg \bigvee_{x \in G} \partial_{G - \{x\}}\varphi \quad (39)$$

This alternative formulation explains Belnap and Perloff's choice of the term 'no inessential members'. As they say [4, p27], 'strictly *stit*', in our notation $\Delta_G^{\min} \varphi$, expresses that 'the bearers of G , without any outside help, *and with the essential input of each of them*, guarantee that φ '.

What about some of the properties of $\Delta_G^{\min}$?

- Obviously we have:

$$\models \Delta_G^{\min} \varphi \rightarrow \partial_G \varphi \quad (40)$$

- It is no longer the case that if $G \subseteq H$ then $\Delta_G^{\min} \varphi \rightarrow \Delta_H^{\min} \varphi$. Indeed, we have for all non-empty subsets G and H of Ag :

$$\models \Delta_G^{\min} \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \perp \quad \text{if } G \subset H \quad (41)$$

- $\Delta_G^{\min} \varphi \wedge \Delta_H^{\min} \varphi$ can be satisfiable for $G \neq H$.

$$\not\models \Delta_G^{\min} \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \perp \quad (G \not\subseteq H, H \not\subseteq G)$$

Suppose there are three agents a , b , and c pushing against a spring-loaded door and keeping it shut. Suppose any two of them pushing together are strong enough to keep the door shut. Let k represent 'keeping-the-door-shut'. When all three push, $[\{a, b\}]^{\min} k$, $[\{b, c\}]^{\min} k$, and $[\{a, c\}]^{\min} k$ are true. Clearly $[\{a, b, c\}]^{\min} k$ is true but $[\{a, b, c\}]^{\min} k$ is not. Since there is a possible alternative in which $\neg k$ is true, we have $\Delta_{\{a, b, c\}}^{\min} k$. (There must be such an alternative. c is not strong enough to keep the door shut by itself, so $\neg[c]k$. And $\neg[c]k$ implies $\neg\Box k$, i.e., $\Diamond\neg k$.) And likewise for $\Delta_{\{b, c\}}^{\min} k$ and $\Delta_{\{a, c\}}^{\min} k$.

- Furthermore, for all subsets G of Ag

$$\models \partial_G \varphi \leftrightarrow \bigvee_{H \subseteq G} \Delta_H^{\min} \varphi \quad (42)$$

$$\models \partial_{Ag} \varphi \leftrightarrow \bigvee_{H \subseteq Ag} \Delta_H^{\min} \varphi \quad (43)$$

In other words, if φ is brought about by any set of agents, then there exists at least one minimal set H such that $\Delta_H^{\min} \varphi$.

3.2 Inessential members

The following definition is from Belnap and Perloff [4, p40, Definition 15], but rephrased in terms of $\partial_G \varphi$ rather than $[G \textit{ stit}: \varphi]$.

Definition 4 x is *essential* for $\partial_G \varphi$ at τ iff $\tau \models \partial_G \varphi \wedge \neg \partial_{G-\{x\}} \varphi$. x is *inessential* for $\partial_G \varphi$ at τ iff $\tau \models \partial_{G-\{x\}} \varphi$.

Belnap and Perloff say that G ‘*strictly sees*’ to it that φ (at τ) when every member of G is essential for $\partial_G \varphi$ at τ . Let us confirm. When G is non-empty:

$$\begin{aligned}
 \forall x (x \in G \Rightarrow \tau \models \partial_G \varphi \wedge \neg \partial_{G-\{x\}} \varphi) \\
 \text{iff } \tau \models \bigwedge_{x \in G} (\partial_G \varphi \wedge \neg \partial_{G-\{x\}} \varphi) \\
 \text{iff } \tau \models \partial_G \varphi \wedge \bigwedge_{x \in G} \neg \partial_{G-\{x\}} \varphi \\
 \text{iff } \tau \models \partial_G \varphi \wedge \bigwedge_{H \subset G} \neg \partial_H \varphi \\
 \text{iff } \tau \models \Delta_G^{\min} \varphi \quad \text{by Eq. 37}
 \end{aligned}$$

Following Belnap and Perloff, let us define $\text{nim}(G, \varphi)$ (for ‘no inessential members’).

Definition 5 For every subset G of Ag :

$$\text{nim}(G, \varphi) =_{\text{def}} \bigwedge_{H \subset G} \neg \partial_H \varphi$$

Belnap and Perloff’s definition of $\text{nim}(G, \varphi)$ is not expressed as a formula. In our notation their definition [4, p41, Theorem 17] is:

$$\tau \models \text{nim}(G, \varphi) \quad \text{iff} \quad \forall H (\emptyset \subset H \subset G \Rightarrow \tau \models \neg \partial_H \varphi) \quad (44)$$

That is equivalent since $\models \neg \partial_\emptyset \varphi \leftrightarrow \top$. $\text{nim}(G, \varphi)$ could also be defined as:

$$\tau \models \text{nim}(G, \varphi) \quad \text{iff} \quad \forall x (x \in G \Rightarrow \tau \models \neg \partial_{G-\{x\}} \varphi) \quad (45)$$

It is perhaps worth noting that $\models \text{nim}(\{x\}, \varphi) \leftrightarrow \top$. (The empty conjunction is true. The empty disjunction is false.)

Re-stated in terms of $\text{nim}(G, \varphi)$, the ‘*strictly stit*’ $\Delta_G^{\min} \varphi$ is:

$$\models \Delta_G^{\min} \varphi \leftrightarrow \partial_G \varphi \wedge \text{nim}(G, \varphi) \quad (46)$$

Belnap and Perloff now prove (for $[G \text{ stit}: \varphi]$ in place of $\partial_G \varphi$) that the following three expressions are equivalent (their Theorem 17 in our notation):

$$\partial_G \varphi \wedge \text{nim}(G, \varphi), \quad \partial_G (\varphi \wedge \text{nim}(G, \varphi)), \quad \partial_G (\partial_G \varphi \wedge \text{nim}(G, \varphi))$$

They show that the first implies the second, the second implies the third, and the third (trivially) implies the first. The essential point is that the first implies the third, i.e.

$$\models \Delta_G^{\min} \varphi \rightarrow \partial_G \Delta_G^{\min} \varphi \quad (47)$$

This says that $\Delta_G^{\min} \varphi$ is (by Definition 2) ‘agentive in G ’.

Further (their Lemma 20, our notation):

$$\models \neg \partial_H \Delta_G^{\min} \varphi \quad \text{for all } H \subset G \quad (48)$$

Taking these together we have

$$\models \Delta_G^{\min} \varphi \rightarrow \partial_G \Delta_G^{\min} \varphi \wedge \bigwedge_{H \subset G} \neg \partial_H \Delta_G^{\min} \varphi \quad (49)$$

and hence

$$\models \Delta_G^{\min} \varphi \rightarrow \Delta_G^{\min} \Delta_G^{\min} \varphi \quad (50)$$

In other words (by Belnap and Perloff's definition), $\Delta_G^{\min} \varphi$ is *strictly agentive* in G .

Definition 6 An expression ψ is *strictly agentive* in G when $\models \psi \leftrightarrow \Delta_G^{\min} \psi$.

It is quite easy to derive these results in the logic by following, more or less, the semantic arguments given by Belnap and Perloff. Here is a slightly more direct derivation of the main result.

Proposition 1

$$\models \Delta_G^{\min} \varphi \rightarrow \Delta_G^{\min} \Delta_G^{\min} \varphi$$

Proof We need to show

- (i) $\models \Delta_G^{\min} \varphi \rightarrow [G] \Delta_G^{\min} \varphi$
- (ii) $\models \Delta_G^{\min} \varphi \rightarrow \bigwedge_{H \subset G} \neg[H] \Delta_G^{\min} \varphi$
- (iii) $\models \Delta_G^{\min} \varphi \rightarrow \neg \Box \Delta_G^{\min} \varphi$

(iii) is not really necessary as it is already implied by (ii) if we allow the special case $H = \emptyset$, but it is easy. $\models \Delta_G^{\min} \varphi \rightarrow \varphi$ so ($\Box$ is normal) $\models \neg \Box \varphi \rightarrow \neg \Box \Delta_G^{\min} \varphi$ and we have, from Definition 3, $\models \Delta_G^{\min} \varphi \rightarrow \neg \Box \varphi$.

For part (i), note first

$$\begin{aligned} &\models \neg[H] \varphi \rightarrow [H] \neg[H] \varphi && ([H] \text{ S5}) \\ &\models \neg[H] \varphi \rightarrow [G] \neg[H] \varphi && \text{for all } H \subseteq G \quad (\text{'superadditivity' Eq. 15}) \\ &\models \bigwedge_{H \subset G} \neg[H] \varphi \rightarrow \bigwedge_{H \subset G} [G] \neg[H] \varphi \\ &\models \bigwedge_{H \subset G} \neg[H] \varphi \rightarrow [G] \bigwedge_{H \subset G} \neg[H] \varphi && ([G] \text{ normal}) \end{aligned}$$

From the definition, $\models \Delta_G^{\min} \varphi \rightarrow [G] \varphi \wedge \bigwedge_{H \subset G} \neg[H] \varphi \wedge \neg \Box \varphi$ and so, with the above, we have:

$$\begin{aligned} &\models \Delta_G^{\min} \varphi \rightarrow [G] \varphi \wedge [G] \bigwedge_{H \subset G} \neg[H] \varphi \wedge \neg \Box \varphi \\ &\models [G] \varphi \rightarrow [G] [G] \varphi \quad ([G] \text{ is type S5}) \text{ and } \models \neg \Box \varphi \rightarrow [G] \neg \Box \varphi \quad (\Box \text{ is type S5 and 'necessity' Eq. 14}) \text{ so:} \end{aligned}$$

$$\begin{aligned} &\models \Delta_G^{\min} \varphi \rightarrow [G] [G] \varphi \wedge [G] \bigwedge_{H \subset G} \neg[H] \varphi \wedge [G] \neg \Box \varphi \\ &\models \Delta_G^{\min} \varphi \rightarrow [G] \Delta_G^{\min} \varphi && ([G] \text{ normal}) \end{aligned}$$

For part (ii) we can prove the stronger:

$$\models \neg[H] \Delta_G^{\min} \varphi \quad \text{for all } H \subset G$$

This is because:

$$\models [H] \Delta_G^{\min} \varphi \rightarrow [H] [G] \varphi \wedge [H] \bigwedge_{G' \subset G} \neg[G'] \varphi$$

(from the definition of $\Delta_G^{\min}$ and $[H]$ is normal) and so:

$$\models [H] \Delta_G^{\min} \varphi \rightarrow [H] [G] \varphi \wedge [H] \neg[H] \varphi \quad \text{if } H \subset G$$

But $\models [H][G]\varphi \rightarrow [H]\varphi$ (because $\models [G]\varphi \rightarrow \varphi$ and $[H]$ is normal) and $\models [H]\neg[H]\varphi \rightarrow \neg[H]\varphi$. So if $H \subset G$ then $\models [H]\Delta_G^{\min}\varphi \rightarrow [H]\varphi \wedge \neg[H]\varphi$. $\square$

Belnap and Perloff's Lemma 20 says (our notation) that $\models \neg\partial_H\Delta_G^{\min}\varphi$ when $H \subset G$. That follows from $\models \neg[H]\Delta_G^{\min}\varphi$ (part (ii) above) and $\models \partial_H\Delta_G^{\min}\varphi \rightarrow [H]\Delta_G^{\min}\varphi$.

3.3 A stronger form of collective agency

Can we find a stronger form of collective agency, denoted Δ_G^{sole} say, such that

- $\models \Delta_G^{\text{sole}}\varphi \rightarrow \Delta_G^{\min}\varphi$
- $\models \Delta_G^{\text{sole}}\varphi \wedge \Delta_H^{\text{sole}}\varphi \rightarrow \perp$ if $G \neq H$

$\Delta_G^{\text{sole}}\varphi$ would be to $\Delta_G^{\min}\varphi$ what $\partial_x^+\varphi$ is to $\partial_x\varphi$.

One possibility, for example, is to try:

$$\Delta'_G\varphi =_{\text{def}} \Delta_G^{\min}\varphi \wedge \neg[Ag-G]\varphi \quad (51)$$

which is also (as it turns out) $\models \Delta'_G\varphi \leftrightarrow \partial_G^+\varphi \wedge \bigwedge_{x \in G} \neg[G-\{x}]\varphi$, and hence

$$\models \Delta'_G\varphi \leftrightarrow \partial_G^+\varphi \wedge \text{nim}(G, \varphi) \quad (52)$$

Δ'_G has some reasonable formal properties but does not seem to correspond to any particularly meaningful notion of collective agency.

The counteraction condition employed in the definition of $\Delta_G^{\min}\varphi$ is:

$$\bigwedge_{x \in G} \neg[G-\{x}]\varphi$$

The *strongest* counteraction condition that can be defined of similar form is:

$$\bigwedge_{x \in G} \neg[Ag-\{x}]\varphi$$

It is stronger because $\models [G-\{x}]\varphi \rightarrow [Ag-\{x}]\varphi$ ('superadditivity' Eq. 15).

This suggests the following construction as a plausible candidate.

Definition 7 For every subset G of Ag :

$$\Delta_G^{\text{sole}}\varphi =_{\text{def}} [G]\varphi \wedge \bigwedge_{x \in G} \neg[Ag-\{x}]\varphi$$

$\Delta_G^{\text{sole}}\varphi$ is satisfied at τ in a model $\mathcal{M}$ when:

- (1) (necessity) $\mathcal{M}, \tau \models [G]\varphi$, that is, the joint actions of G guarantee that φ ; and
- (2) (counteraction) for every $x \in G$, there is an alternative in which φ is false and in which *all* other agents $Ag-\{x\}$, not only those in G , act in the same way as they do in τ .

Since $\models \neg[Ag-\{x}]\varphi \rightarrow \neg\Box\varphi$ for every $x \in Ag$ ('necessity' Eq. 14), the definition can be expressed equivalently as:

$$\Delta_G^{\text{sole}}\varphi \leftrightarrow \partial_G\varphi \wedge \bigwedge_{x \in G} \neg\partial_{Ag-\{x}}\varphi \quad (53)$$

Furthermore, since $\models \neg[Ag-\{x\}]\varphi \rightarrow \neg[G-\{x\}]\varphi$ ('superadditivity' Eq. 15), we have

$$\models \Delta_G^{\text{sole}}\varphi \leftrightarrow [G]\varphi \wedge \bigwedge_{x \in G} \neg[G-\{x}]\varphi \wedge \bigwedge_{x \in G} \neg[Ag-\{x}]\varphi \quad (54)$$

and hence

$$\models \Delta_G^{\text{sole}}\varphi \leftrightarrow \Delta_G^{\min}\varphi \wedge \bigwedge_{x \in G} \neg\partial_{Ag-\{x}\}\varphi$$

Some properties of Δ_G^{sole} are immediate.

- For the singleton case $G = \{x\}$, $\Delta_{\{x\}}^{\text{sole}}\varphi = [x]\varphi \wedge \neg[Ag-\{x}]\varphi$, and so

$$\models \Delta_{\{x\}}^{\text{sole}}\varphi \leftrightarrow \partial_x^+\varphi$$

- For any $G \subseteq Ag$:

$$\models \Delta_G^{\text{sole}}\varphi \rightarrow \Delta_G^{\min}\varphi \quad (55)$$

(and also $\models \Delta_G^{\text{sole}}\varphi \rightarrow \Delta'_G\varphi$).

- For any subsets G and H of Ag :

$$\models \Delta_G^{\text{sole}}\varphi \wedge \Delta_H^{\text{sole}}\varphi \rightarrow \perp \quad \text{if } G \neq H \quad (56)$$

In the pushing-the-door example where any two agents pushing together are strong enough to keep the door shut, we have $\Delta_G^{\min}k$ true for any pair G of distinct agents from Ag (at least two agents are required to keep the door shut, and if $\Delta_G^{\min}k$ is true then $\Delta_H^{\min}k$ is not true for any $H \supset G$). $\Delta_G^{\text{sole}}k$ on the other hand is not true for any particular pair $G \subseteq Ag$. This seems quite natural and satisfactory.

The definition of $\Delta_G^{\text{sole}}\varphi$ may seem contrived: the only motivation offered was an appeal to the strongest counteraction condition of a particular form. Here is an alternative characterisation.

Lemma 1 Let G and H be (non-empty) subsets of Ag .

$$\models \Delta_G^{\min}\varphi \rightarrow (\bigvee_{G \neq H} \Delta_H^{\min}\varphi \leftrightarrow \bigvee_{x \in G} [Ag-\{x}]\varphi)$$

Proof For the first half (left-to-right): suppose $\tau \models \Delta_G^{\min}\varphi \wedge \Delta_H^{\min}\varphi$. If $G \neq H$ then $G \not\subseteq H$ (otherwise $\Delta_G^{\min}\varphi$ and $\Delta_H^{\min}\varphi$ could not both be true at τ). So there must be some $x \in G$ such that $x \notin H$, i.e., such that $H \subseteq Ag-\{x\}$. In that case $\tau \models [H]\varphi \rightarrow [Ag-\{x}]\varphi$. Since $\models \Delta_H^{\min}\varphi \rightarrow [H]\varphi$, if $x \in G$ and $x \notin H$, then $\models \Delta_G^{\min}\varphi \wedge \Delta_H^{\min}\varphi \rightarrow [Ag-\{x}]\varphi$. It follows (propositional logic) that if $G \neq H$, then $\models \Delta_G^{\min}\varphi \wedge \Delta_H^{\min}\varphi \rightarrow \bigvee_{x \in G} [Ag-\{x}]\varphi$.

For the other half, we prove $\models (\bigvee_{x \in G} [Ag-\{x}]\varphi \rightarrow \bigvee_{G \neq H} \Delta_H^{\min}\varphi)$. Suppose $\tau \models [Ag-\{x}]\varphi$ for some $x \in G$. Then by Eq. 42, $\tau \models \Delta_H^{\min}\varphi$ for some $H \subseteq Ag-\{x\}$. Since $x \in G$ and $x \notin H$, there exists $G \neq H$ such that $\Delta_H^{\min}\varphi$. $\square$

Proposition 2 Let G and H be (non-empty) subsets of Ag .

$$\models \Delta_G^{\text{sole}}\varphi \leftrightarrow \Delta_G^{\min}\varphi \wedge \neg \bigvee_{H \neq G} \Delta_H^{\min}\varphi$$

☐

In other words, $\Delta_G^{\text{sole}} \varphi$ holds when G is the only set (if it exists) such that $\Delta_G^{\text{min}} \varphi$. By analogy with $\partial_G^+ \varphi$, properties Eqs. 29 and 30, we can also derive the following.

Proposition 3

$$\begin{aligned} & \models \Delta_G^{\text{sole}} \Delta_G^{\text{sole}} \varphi \leftrightarrow [G] \Delta_G^{\text{sole}} \varphi \\ & \models (\Delta_G^{\text{sole}} \varphi \rightarrow \Delta_G^{\text{sole}} \Delta_G^{\text{sole}} \varphi) \leftrightarrow (\bigwedge_{x \in G} \neg[Ag-\{x\}] \varphi \rightarrow [G] \bigwedge_{x \in G} \neg[Ag-\{x\}] \varphi) \end{aligned}$$

Proof As in the derivation of the corresponding properties Eqs. 29 and 30 of $\partial_G^+ \varphi$ $\square$

Finally, we noted earlier that

$$\models \partial_G^+ \varphi \leftrightarrow \partial_G \varphi \wedge \neg \partial_{Ag-G} \varphi$$

Perhaps the construction

$$\Delta_G^{\min} \varphi \wedge \neg \Delta_{Ag-G}^{\min} \varphi \quad (57)$$

yields something interesting? It does not. We have:

$$\models \Delta_G^{\text{sole}} \varphi \rightarrow \Delta_G^{\text{min}} \varphi \wedge \neg \Delta_{Ag-G}^{\text{min}} \varphi \quad (58)$$

but this is not an equivalence: in the pushing-the-door-shut example we have $\Delta_{\{a,b\}}^{\min} k$ and $\neg \Delta_{\{c\}}^{\min} k$ but not $\Delta_{\{a,b\}}^{\text{sole}} k$. The construction Eq. 57 seems to have no particular significance. Similarly, we have

$$\models \Delta_G^{\text{sole}} \varphi \rightarrow \Delta'_G \varphi \wedge \neg \Delta'_{Ag-G} \varphi$$

But this is not an equivalence. It says only that $\models \Delta_G^{\text{sole}} \varphi \rightarrow \Delta'_G \varphi$.

3.4 Summary

We have identified a range of possible forms of collective agency with implications between them as summarised in the following diagram:

Of these, it is $\Delta_G^{\min}$ and Δ_G^{sole} that are deserving of attention. They are the analogues of ∂_x and ∂_x^+ , respectively, in that $\Delta_G^{\min}\varphi$ allows for the possibility

that several distinct sets G of agents bring about that φ , while $\Delta_G^{\text{sole}}\varphi$ expresses that any such set G , if it exists, is unique. Both imply that the set G of agents is minimal: $\models \Delta_G^{\text{min}}\varphi \rightarrow [G]^{\text{min}}\varphi$, and $\models \Delta_G^{\text{sole}}\varphi \rightarrow \Delta_G^{\text{min}}\varphi$. Δ_G^{min} corresponds to what Belnap and Perloff call 'strictly *stit*'. $\Delta_G^{\text{min}}\varphi$ ('strictly *stit*') implies $\partial_G\varphi$ ('*stit*') but, obviously, not the converse. $\Delta_G^{\text{min}}\varphi$ can also be defined in terms of 'no inessential members' as $\partial_G\varphi \wedge \text{nim}(G, \varphi)$. Δ'_G has a natural technical definition as $\partial_G^+\varphi \wedge \text{nim}(G, \varphi)$ and some reasonable formal properties but does not appear to express any meaningful notion of collective agency. For the special case of singleton sets, $\models \Delta_{\{x\}}^{\text{min}}\varphi \leftrightarrow \partial_x\varphi$ and $\models \Delta_{\{x\}}^{\text{sole}}\varphi \leftrightarrow \partial_x^+\varphi$.

4 Contributors and bystanders

4.1 Contributors: $\Delta_G^{\text{max}}\varphi$

Consider a version of the pushing-the-door example in which agents a and b are strong enough to keep the door shut (k) when both push and c is strong enough on its own. Suppose a , b and c are all pushing. Then $\Delta_{\{a,b\}}^{\text{min}}k$ and $\Delta_{\{c\}}^{\text{min}}k$ are true. Suppose that d is an agent whose actions do not affect the door, directly or indirectly. There is a sense in which it is $\{a, b, c\}$ but not $\{a, b, c, d\}$ who collectively bring about that the door is shut. In what sense? In the sense that $\{a, b, c\}$ is the union of all sets G such that $\Delta_G^{\text{min}}k$. d on the other hand does not contribute to the bringing about of k : d does not belong to any of the sets G such that $\Delta_G^{\text{min}}k$.

Definition 8 For every subset G of Ag :

$$\tau \models \Delta_G^{\text{max}}\varphi \quad \text{iff} \quad \tau \models \partial_{\text{Ag}}\varphi \quad \& \quad G = \bigcup_H (\tau \models \Delta_H^{\text{min}}\varphi)$$

G is the set of *contributors* to φ at τ .

The condition $\tau \models \partial_{\text{Ag}}\varphi$ is to guard against the possibility that there are no sets H such that $\tau \models \Delta_H^{\text{min}}\varphi$. $\tau \models \partial_{\text{Ag}}\varphi$ is equivalent to saying that $\tau \models \Delta_H^{\text{min}}\varphi$ for some $H \subseteq \text{Ag}$. In the definition it could be replaced equivalently by $G \neq \emptyset$.

The following properties follow more or less immediately.

Proposition 4

- (i) $\models \Delta_G^{\text{max}}\varphi \rightarrow \partial_G\varphi$
- (ii) $\models \Delta_G^{\text{max}}\varphi \wedge \Delta_H^{\text{max}}\varphi \rightarrow \perp \quad \text{for } G \neq H$
- (iii) if $\tau \models \Delta_G^{\text{max}}\varphi \wedge \Delta_H^{\text{min}}\varphi$ then $H \subseteq G$ ($\models \Delta_G^{\text{max}}\varphi \rightarrow \neg \Delta_H^{\text{min}}\varphi$ if $H \not\subseteq G$)
- (iv) $\models \Delta_H^{\text{min}}\varphi \rightarrow \bigvee_{H \subseteq G} \Delta_G^{\text{max}}\varphi \quad (\text{'there exists } G \supseteq H \text{ such that } \Delta_G^{\text{max}}\varphi')$
- (v) if $\tau \models \Delta_G^{\text{max}}\varphi$ & $x \in G$ then $\exists H (\tau \models \Delta_H^{\text{min}}\varphi \text{ & } x \in H)$

(vi) *From the two previous observations*

$$\exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H) \quad \text{iff} \quad \exists G (\tau \models \Delta_G^{\max} \varphi \ \& \ H \subseteq G)$$

Proposition 5

$$\models \Delta_G^{\text{sole}} \varphi \leftrightarrow \Delta_G^{\min} \varphi \wedge \Delta_G^{\max} \varphi$$

Proof We know Eq. 55 that $\models \Delta_G^{\text{sole}} \varphi \rightarrow \Delta_G^{\min} \varphi$. If $\tau \models \Delta_G^{\text{sole}} \varphi$ then $\{H \mid \tau \models \Delta_H^{\min} \varphi\} = \{G\}$, and so $\tau \models \Delta_G^{\max} \varphi$. For the other half: if $\tau \models \Delta_G^{\min} \varphi$ and $\tau \not\models \Delta_G^{\text{sole}} \varphi$ then there must be some $H \neq G$ such that $\tau \models \Delta_H^{\min} \varphi$. It cannot be that $G \subset H$ or that $H \subset G$, so $\{G' \mid \tau \models \Delta_{G'}^{\min} \varphi\} \supseteq \{G, H\}$. $G \cup H \neq G$, and so $\tau \not\models \Delta_G^{\max} \varphi$. $\square$

For the case of singleton sets, $\models \Delta_{\{x\}}^{\max} \varphi \leftrightarrow \partial_x^+ \varphi$.

4.2 Bystanders

We will now relate the definition of $\Delta_G^{\max}$ and ‘contributors’ to some suggestions by Belnap and Perloff [4]. The following are their definitions [4, p40, Definition 16] (but with ∂_G in place of their *stit*).

Definition 9 x is *essential* [*inessential*, *a mere bystander*, *not a mere bystander*] for φ at τ iff $\exists G' \tau \models \partial_{G'} \varphi$ and for every [*not all*, *no*, *some*] G such that $\tau \models \partial_G \varphi$, x is essential for $\partial_G \varphi$.

(There is a typographical error in Belnap and Perloff’s statement of the definition but it is clear from context what was intended.)

As in Belnap and Perloff, our main interest is in (mere) bystanders, and in particular the case where $\partial_G \varphi$ is true and G contains no mere bystanders for φ . But first, notice that ‘essential for φ ’ as defined above is not the same as, and is much stronger than, the notion of ‘essential for $\partial_G \varphi$ ’ in terms of which $\text{nim}(G, \varphi)$ was earlier defined. ‘essential for $\partial_G \varphi$ ’ (Definition 4) is $\partial_G \varphi \wedge \neg \partial_{G-\{x\}} \varphi$. The terminology is unfortunate and potentially confusing. One concept (Definition 4) is group dependent. The other, stronger (Definition 9) is group independent.

Definition 9 partitions the set Ag of agents, when $\tau \models \partial_{Ag} \varphi$, in two different ways: ‘essential for φ ’ and ‘inessential for φ ’ on the one hand, and ‘mere bystanders for φ ’ and ‘not mere bystanders for φ ’ on the other. ‘essential for φ ’ implies ‘not mere bystander for φ ’; ‘inessential for φ ’ does not imply ‘mere bystander for φ ’.

The group independent concept ‘essential for φ ’ is extremely strong indeed. Here is an equivalent formulation which is easier to work with.

Proposition 6 x is *essential* for φ at τ iff:

$$\tau \models \partial_{Ag} \varphi \ \& \ \forall H (\tau \models \Delta_H^{\min} \varphi \Rightarrow x \in H)$$

Proof From Definition 9:

x is essential for φ at τ

iff $\exists G' \tau \models \partial_{G'} \varphi \ \& \ \forall G (\tau \models \partial_G \varphi \Rightarrow x \text{ is essential for } \partial_G \varphi \text{ at } \tau)$

iff $\exists G' \tau \models \partial_{G'} \varphi \ \& \ \forall G (\tau \models \partial_G \varphi \Rightarrow \tau \models \neg \partial_{G-\{x\}} \varphi)$

iff $\tau \models \partial_{A_g} \varphi \ \& \ \forall H (\tau \models \Delta_H^{\min} \varphi \Rightarrow x \in H)$

The last step of the derivation follows because $\exists G \tau \models (\partial_G \varphi \wedge \partial_{G-\{x\}} \varphi)$ iff $\exists G \tau \models \partial_{G-\{x\}} \varphi$ ('superadditivity', schema Eq. 31), and $\exists G \tau \models \partial_{G-\{x\}} \varphi$ iff $\exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \notin H)$. For suppose that $\tau \models \partial_{G-\{x\}} \varphi$ for some G . Then $\tau \models \Delta_H^{\min} \varphi$ for some $H \subseteq G-\{x\}$, and since $H \subseteq G-\{x\}$ then $x \notin H$. For the other direction, suppose $\tau \models \Delta_H^{\min} \varphi$ for some H such that $x \notin H$. If $x \notin H$ then $H-\{x\} = H$, and so $\tau \models \partial_{H-\{x\}} \varphi$ and $\exists G \tau \models \partial_{G-\{x\}} \varphi$ as required. $\square$

We see that x is essential for φ at τ when x is a member of every set H such that $\tau \models \Delta_H^{\min} \varphi$. That is extremely strong. We will comment briefly in Section 4.5 on what it might signify.

Let us turn to mere bystanders. In discussion of directions for further work, Belnap and Perloff [4, p46] say that G contains no mere bystander for φ when every x in G is such that $[H \text{ stit}: \varphi]$ for some $x \in H$. There seems to be a mistake here, though it might be just a typographical slip: from the definitions it seems that *strictly stit* $[H \text{ sstit}: \varphi]$ was intended rather than plain *stit*, i.e., in our notation $\Delta_H^{\min} \varphi$ rather than $\partial_H \varphi$. Let us confirm that.

Proposition 7 x is not a mere bystander for φ at τ iff:

$$\exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H)$$

Proof From Definition 9:

x is not a mere bystander for φ at τ

iff $\exists G' \tau \models \partial_{G'} \varphi \ \& \ \exists G (\tau \models \partial_G \varphi \ \& \ x \text{ is essential for } \partial_G \varphi \text{ at } \tau)$

iff $\exists G (\tau \models \partial_G \varphi \ \& \ x \text{ is essential for } \partial_G \varphi \text{ at } \tau)$

iff $\exists G \tau \models (\partial_G \varphi \wedge \neg \partial_{G-\{x\}} \varphi)$

iff $\exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H)$

For the last step: suppose that $\tau \models \partial_G \varphi$ and $\tau \models \neg \partial_{G-\{x\}} \varphi$ for some G . Then $\tau \models \Delta_H^{\min} \varphi$ for some $H \subseteq G$. If $x \notin H$ then $H = H-\{x\}$ and we have $\tau \models \partial_{H-\{x\}} \varphi$. Since $H-\{x\} \subseteq G-\{x\}$, it follows that $\tau \models \partial_{G-\{x\}} \varphi$, which contradicts the assumption. For the other direction, suppose $\tau \models \Delta_H^{\min} \varphi$. If $x \in H$ then $\models \Delta_H^{\min} \varphi \Rightarrow (\partial_H \varphi \wedge \neg \partial_{H-\{x\}} \varphi)$. $\square$

Belnap and Perloff refer to $nmb(G, \varphi)$ 'contains no mere bystander'. Let us define it as follows.

Definition 10 For every subset G of Ag :

$$\begin{aligned}\tau \models nmb(G, \varphi) \quad \text{iff} \quad \forall x (x \in G \Rightarrow \exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H)) \\ \tau \models nmb(G, \varphi) \quad \text{iff} \quad G \subseteq \bigcup_H (\tau \models \Delta_H^{\min} \varphi)\end{aligned}$$

Expressed as a formula: $nmb(G, \varphi) =_{\text{def}} \bigwedge_{x \in G} \bigvee_{x \in H} \Delta_H^{\min} \varphi$.

Note that $nmb(G, \varphi)$ is always true for $G = \emptyset$. (The empty conjunction is true.) On the other hand, if $G \neq \emptyset$ then $\models nmb(G, \varphi) \rightarrow \partial_{Ag} \varphi$.

‘Evidently’, according to Belnap and Perloff [4, p46], $nim(G, \varphi)$ implies $nmb(G, \varphi)$. Informally, that is true: from Definition 9 ‘ x is essential for φ ’ implies ‘ x is not a mere bystander for φ ’. But that is not the validity of $nim(G, \varphi) \rightarrow nmb(G, \varphi)$ because $nim(G, \varphi)$ is defined in terms of the group dependent ‘ x is essential for $\partial_G \varphi$ ’ as in Definition 5 not in terms of the group independent ‘ x is essential for φ ’. However, the basic intuition is correct, with a minor adjustment. From Eq. 46 we have:

$$\tau \models \partial_G \varphi \wedge nim(G, \varphi) \Rightarrow \tau \models \Delta_G^{\min} \varphi$$

Furthermore (trivially)

$$\tau \models \Delta_G^{\min} \varphi \Rightarrow \forall x (x \in G \Rightarrow \exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H))$$

which is just $\tau \models \Delta_G^{\min} \varphi \Rightarrow \tau \models nmb(G, \varphi)$. So together we have:

$$\models \partial_G \varphi \rightarrow (nim(G, \varphi) \rightarrow nmb(G, \varphi)) \quad (59)$$

Now Belnap and Perloff [4, p46]: ‘There is the statement $omb(G, \varphi)$ that outside of G there are only mere bystanders for φ ’. They go on to suggest that $\Delta_G^{\min} \varphi \wedge omb(G, \varphi)$ might then express that ‘ G is the one and only joint agent for φ ’. $\Delta_G^{\min} \varphi \wedge omb(G, \varphi)$ is equivalent to:

$$\partial_G \varphi \wedge nim(G, \varphi) \wedge omb(G, \varphi) \quad (60)$$

And indeed that does turn out to express *one sense* in which G is the one and only joint agent for φ . We will look at it presently.

First, the more obvious construction to look at is the following:

$$\partial_G \varphi \wedge nmb(G, \varphi) \wedge omb(G, \varphi) \quad (61)$$

That says $\partial_G \varphi$ is true, there are no mere bystanders for φ in G , and outside G there are only mere bystanders for G : in other words, that G is precisely the set of agents who are not mere bystanders for φ . Belnap and Perloff’s suggestion Eq. 60 is different. It has $nim(G, \varphi)$ (no inessential members, in the group dependent sense) in place of $nmb(G, \varphi)$ (no mere bystanders, group independent).

It remains to define $omb(G, \varphi)$, ‘outside G there are only mere bystanders for φ ’:

$$\begin{aligned}\tau \models omb(G, \varphi) \quad \text{iff} \quad \forall x (x \notin G \Rightarrow x \text{ is a mere bystander for } \varphi \text{ at } \tau) \\ \tau \models omb(G, \varphi) \quad \text{iff} \quad \forall x (x \text{ is not a mere bystander for } \varphi \text{ at } \tau \Rightarrow x \in G)\end{aligned}$$

That is

$$\tau \models omb(G, \varphi) \quad \text{iff} \quad \forall x (\exists H (\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H) \Rightarrow x \in G) \quad (62)$$

or equivalently

$$\tau \models omb(G, \varphi) \quad \text{iff} \quad \forall x \forall H ((\tau \models \Delta_H^{\min} \varphi \ \& \ x \in H) \Rightarrow x \in G)$$

$$\tau \models omb(G, \varphi) \quad \text{iff} \quad \forall H (\tau \models \Delta_H^{\min} \varphi \Rightarrow \forall x (x \in H \Rightarrow x \in G))$$

We thus take the following definition.

Definition 11 For every subset G of Ag :

$$\tau \models omb(G, \varphi) \quad \text{iff} \quad \forall H (\tau \models \Delta_H^{\min} \varphi \Rightarrow H \subseteq G)$$

$$\tau \models omb(G, \varphi) \quad \text{iff} \quad \bigcup_H (\tau \models \Delta_H^{\min} \varphi) \subseteq G$$

Expressed as a formula, e.g. via Eq. 62: $omb(G, \varphi) =_{\text{def}} \bigwedge_{x \notin G} \neg \bigvee_{x \in H} \Delta_H^{\min} \varphi$.

Note that if there is no H such that $\tau \models \Delta_H^{\min} \varphi$ then $\tau \models omb(G, \varphi)$ for all G .

Consider now the construction Eq. 61 above: $\partial_G \varphi \wedge nmb(G, \varphi) \wedge omb(G, \varphi)$. This expression can be simplified slightly. As shown next, it is equivalent to $\partial_{Ag} \varphi \wedge nmb(G, \varphi) \wedge omb(G, \varphi)$.

Proposition 8

$$\models \partial_{Ag} \varphi \wedge omb(G, \varphi) \rightarrow \partial_G \varphi$$

Proof Suppose $\tau \models \partial_{Ag} \varphi \wedge omb(G, \varphi)$. $\tau \models \partial_{Ag} \varphi$ implies there exists H such that $\tau \models \Delta_H^{\min} \varphi$. By Definition 11, $\tau \models omb(G, \varphi)$ then implies $H \subseteq G$. $H \subseteq G$ and $\tau \models \partial_H \varphi$ implies $\tau \models \partial_G \varphi$ ('superadditivity', schema Eq. 31). $\square$

Now we can see what construction Eq. 61 expresses. It is $\Delta_G^{\max} \varphi$.

Proposition 9

$$\models \Delta_G^{\max} \varphi \leftrightarrow \partial_G \varphi \wedge nmb(G, \varphi) \wedge omb(G, \varphi)$$

$$\models \Delta_G^{\max} \varphi \leftrightarrow \partial_{Ag} \varphi \wedge nmb(G, \varphi) \wedge omb(G, \varphi)$$

Proof The first is Definition 8 expressed in terms of nmb and omb . The second follows by Proposition 8. $\square$

Corollary 1 $\Delta_G^{\max} \varphi$ can be expressed as a formula, as follows:

$$\Delta_G^{\max} \varphi =_{\text{def}} \partial_{Ag} \varphi \wedge \bigwedge_{x \in G} \bigvee_{x \in H} \Delta_H^{\min} \varphi \wedge \bigwedge_{x \notin G} \neg \bigvee_{x \in H} \Delta_H^{\min} \varphi$$

What we call the *contributors* to φ at τ , that is, the set G such that $\tau \models \Delta_G^{\max} \varphi$, is exactly what Belnap and Perloff call the set of *not mere bystanders* for φ at τ .

Finally, what of Belnap and Perloff's conjecture Eq. 60 that $\Delta_G^{\min} \varphi \wedge omb(G, \varphi)$ expresses that ' G is the one and only joint agent for φ '? That turns out to be $\Delta_G^{\text{sole}} \varphi$.

Proposition 10

$$\models \Delta_G^{\min} \varphi \wedge omb(G, \varphi) \leftrightarrow \Delta_G^{\text{sole}} \varphi$$

Proof

$$\begin{aligned} \models \Delta_G^{\min} \varphi \wedge omb(G, \varphi) &\leftrightarrow \partial_G \varphi \wedge nim(G, \varphi) \wedge omb(G, \varphi) \\ &\leftrightarrow \partial_G \varphi \wedge nim(G, \varphi) \wedge nmb(G, \varphi) \wedge omb(G, \varphi) \\ &\quad \text{by Eq. 59} \\ &\leftrightarrow \Delta_G^{\min} \varphi \wedge \Delta_G^{\max} \varphi \quad \text{by Proposition 9} \\ &\leftrightarrow \Delta_G^{\text{sole}} \varphi \quad \text{by Proposition 5} \end{aligned}$$

□

4.3 Summary

For any $\tau \models \partial_{Ag} \varphi$ the set of agents Ag is partitioned into 'essential for φ '/'inessential for φ ' and 'contributors to φ '/'mere bystanders for φ ' as depicted in the following diagram.

The notation in the diagram is intended to be mnemonic. More exactly:

$$\begin{aligned} \tau \models \Delta_G^{\max} \varphi \quad \text{iff} \quad \partial_{Ag} \varphi \quad \& \quad G = \bigcup_H (\tau \models \Delta_H^{\min} \varphi) \\ \models \Delta_G^{\text{sole}} \varphi &\leftrightarrow \Delta_G^{\min} \varphi \wedge \Delta_G^{\max} \varphi \end{aligned}$$

Given φ and τ , the set G such that $\tau \models \Delta_G^{\max} \varphi$ is unique. G is the set of 'contributors' to φ : they are the agents who are not mere bystanders for φ . If $\tau \models \partial_{Ag} \varphi$, i.e., $\exists G' \tau \models \partial_{G'} \varphi$, then $\Delta_G^{\max} \varphi$ is always true at τ for some non-empty set G . $\Delta_G^{\text{sole}} \varphi$ is stronger. $\Delta_G^{\text{sole}} \varphi$ implies $\Delta_G^{\max} \varphi$ but not the other way round. It may be that there is no G such that $\tau \models \Delta_G^{\text{sole}} \varphi$.

Consider a version of the pushing-the-door-shut example where a , b and c are all pushing and any two of them are strong enough to keep the door shut (k). Then $\Delta_{\{a,b\}}^{\min} k$, $\Delta_{\{b,c\}}^{\min} k$, $\Delta_{\{a,c\}}^{\min} k$ and $\Delta_{\{a,b,c\}}^{\max} k$ are true, but $\Delta_G^{\text{sole}} k$ is not true for any G .

If $\tau \models \Delta_G^{\text{sole}} \varphi$ for some G then $\tau \models \Delta_G^{\text{max}} \varphi$ and $\tau \models \Delta_G^{\text{min}} \varphi$: in that case the contributors to φ are all essential for φ , and the agents who are not essential for φ are the mere bystanders for φ . The essential/inessential and contributors/mere bystanders partitions then coincide.

Example (Door) Here is another variant of the pushing-the-door example, for illustration. Suppose there are agents a, b, c , and d , and a spring-loaded door which can be pushed shut (k) or allowed to spring open ($\neg k$). a and b together and c by itself are strong enough to push the door shut, but only if d does not push (from the other side). If d pushes (from the other side) then a, b and c must all push to keep the door shut otherwise it opens.

There are four situations in which the door is kept shut (k). Situations in which the door is open ($\neg k$) are not shown in the following table.

<i>pushers</i>	$\Delta_G^{\text{min}} k$	$\Delta_G^{\text{max}} k$	<i>essential</i>	<i>bystanders</i>
(a, b, c, d)	$\{a, b, c\}$	$\{a, b, c\}$	$\{a, b, c\}$	$\{d\}$
$(a, b, c, \cdot)$	$\{a, b, d\}, \{c, d\}, \{a, b, c\}$	$\{a, b, c, d\}$	$\{\}$	$\{\}$
$(a, b, \cdot, \cdot)$	$\{a, b, d\}$	$\{a, b, d\}$	$\{a, b, d\}$	$\{c\}$
$(\cdot, \cdot, c, \cdot)$	$\{c, d\}$	$\{c, d\}$	$\{c, d\}$	$\{a, b\}$

In the first row, a, b , and c all push and so the door remains shut even though d also pushes from its side. d is not a contributor to k and is a 'mere bystander'. In the second row, pushing by $\{a, b, c\}$ is sufficient to guarantee the door is shut, irrespective of whether d pushes or not. Pushing by $\{a, b\}$ or by $\{c\}$ is also sufficient to keep the door shut—as long as d does not push: d is a contributor to k (not a mere bystander). In the last two rows d does not push. There $\{a, b\}$ and $\{c\}$ by themselves guarantee the door remains shut, given that d does not push. d is a contributor (not a mere bystander). The table also shows the agents 'essential for k ' in each case, in Belnap and Perloff's strong group-independent sense.

4.4 Is $\Delta_G^{\text{max}} \varphi$ agentive?

A standard question in writings on *stit* is whether an expression of interest is 'agentive'. Is $\Delta_G^{\text{max}} \varphi$ agentive in any of the following senses?

- $\models \Delta_G^{\text{max}} \varphi \rightarrow \partial_G \Delta_G^{\text{max}} \varphi$?
- $\models \Delta_G^{\text{max}} \varphi \rightarrow \Delta_G^{\text{min}} \Delta_G^{\text{max}} \varphi$?
- $\models \Delta_G^{\text{max}} \varphi \rightarrow \Delta_G^{\text{max}} \Delta_G^{\text{max}} \varphi$?

No: even the first of these is false.

$$\not\models \Delta_G^{\text{max}} \varphi \rightarrow [G] \Delta_G^{\text{max}} \varphi$$

For consider again the simple version of the pushing-the-door example, without d , where a and b together and c by itself are strong enough to keep the door shut (k). Suppose a and b are pushing and c is not. Here $\Delta_{\{a,b\}}^{\text{min}} k$ and $\Delta_{\{a,b\}}^{\text{max}} k$. c is a bystander

for k . But in the $\{a, b\}$ -similar alternative in which a, b and c all push, $\Delta_{\{a,b,c\}}^{\max} k$. Here $\Delta_{\{a,b\}}^{\max} k$ is false.

And that is surely to be expected, for whether $\Delta_G^{\max} \varphi$ is true will depend not only on the actions of the members of G but also potentially on the actions of agents outside G , who might perhaps have contributed to bringing about that φ had they acted otherwise. (They could not have contributed to bringing about that $\neg\varphi$ by acting otherwise, because then $[G]\varphi$ would be false.)

4.5 A comment about ‘essential for φ ’

As defined by Belnap and Perloff, the group-independent notion x is essential for φ at τ is equivalent to saying (Proposition 6) that x belongs to all sets H such that $\tau \models \Delta_H^{\min} \varphi$. That is very much stronger than merely being a ‘contributor’ (not mere bystander). Having introduced it, nothing much is made of it thereafter by Belnap and Perloff. What might it represent? Intuitively,⁵ it can be seen as saying that, given the actions of the other agents, the agent x had the power of unilaterally avoiding that φ was brought about. x is not just involved in the bringing about of φ ; with the actions of the other agents fixed, x ’s action is necessary to guarantee that φ .

How might this reading be expressed? It turns out to be simple.

Proposition 11 x is essential for φ at τ iff $\tau \models \partial_{Ag} \varphi \wedge \neg[Ag - \{x\}]\varphi$.

Proof First, for the degenerate case $Ag = \{x\}$, x is essential for φ at τ iff $\tau \models \partial_{\{x\}} \varphi$. By definition, $\tau \models \partial_{\{x\}} \varphi$ implies $\tau \models \neg\Box\varphi$, and $\tau \models \Box\varphi \leftrightarrow [\emptyset]\varphi$. For the general case $Ag \neq \{x\}$, it is enough to show that if $\tau \models \partial_{Ag} \varphi$, then x is essential for φ at τ iff $\tau \models \neg[Ag - \{x\}]\varphi$.

Suppose $\tau \models \partial_{Ag} \varphi$ and x is not essential for φ at τ . Then there is some $H \subseteq Ag$ such that $\tau \models \Delta_H^{\min} \varphi$ and $x \notin H$. $\tau \models \Delta_H^{\min} \varphi$ implies $\tau \models [H]\varphi$, $x \notin H$ implies $H \subseteq Ag - \{x\}$, and by ‘superadditivity’ Eq. 15, $\tau \models [Ag - \{x\}]\varphi$. Conversely, suppose $\tau \models \partial_{Ag} \varphi$ and $\tau \models [Ag - \{x\}]\varphi$. Then $\tau \models \partial_{Ag - \{x\}} \varphi$ and so there is some $H \subseteq Ag - \{x\}$ such that $\tau \models \Delta_H^{\min} \varphi$ (c.f. Eq. 42). Since $x \notin H$, x is not essential for φ at τ . $\square$

Example (Bomb) a and b are officers. c and d are technicians. a commands both c and d . b commands only d . In order to detonate a certain bomb, at least one officer together with at least one of her subordinate technicians must press their buttons. The following table shows all the combinations of button pushes which detonate the bomb (k). Combinations which do not detonate the bomb, or do not guarantee the bomb is detonated, are not shown in the table.

⁵I am grateful to one of the anonymous referees for suggesting this reading, which I quote here almost verbatim.

<i>pushers</i>	$\Delta_G^{\min} k$	$\Delta_G^{\max} k$	<i>essential</i>	<i>bystanders</i>
(a, b, c, d)	$\{a, c\}, \{a, d\}, \{b, d\}$	$\{a, b, c, d\}$	$\{\}$	$\{\}$
$(a, b, c, \cdot)$	$\{a, c\}$	$\{a, c\}$	$\{a, c\}$	$\{b, d\}$
$(a, b, \cdot, d)$	$\{a, d\}, \{b, d\}$	$\{a, b, d\}$	$\{d\}$	$\{c\}$
$(a, \cdot, c, d)$	$\{a, c\}, \{a, d\}$	$\{a, c, d\}$	$\{a\}$	$\{b\}$
$(a, \cdot, c, \cdot)$	$\{a, c\}$	$\{a, c\}$	$\{a, c\}$	$\{b, d\}$
$(a, \cdot, \cdot, d)$	$\{a, d\}$	$\{a, d\}$	$\{a, d\}$	$\{b, c\}$
$(\cdot, b, c, d)$	$\{b, d\}$	$\{b, d\}$	$\{b, d\}$	$\{a, c\}$
$(\cdot, b, \cdot, d)$	$\{b, d\}$	$\{b, d\}$	$\{b, d\}$	$\{a, c\}$

One can see from the table that, in general, the set of agents essential for φ at τ do not bring about φ at τ . They are essential but their actions are not necessarily sufficient to guarantee that φ .

What if the set G of agents essential for φ at τ are sufficient to guarantee that φ at τ ? That would be, by Proposition 11:

$$\tau \models [G]\varphi \wedge \bigwedge_{x \in G} \neg[Ag-\{x\}]\varphi \quad (63)$$

That is (Definition 7) $\tau \models \Delta_G^{\text{sole}}\varphi$. In that case we have $\tau \models \Delta_G^{\min}\varphi$ and $\tau \models \Delta_G^{\max}\varphi$, and all contributors to φ are also all essential for φ .

It is not necessary in Eq. 63 to specify explicitly that G is the set of all agents essential for φ at τ : that outside G there is no agent essential for φ at τ . For suppose there were: suppose $y \notin G$ is essential for φ at τ . Then since $\models [G]\varphi \rightarrow [G \cup \{y}]\varphi$ ('superadditivity'), we would have $\tau \models [G \cup \{y}]\varphi \wedge \bigwedge_{x \in G \cup \{y\}} \neg[Ag-\{x\}]\varphi$, and hence $\tau \models \Delta_{G \cup \{y\}}^{\text{sole}}\varphi$ and $\tau \models \Delta_{G \cup \{y\}}^{\min}\varphi$. But that cannot be because $\tau \models \Delta_G^{\text{sole}}\varphi$ and $\tau \models \Delta_G^{\min}\varphi$.

5 Impotent bystanders

Example (Poisoners) Suppose a and b have poison and c and d have antidote. The doses are of equal strength and will cancel out if administered in equal measure. If a and b administer their poison, and c adds its antidote and d does not, then the poison outweighs the antidote and the concoction is lethal (k). We have $\Delta_{\{a,b,d\}}^{\min} k$; if d does not add its antidote then whether or not c adds its antidote makes no difference to k . In this example, $\Delta_{\{a,b,d\}}^{\max} k$ is also true, so c is a bystander for k . (It may seem a little strange to say that c , who administered antidote, is a mere bystander for k whereas d , who administered neither antidote nor poison, is not a mere bystander, but the point is not who did or did not administer antidote but whose actions are essential/inessential for bringing about k . From that point of view the term 'non contributor' is perhaps more apt than 'bystander'.)

Compare c with some other agent e who has neither poison nor antidote, nor any way of affecting a , b , c or d 's actions. e is also a bystander (non contributor) for k . The difference is that c is a bystander for k on this particular occasion. There is another, counterfactual situation in which c is not a bystander for k : had c not added

antidote while a , b and d acted as they did then $\Delta_{\{a,b,c\}}^{\min} k$ and $\Delta_{\{a,b,d\}}^{\min} k$ would be true and c would not be a bystander for k . e on the other hand is always a bystander for k : there is no alternative situation in which e is not a bystander for k .

This suggests distinguishing between two kinds of bystanders: those whose actions did not contribute to the bringing about of φ on this occasion but who *might* have contributed had they acted otherwise—let us call them ‘potentially contributing bystanders’—and those others whose actions would make no difference to the bringing about of φ , by any G . These bystanders are not potentially contributing: let us call them ‘impotent bystanders’.

To take another example: in the Brexit referendum of 2016 in which the majority of those voting indicated a wish for the United Kingdom to leave the European Union, it was the ‘leave’ voters and the abstainers together who brought about the outcome. They were the ‘contributors’. Those who voted ‘remain’ were ‘bystanders’. An abstainer is someone who is entitled to vote but does not. Abstainers contributed to the outcome because had they voted ‘remain’ the outcome might have been different. The ‘remain’ voters were ‘bystanders’—but not ‘impotent bystanders’. Had they voted ‘leave’ or abstained they would have been ‘contributors’. The ‘impotent bystanders’ were those who were not entitled to vote.

5.1 Potentially contributing and impotent bystanders

Definition 12 x is a *potentially contributing bystander* for φ at τ iff x is a mere bystander for φ at τ and x is not a mere bystander for φ at some alternative $\tau' \sim \tau$.

x is an *impotent bystander* for φ at τ iff x is a mere bystander for φ at τ but not a potentially contributing bystander for φ at τ .

Now instead of $nmb(G, \varphi)$ and $omb(G, \varphi)$ we will have $nib(G, \varphi)$ and $oib(G, \varphi)$ (for ‘no impotent bystanders’ and ‘outside only impotent bystanders’).

Proposition 12 x is an *impotent bystander* for φ at τ iff:

$$\tau \models \partial_{Ag} \varphi \ \& \ \nexists H (\tau \models \Diamond \Delta_H^{\min} \varphi \ \& \ x \in H)$$

Proof x is a *potentially contributing bystander* for φ at τ

$$\begin{aligned} \text{iff } \tau \models \partial_{Ag} \varphi \ \& \ \nexists G (\tau \models \Delta_G^{\min} \varphi \ \& \ x \in G) \\ \& \ \exists H (\tau' \models \Delta_H^{\min} \varphi \text{ for some } \tau' \sim \tau \ \& \ x \in H) \end{aligned}$$

$$\begin{aligned} \text{iff } \tau \models \partial_{Ag} \varphi \ \& \ \nexists G (\tau \models \Delta_G^{\min} \varphi \ \& \ x \in G) \\ \& \ \exists H (\tau \models \Diamond \Delta_H^{\min} \varphi \ \& \ x \in H) \end{aligned}$$

x is an *impotent bystander* for φ at τ

$$\text{iff } \tau \models \partial_{Ag} \varphi \ \& \ \nexists G (\tau \models \Delta_G^{\min} \varphi \ \& \ x \in G) \ \& \ \nexists H (\tau \models \Diamond \Delta_H^{\min} \varphi \ \& \ x \in H)$$

$$\text{iff } \tau \models \partial_{Ag} \varphi \ \& \ \nexists H (\tau \models \Diamond \Delta_H^{\min} \varphi \ \& \ x \in H)$$

The last simplification step is because $\models \Delta_H^{\min} \varphi \rightarrow \Diamond \Delta_H^{\min} \varphi$ ($\Box$ is type S5). $\square$

So let us define 'G contains no impotent bystander for φ ' as follows.

Definition 13 For every subset G of Ag :

$$\begin{aligned} \tau \models nib(G, \varphi) \quad \text{iff} \quad \forall x (x \in G \Rightarrow \exists H (\tau \models \Diamond \Delta_H^{\min} \varphi \ \& \ x \in H)) \\ \tau \models nib(G, \varphi) \quad \text{iff} \quad G \subseteq \bigcup_H (\tau \models \Diamond \Delta_H^{\min} \varphi) \end{aligned}$$

Expressed as a formula: $nib(G, \varphi) =_{\text{def}} \bigwedge_{x \in G} \bigvee_{x \in H} \Diamond \Delta_H^{\min} \varphi$.

Note that we have chosen to omit the $\partial_{Ag} \varphi$ condition: $nib(G, \varphi)$ is true when there are no bystanders for φ in G of any kind, not only impotent ones. To say that there are bystanders for φ in G but no impotent bystanders for φ in G we would write $\partial_{Ag} \varphi \wedge nib(G, \varphi)$.

Clearly (because $\models \Delta_H^{\min} \varphi \rightarrow \Diamond \Delta_H^{\min} \varphi$):

$$\models nmb(G, \varphi) \rightarrow nib(G, \varphi) \quad (64)$$

Now define 'outside G there are only impotent bystanders for φ ':

$$\tau \models oib(G, \varphi) \quad \text{iff} \quad \forall x (\exists H (\tau \models \Diamond \Delta_H^{\min} \varphi \ \& \ x \in H) \Rightarrow x \in G) \quad (65)$$

or equivalently (cf. the treatment of $omb(G, \varphi)$), as follows.

Definition 14 For every subset G of Ag :

$$\begin{aligned} \tau \models oib(G, \varphi) \quad \text{iff} \quad \forall H (\tau \models \Diamond \Delta_H^{\min} \varphi \Rightarrow H \subseteq G) \\ \tau \models oib(G, \varphi) \quad \text{iff} \quad \bigcup_H (\tau \models \Diamond \Delta_H^{\min} \varphi) \subseteq G \end{aligned}$$

Expressed as a formula, e.g. via Eq. 65: $oib(G, \varphi) =_{\text{def}} \bigwedge_{x \notin G} \neg \bigvee_{x \in H} \Diamond \Delta_H^{\min} \varphi$.

Clearly:

$$\models oib(G, \varphi) \rightarrow omb(G, \varphi) \quad (66)$$

Now we will define $\Gamma_G \varphi$ to say that G brings it about that φ and consists exactly of those agents who are not impotent bystanders for φ . As in the case of $\Delta_G^{\max} \varphi$, there is a simplification of the definition, because of the following general property.

Proposition 13

$$\models \partial_{Ag} \varphi \wedge oib(G, \varphi) \rightarrow \partial_G \varphi$$

Proof From Proposition 8 and $\models oib(G, \varphi) \rightarrow omb(G, \varphi)$. In full: suppose $\tau \models \partial_{Ag} \varphi \wedge oib(G, \varphi)$. From $\tau \models \partial_{Ag} \varphi$, there exists H such that $\tau \models \Delta_H^{\min} \varphi$, and so also $\tau \models \Diamond \Delta_H^{\min} \varphi$. Now $\tau \models oib(G, \varphi)$ implies $H \subseteq G$. So there exists $H \subseteq G$ such that $\tau \models \Delta_H^{\min} \varphi$, which implies $\tau \models \partial_G \varphi$. $\square$

So we define:

Definition 15 For every subset G of Ag :

$$\Gamma_G \varphi =_{\text{def}} \partial_{Ag} \varphi \wedge nib(G, \varphi) \wedge oib(G, \varphi)$$

that is, in full:

$$\tau \models \Gamma_G \varphi \quad \text{iff} \quad \tau \models \partial_{Ag} \varphi \ \& \ G = \bigcup_H (\tau \models \Diamond \Delta_H^{\min} \varphi)$$

Expressed as a formula:

$$\Gamma_G \varphi =_{\text{def}} \partial_{Ag} \varphi \wedge \bigwedge_{x \in G} \bigvee_{x \in H} \Diamond \Delta_H^{\min} \varphi \wedge \bigwedge_{x \notin G} \neg \bigvee_{x \in H} \Diamond \Delta_H^{\min} \varphi$$

$\Gamma_G \varphi$ can also be defined in terms of $\Diamond \Delta_H^{\max} \varphi$ instead of $\Diamond \Delta_H^{\min} \varphi$, because of the following.

Proposition 14

$$\bigcup_H (\tau \models \Diamond \Delta_H^{\max} \varphi) = \bigcup_G (\tau \models \Diamond \Delta_G^{\min} \varphi)$$

Proof This is a routine derivation using only that $\square$ is type S5. It is a straightforward modification of the (trivial) observation that $\bigcup_H (\tau \models \Delta_H^{\max} \varphi) = \bigcup_G (\tau \models \Delta_G^{\min} \varphi)$. In detail:

$$\begin{aligned} x \in \bigcup_H (\tau \models \Diamond \Delta_H^{\max} \varphi) & \quad \text{iff} \quad \exists H (\tau \models \Diamond \Delta_H^{\max} \varphi \ \& \ x \in H) \\ & \quad \text{iff} \quad \exists H \exists \tau' (\tau \sim \tau' \ \& \ \tau' \models \Delta_H^{\max} \varphi \ \& \ x \in H) \\ & \quad \text{iff} \quad \exists \tau' (\tau \sim \tau' \ \& \ \exists H (\tau' \models \Delta_H^{\max} \varphi \ \& \ x \in H)) \\ & \quad \text{iff} \quad \exists \tau' (\tau \sim \tau' \ \& \ x \in \bigcup_G (\tau' \models \Delta_G^{\min} \varphi)) \\ & \quad \text{iff} \quad \exists \tau' (\tau \sim \tau' \ \& \ \exists G (\tau' \models \Delta_G^{\min} \varphi \ \& \ x \in G)) \\ & \quad \text{iff} \quad \exists G \exists \tau' (\tau \sim \tau' \ \& \ \tau' \models \Delta_G^{\min} \varphi \ \& \ x \in G) \\ & \quad \text{iff} \quad \exists G (\tau \models \Diamond \Delta_G^{\min} \varphi \ \& \ x \in G) \\ & \quad \text{iff} \quad x \in \bigcup_G (\tau \models \Diamond \Delta_G^{\min} \varphi) \end{aligned} \quad \square$$

Corollary 2

$$\tau \models \Gamma_G \varphi \quad \text{iff} \quad \tau \models \partial_{Ag} \varphi \ \& \ G = \bigcup_H (\tau \models \Diamond \Delta_H^{\max} \varphi)$$

The following properties, among others, follow from the definitions and preceding discussion.

Proposition 15

- (i) $\models \Gamma_G \varphi \rightarrow \partial_G \varphi$
- (ii) $\models \Gamma_G \varphi \wedge \Gamma_H \varphi \rightarrow \perp$ for $G \neq H$
- (iii) if $\tau \models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi$ then $H \subseteq G$ (because $\models \Delta_H^{\min} \varphi \rightarrow \Diamond \Delta_H^{\min} \varphi$)
- (iv) if $\tau \models \Gamma_G \varphi \wedge \Delta_H^{\max} \varphi$ then $H \subseteq G$
- (v) $\models \Delta_H^{\max} \varphi \rightarrow \bigvee_{H \subseteq G} \Gamma_G \varphi$ ('there exists $G \supseteq H$ such that $\Gamma_G \varphi$ ')

What else can we say about $\Gamma_G \varphi$? That it is 'agentive' in G ? That it is 'strictly agentive' in G ? That it is 'collectively agentive' in G , in the sense that $\models \Gamma_G \varphi \rightarrow \Gamma_G \Gamma_G \varphi$?

First, two useful and rather obvious properties.

Lemma 2

- (i) $\models nib(G, \varphi) \rightarrow \Box nib(G, \varphi)$
- (ii) $\models oib(G, \varphi) \rightarrow \Box oib(G, \varphi)$

Proof The first relies on $\models \Diamond \Delta_H^{\min} \varphi \leftrightarrow \Box \Diamond \Delta_H^{\min} \varphi$. The second relies on $\models \neg \Diamond \Delta_H^{\min} \varphi \leftrightarrow \Box \neg \Diamond \Delta_H^{\min} \varphi$. Both are valid because $\Box$ is of type S5.

$$\begin{aligned}
 nib(G, \varphi) &\leftrightarrow \bigwedge_{x \in G} \bigvee_{x \in H} \Diamond \Delta_H^{\min} \varphi && \text{(Definition 13, formula)} \\
 &\leftrightarrow \bigwedge_{x \in G} \Diamond (\bigvee_{x \in H} \Delta_H^{\min} \varphi) \\
 &\leftrightarrow \bigwedge_{x \in G} \Box \Diamond (\bigvee_{x \in H} \Delta_H^{\min} \varphi) && (\Box \text{ S5}) \\
 &\leftrightarrow \Box (\bigwedge_{x \in G} \Diamond (\bigvee_{x \in H} \Delta_H^{\min} \varphi)) \\
 &\leftrightarrow \Box nib(G, \varphi)
 \end{aligned}$$

$$\begin{aligned}
 oib(G, \varphi) &\leftrightarrow \bigwedge_{x \notin G} \neg \bigvee_{x \in H} \Diamond \Delta_H^{\min} \varphi && \text{(Definition 14, formula)} \\
 &\leftrightarrow \bigwedge_{x \notin G} \bigwedge_{x \in H} \neg \Diamond \Delta_H^{\min} \varphi \\
 &\leftrightarrow \bigwedge_{x \notin G} \bigwedge_{x \in H} \Box \neg \Diamond \Delta_H^{\min} \varphi && (\Box \text{ S5}) \\
 &\leftrightarrow \Box (\bigwedge_{x \notin G} \bigwedge_{x \in H} \neg \Diamond \Delta_H^{\min} \varphi) \\
 &\leftrightarrow \Box oib(G, \varphi)
 \end{aligned}$$

□

Remark Suppose $\tau \models \Gamma_G \varphi$ for some τ . Since Γ_G is determined from the potentially contributing members of G at all alternatives τ' of τ , it might seem that $\Gamma_G \varphi$ must also be true at all alternatives τ' of τ at which φ is true, i.e., that the following is valid: $\Gamma_G \varphi \rightarrow \Box (\varphi \rightarrow \Gamma_G \varphi)$. But that is not so: φ may be true at some τ' but without there being a set of agents who bring about that φ at τ' . Consider another variant of the pushing-the-door example. Suppose that if a and b push together they are strong enough to keep a door shut (k) but suppose now that if only one of them pushes then, non-deterministically, the door springs open or not. In the case τ_{ab} where both a and b push we have $\Delta_{\{a,b\}}^{\min} k$ and $\Gamma_{\{a,b\}} k$. But there is an alternative τ_a in which only a pushes and k is true (there is also one in which only a pushes and k is false). In τ_a we have $\tau_a \models k$ but $\tau_a \not\models \partial_{\{a,b\}} k$: the actions of a and b (a pushes and b does not) do not guarantee that the door will stay shut, k . It follows that $\tau_a \not\models \Gamma_{\{a,b\}} k$. So for τ_{ab} we have $\tau_{ab} \models \Gamma_{\{a,b\}} k$ but $\tau_{ab} \models \Diamond (k \wedge \neg \Gamma_{\{a,b\}} k)$.

However, if $\tau \models \Gamma_G \varphi$ and τ' is an alternative to τ in which some set of agents does bring it about that φ , then indeed $\Gamma_G \varphi$ is also true at τ' . The following result confirms that.

Proposition 16

$$\models \Gamma_G \varphi \rightarrow \Box(\partial_{Ag} \varphi \rightarrow \Gamma_G \varphi)$$

Proof $\Box$ is normal so $\models \Box nib(G, \varphi) \rightarrow \Box(\partial_{Ag} \varphi \rightarrow nib(G, \varphi))$. By Lemma 2 we have $\models nib(G, \varphi) \rightarrow \Box(\partial_{Ag} \varphi \rightarrow nib(G, \varphi))$. And similarly for $oib(G, \varphi)$. Clearly $\models \Box(\partial_{Ag} \varphi \rightarrow \partial_{Ag} \varphi)$. So: $\models nib(G, \varphi) \wedge oib(G, \varphi) \rightarrow \Box(\partial_{Ag} \varphi \rightarrow (\partial_{Ag} \varphi \wedge nib(G, \varphi) \wedge oib(G, \varphi)))$. i.e., $\models nib(G, \varphi) \wedge oib(G, \varphi) \rightarrow \Box(\partial_{Ag} \varphi \rightarrow \Gamma_G \varphi)$. $\square$

The fact that $\Gamma_G \varphi$ is ‘agentive’ in G now follows easily.

Proposition 17

$$\models \Gamma_G \varphi \rightarrow \partial_G \Gamma_G \varphi$$

Proof A routine derivation using $\models \partial_G \varphi \rightarrow \partial_G \partial_G \varphi$ (schema Eq. 27), $\models (\partial_G \psi \wedge \Box \psi') \rightarrow \partial_G (\psi \wedge \psi')$ (schema Eq. 26), and Lemma 2:

$$\begin{aligned} &\models \Gamma_G \varphi \rightarrow \partial_G \varphi \wedge nib(G, \varphi) \wedge oib(G, \varphi) \\ &\models \Gamma_G \varphi \rightarrow \partial_G \partial_G \varphi \wedge \Box nib(G, \varphi) \wedge \Box oib(G, \varphi) \quad (\text{Eq. 27 and Lemma 2}) \\ &\models \Gamma_G \varphi \rightarrow \partial_G (\partial_G \varphi \wedge nib(G, \varphi) \wedge oib(G, \varphi)) \quad (\text{Eq. 26 and } \Box \text{ of type S5}) \end{aligned}$$

$\square$

We would expect

$$\not\models \Gamma_G \varphi \rightarrow \Delta_G^{\min} \Gamma_G \varphi$$

And that is clearly right. Consider a simple version of the pushing-the-door-shut example, in which each of a and b by themselves are strong enough to keep the door shut (k). In any situation where the door is pushed shut, we have $\Gamma_{\{a,b\}} k$. However, in that case $\Delta_{\{a,b\}}^{\min} \Gamma_{\{a,b\}} k$ is false: if a pushes (with or without b) we have $[a] \Gamma_{\{a,b\}} k$; if b pushes (with or without a) we have $[b] \Gamma_{\{a,b\}} k$; in no case do we have $\Delta_{\{a,b\}}^{\min} \Gamma_{\{a,b\}} k$.

Proposition 18

$$\models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi \leftrightarrow \Delta_H^{\min} \Gamma_G \varphi$$

Proof Left-to-right: first we prove $\models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \partial_H \partial_G \varphi$. If $\tau \models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi$ then $H \subseteq G$ (Proposition 15). If $H \subseteq G$ then $\models \partial_H \varphi \rightarrow \partial_H \partial_G \varphi$. This is because $\models \partial_H \varphi \rightarrow \partial_H \partial_H \varphi$ (schema Eq. 27). If $H \subseteq G$ then $\models \partial_H \varphi \rightarrow \partial_G \varphi$, and (schema Eq. 25) $\models \neg \Box \partial_G \varphi \rightarrow (\partial_H \partial_H \varphi \rightarrow \partial_H \partial_G \varphi)$. But $\models \partial_H \varphi \rightarrow \neg \Box \varphi$, $\models \neg \Box \varphi \rightarrow \neg \Box \partial_G \varphi$ ($\Box$ is normal and $\models \partial_G \varphi \rightarrow \varphi$) and so $\models \partial_H \varphi \rightarrow \partial_H \partial_G \varphi$. So if $\tau \models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi$ then $H \subseteq G$, $\tau \models \partial_H \varphi$, and $\tau \models \partial_H \partial_G \varphi$.

Now we show $\models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \partial_H \Gamma_G \varphi$. We have:

$$\begin{aligned} &\models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \partial_H \partial_G \varphi \wedge nib(G, \varphi) \wedge oib(G, \varphi) \\ &\models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \partial_H \partial_G \varphi \wedge \Box nib(G, \varphi) \wedge \Box oib(G, \varphi) \quad (\text{Lemma 2}) \\ &\models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi \rightarrow \partial_H (\partial_G \varphi \wedge nib(G, \varphi) \wedge oib(G, \varphi)) \end{aligned}$$

The last step above is because $\models (\partial_H \psi \wedge \Box \psi') \rightarrow \partial_H (\psi \wedge \psi')$ (schema Eq. 26).

So suppose for contradiction that $\tau \models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi$ and $\tau \not\models \Delta_H^{\min} \Gamma_G \varphi$. We have just shown that $\tau \models \partial_H \Gamma_G \varphi$ so it must be that $\tau \models \partial_{H'} \Gamma_G \varphi$ for some proper subset $H' \subset H$. But $\models \partial_{H'} \Gamma_G \varphi \rightarrow \partial_{H'} \varphi$. (This is because $\models \Gamma_G \varphi \rightarrow \varphi$, $\models \Gamma_G \varphi \rightarrow \neg \Box \varphi$, and hence $\models \partial_{H'} \Gamma_G \varphi \rightarrow \partial_{H'} \varphi$ by Eq. 25.) So then $\tau \models \partial_{H'} \varphi$ for some $H' \subset H$ and $\tau \not\models \Delta_H^{\min} \varphi$, which contradicts the assumption.

For the converse: note first that (just shown above) $\models \partial_H \Gamma_G \varphi \rightarrow \partial_H \varphi$ and hence $\models \Delta_H^{\min} \Gamma_G \varphi \rightarrow \partial_H \varphi$. Clearly $\models \Delta_H^{\min} \Gamma_G \varphi \rightarrow \Gamma_G \varphi$. It just remains to show that $\models \Delta_H^{\min} \Gamma_G \varphi \rightarrow \Delta_H^{\min} \varphi$. Suppose not. Then it must be that for some τ , $\tau \models \Delta_H^{\min} \Gamma_G \varphi \wedge \Delta_{H'}^{\min} \varphi$ for some $H' \subset H$. If that were the case, we would have $\tau \models \Delta_{H'}^{\min} \varphi \wedge \Gamma_G \varphi$, and hence $\tau \models \Delta_{H'}^{\min} \Gamma_G \varphi$ by the left-to-right half. But since $H' \subset H$ that would imply $\tau \not\models \Delta_H^{\min} \Gamma_G \varphi$. $\square$

Corollary 3

$$\models \Delta_H^{\min} \varphi \rightarrow (\Gamma_G \varphi \leftrightarrow \Delta_H^{\min} \Gamma_G \varphi)$$

So although $\Gamma_G \varphi$ is not 'strictly agentive' in G , $\Gamma_G \varphi$ is 'strictly agentive' in any subset $H \subseteq G$ such that $\Delta_H^{\min} \varphi$. This is in contrast to $\Delta_G^{\max} \varphi$: the example used in Section 4.4 to show that $\not\models \Delta_G^{\max} \varphi \rightarrow [G] \Delta_G^{\max} \varphi$ also shows that $\not\models \Delta_H^{\min} \varphi \rightarrow (\Delta_G^{\max} \varphi \rightarrow \Delta_H^{\min} \Delta_G^{\max} \varphi)$.

Lastly, the following is worth recording. $\Gamma_G \varphi$ is 'collectively agentive' in G , in the following sense.

Proposition 19

$$\models \Gamma_G \varphi \rightarrow \Gamma_G \Gamma_G \varphi$$

Lemma

$$\models \Gamma_G \varphi \rightarrow \text{nib}(G, \Gamma_G \varphi)$$

Suppose $\tau \models \Gamma_G \varphi$. Suppose $x \in G$. Then $\tau \models \Diamond \Delta_H^{\min} \varphi$ for some H such that $x \in H$, i.e., $x \in H$ and $\tau' \models \Delta_H^{\min} \varphi$ for some $\tau \sim \tau'$. $\tau' \models \Delta_H^{\min} \varphi$ implies $\tau' \models \partial_{A_g} \varphi$. We have $\tau \models \Gamma_G \varphi$ and so by Proposition 16, $\tau' \models \Gamma_G \varphi$. Now $\tau' \models \Gamma_G \varphi \wedge \Delta_H^{\min} \varphi$ implies $\tau' \models \Delta_H^{\min} \Gamma_G \varphi$ by Proposition 18. So for every x we have $\tau \models \Diamond \Delta_H^{\min} \Gamma_G \varphi$ for some H such that $x \in H$, i.e., $\tau \models \text{nib}(G, \Gamma_G \varphi)$.

Lemma

$$\models \text{oib}(G, \varphi) \rightarrow \text{oib}(G, \Gamma_G \varphi)$$

Suppose $\tau \models \text{oib}(G, \varphi)$. We show that if $\tau \models \Diamond \Delta_H^{\min} \Gamma_G \varphi$ then $H \subseteq G$. Suppose $\tau \models \Diamond \Delta_H^{\min} \Gamma_G \varphi$. Then $\tau' \models \Delta_H^{\min} \Gamma_G \varphi$ for some $\tau \sim \tau'$. Now $\tau' \models \Delta_H^{\min} \varphi$ by Proposition 18, and so $\tau \models \Diamond \Delta_H^{\min} \varphi$. But $\tau \models \text{oib}(G, \varphi)$ so $H \subseteq G$.

Proof Proposition 17 gives $\models \Gamma_G \varphi \rightarrow \partial_G \Gamma_G \varphi$. From the definition, $\models \Gamma_G \varphi \rightarrow \text{oib}(G, \varphi)$. With the two lemmas we have

$$\models \Gamma_G \varphi \rightarrow \partial_G \Gamma_G \varphi \wedge \text{nib}(G, \Gamma_G \varphi) \wedge \text{oib}(G, \Gamma_G \varphi)$$

whose consequent is, by definition $\Gamma_G \Gamma_G \varphi$. $\square$

$\Gamma_G \varphi$ expresses a meaningful sense in which it is the set G of agents that collectively brings it about that φ . In the Brexit referendum all those entitled to vote, including the ‘remain’ voters, brought about the outcome in this sense. It must be noted however that $\Gamma_G \varphi$ does not represent any sense of *responsibility* for φ : a potentially contributing bystander for φ is still a bystander for φ and could not reasonably be counted as belonging to the set of agents whose actions are responsible for φ . Those who voted ‘remain’ in the Brexit referendum could not reasonably be counted among those responsible for bringing about that the vote was to leave.

5.2 Necessarily ‘essential for φ ’

Belnap and Perloff’s group-independent notion ‘ x is essential for φ ’ (Definition 9) is already very strong. It says that x belongs to all sets H such that $\Delta_H^{\min} \varphi$ (Proposition 6) or equivalently, assuming $\partial_{Ag} \varphi$ is true, that given the actions of the other agents, without x the outcome φ is not guaranteed: $[Ag - \{x\}] \varphi$ is false (Proposition 11). One might wonder⁶ whether there is a further distinction that can be made analogous to that between ‘impotent bystanders’ and ‘potentially contributing bystanders’, in particular, whether one can distinguish between agents that are merely essential for φ , and agents that are necessarily essential for φ , in the sense that they are essential for φ whenever φ is true, or rather, whenever $\partial_{Ag} \varphi$ is true.

Let us say that x is *necessarily essential* for φ at τ if x is essential for φ at τ and x is essential for φ at every alternative $\tau' \sim \tau$ where $\tau' \models \partial_{Ag} \varphi$.

That can be expressed (c.f. the treatment of impotent bystanders in Section 5 and the observation that $\models \Delta_H^{\min} \varphi \rightarrow \partial_{Ag} \varphi$) as:

$$\tau \models \partial_{Ag} \varphi \ \& \ \forall H \ (\tau \models \Diamond \Delta_H^{\min} \varphi \Rightarrow x \in H) \quad (67)$$

or equivalently (Proposition 11) as a formula thus:

$$\tau \models \partial_{Ag} \varphi \wedge \Box (\partial_{Ag} \varphi \rightarrow \neg [Ag - \{x\}] \varphi) \quad (68)$$

Example (Bomb, modified) As in the previous bomb example, a and b are officers, and c and d are technicians. a commands both c and d . b commands only d . To detonate the bomb an officer together with at least one of her technicians must press their buttons. This time, however, for this (different) bomb, no more than one officer may push her button: if both officers push, with or without their technicians, the bomb does not detonate. The following table shows all the combinations of button pushes which detonate the bomb (k).

⁶This question was raised by one of the reviewers.

<i>pushers</i>	$\Delta_G^{\min} k$	$\Delta_G^{\max} k$	<i>essential</i>	<i>bystanders</i>
$(a, \cdot, c, d)$	$\{a, b, c\}, \{a, b, d\}$	$\{a, b, c, d\}$	$\{a, b\}$	$\{\}$
$(a, \cdot, c, \cdot)$	$\{a, b, c\}$	$\{a, b, c\}$	$\{a, b, c\}$	$\{d\}$
$(a, \cdot, \cdot, d)$	$\{a, b, d\}$	$\{a, b, d\}$	$\{a, b, d\}$	$\{c\}$
$(\cdot, b, c, d)$	$\{a, b, d\}$	$\{a, b, d\}$	$\{a, b, d\}$	$\{c\}$
$(\cdot, b, \cdot, d)$	$\{a, b, d\}$	$\{a, b, d\}$	$\{a, b, d\}$	$\{c\}$

In the first three rows, where a pushes the button and b does not, b 's action is required for the detonation of the bomb because if b acts otherwise and pushes, the bomb does not detonate. Likewise in the last two rows, where b pushes and a does not: if both were to push the bomb would not detonate. One can see from the table that a and b are 'necessarily essential' for the detonation of the bomb. In no combination however are the actions of $\{a, b\}$ alone sufficient to detonate the bomb: a technician is always required as well.

6 Conclusion

In the 2016 Brexit referendum it was the 'leave' voters and the abstainers together who brought about the outcome. They were the 'contributors'. Those who voted 'remain' were not contributors; they were 'mere bystanders' in the terminology of Belnap and Perloff [4]. There is a sense, a very weak sense, in which all of us—those who voted 'leave', those who voted 'remain', abstainers, those who were not even entitled to vote, all of us together—brought about the referendum result. This is the sense of agency expressed by plain group *stit*, which has the property of 'superadditivity'. In principle, with access to individual voting records, one could identify the minimal subsets of 'leave' voters and abstainers whose voting actions brought about the outcome. Each of those minimal subsets would 'strictly *stit*' the outcome in the terminology of Belnap and Perloff [4]. It is also meaningful to say, in yet another sense, that it was those who were entitled to vote who brought about the referendum result. Those who were not entitled to vote were 'impotent bystanders'.

There are various different senses in which a set G of agents, collectively, can be said to bring about a certain outcome. Besides *stit* itself, we identified three forms for special attention. $\Delta_G^{\min} \varphi$, which corresponds exactly to what Belnap and Perloff [4] call 'strictly *stit*', expresses that the set G of agents is a minimal (not necessarily unique) set of agents whose joint actions collectively bring about φ . A variant $\Delta_G^{\text{sole}} \varphi$ holds if there is exactly one such minimal set. It turns out that $\Delta_G^{\text{sole}} \varphi$ is Belnap and Perloff's suggestion for 'the one and only joint agent for φ '. It is noteworthy that $\Delta_G^{\min} \varphi$ is 'strictly agentive' in G , in the sense that $\Delta_G^{\min} \varphi$ implies (and is equivalent to) $\Delta_G^{\min} \Delta_G^{\min} \varphi$. One can also derive the conditions under which $\Delta_G^{\text{sole}} \varphi$ is equivalent to $\Delta_G^{\text{sole}} \Delta_G^{\text{sole}} \varphi$. A second general form, $\Delta_G^{\max} \varphi$, expresses that G is the union of all minimal sets H such that $\Delta_H^{\min} \varphi$. G is the set of 'contributors' to φ . It turns out that the contributors G are what Belnap and Perloff called 'not mere bystanders' for φ . If φ is brought about by any set of agents then such a G always exists and

is unique. It is another natural candidate for ‘the one and only joint agent for φ ’. $\Delta_G^{\max} \varphi$ is not ‘strictly agentive’ nor even ‘agentive’ in G . The third form, $\Gamma_G \varphi$, distinguishes further between ‘potentially contributing bystanders’ (who happen to be ‘mere bystanders’ but might not have been) and ‘impotent bystanders’, who are necessarily bystanders. $\Gamma_G \varphi$ says that G brings about that φ but excludes exactly the ‘impotent bystanders’ for φ . Although $\Gamma_G \varphi$ is not ‘strictly agentive’ in G it is ‘strictly agentive’ in any subset H of G for which $\Delta_H^{\min} \varphi$ holds. $\Gamma_G \varphi$ is also ‘collectively agentive’ in G , in the sense that $\Gamma_G \varphi$ implies (and is equivalent to) $\Gamma_G \Gamma_G \varphi$.

All of these constructions are definable in *stit* logics. The formal results of the paper hold for the ‘deliberative *stit*’ *dstit*, since models for *dstit* are (or can be seen as) a special case of the model structures used in the paper. The formal results also hold for all forms of temporal *stit* in which an atemporal *stit* operator is combined with a separate temporal logic, as in the transition-based *stit*-like formalism in [21, 22] for example and the temporal *stit* logics of e.g. [12, 18]. It is likely that the formal results hold also for those forms of *stit* where the temporal aspects are incorporated in the *stit* operator itself, such as Broersen’s *xstit* [7] and, in particular, the general form of *stit* (the ‘achievement *stit*’) used by Belnap and Perloff in [4]. Details for these other forms of *stit* remain to be checked.

The logic of possibly unwitting, mere behavioural collective agency is surprisingly rich. There are other accounts of agency and responsibility, besides *stit*, that could also be explored in similar fashion. We might also think about adding more features, such as communication between agents, mutual awareness, and joint intention, which are essential ingredients of genuine collective action and have been the focus of other studies.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Balbiani, P., Herzig, A., Troquard, N. (2008). Alternative axiomatics and complexity of deliberative *stit* theories. *Journal of Philosophical Logic*, 37(4), 387–406.
2. Belnap, N., & Perloff, M. (1988). Seeing to it that: A canonical form for agentives. *Theoria*, 54, 175–199. Corrected version in: H.E. Kyburg Jr., R.P. Loui, G.N. Carlson (eds.) *Knowledge Representation and Defeasible Reasoning*, Studies in Cognitive Systems, vol. 5, pp. 167–190. Dordrecht, Boston, London: Kluwer (1990).
3. Belnap, N., & Perloff, M. (1992). The way of the agent. *Studia Logica*, 51, 463–484.
4. Belnap, N., & Perloff, M. (1993). In the realm of agents. *Annals of Mathematics and Artificial Intelligence*, 9(1–2), 25–48.
5. Belnap, N., Perloff, M., Xu, M. (2001). *Facing the future: Agents and choices in our indeterminist world*. Oxford: Oxford University Press.

6. Boudou, J., & Lorini, E. (2018). Concurrent game structures for temporal STIT logic. In Dastani, M., Sukthankar, G., André, E., Koenig, S. (Eds.) *Proceedings 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm* (381–389). ACM Press.
7. Broersen, J. (2011). Deontic epistemic stit logic distinguishing modes of mens rea. *Journal of Applied Logic*, 9(2), 127–152.
8. Broersen, J., Herzig, A., Troquard, N. (2006). Embedding ATL in strategic STIT logic of agency. *Journal of Logic and Computation*, 16(5), 559–578.
9. Broersen, J., Herzig, A., Troquard, N. (2006). From Coalition Logic to STIT. *Electronic Notes in Theoretical Computer Science*, 157, 23–35.
10. Broersen, J., Herzig, A., Troquard, N. (2009). What groups do, can do, and know they can do: an analysis in normal modal logics. *Journal of Applied Non-classical Logics*, 19(3), 261–291.
11. Carmo, J. (2010). Collective agency, direct action and dynamic operators. *Logic Journal of IGPL*, 18(1), 66–98.
12. Ciuni, R., & Lorini, E. (2018). Comparing semantics for temporal STIT logic. *Logique et Analyse*, 61(243), 299–339.
13. Demolombe, R. (2012). Causality in the context of multiple agents. In Ågotnes, T., Broersen, J.M., Elgesem, D. (Eds.) *Proceedings 11th International Conference on Deontic Logic in Computer Science (DEON 2012), Bergen* (pp. 1–15). LNAI 7393, Bergen: Springer.
14. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y. (1995). *Reasoning about knowledge*. Cambridge: MIT Press.
15. Herzig, A., & Schwarzenruber, F. (2008). Properties of logics of individual and group agency. In Areces, C., & Goldblatt, R. (Eds.) *Advances in modal logic*, (Vol. 7 pp. 133–149). London: College Publications.
16. Horty, J.F. (2001). *Agency and deontic logic*. Oxford: Oxford University Press.
17. Horty, J.F., & Belnap, N. (1995). The deliberative stit: a study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24(6), 583–644.
18. Lorini, E. (2013). Temporal STIT logic and its application to normative reasoning. *Journal of Applied Non-classical Logics*, 23(4), 372–399.
19. Pauly, M. (2002). A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12, 146–166.
20. Pörn, I. (1977). *Action theory and social science: Some formal models*. Synthese Library No 120, Dordrecht: D. Reidel.
21. Sergot, M. (2008). Action and agency in norm-governed multi-agent systems. In Artikis, A., O'Hare, G., Stathis, K., Vouros, G. (Eds.) *Engineering Societies in the Agents World VIII, Athens, October 2007, Revised Selected Papers, LNCS 4995*, pp. 1–54. Berlin: Springer.
22. Sergot, M. (2014). Some examples formulated in a 'seeing to it that' logic: Illustrations, observations, problems. In Müller, T. (Ed.) *Nuel Belnap on Indeterminism and Free Action*, pp. 223–256. *Outstanding Contributions to Logic*, Vol. 2. Berlin: Springer Open.