Logic and Logical Philosophy
Volume 27 (2018), 375–413
DOI: 10.12775/LLP.2017.022

Gillman Payette

NORMATIVE INCONSISTENCY: An Xstit Account

Abstract. In this paper we take inspiration from a couple of authors on how to think about normative (in)consistency, and then show how to conceive of normative inconsistency in an xstit framework. One view on normative inconsistency is from von Wright, and the other from Hamblin. These two accounts share a conception of normative inconsistency, but their formal frameworks are very different. We propose a way to get the best of both views on normative inconsistency by using an xstit framework, mixed with a version of Anderson's reduction of deontic logic to alethic modal logic. We consider variations on those ideas and relate it to a work of Ruth Barcan Marcus.

Keywords: xstit; normative consistency; action logic; norms

1. Introduction

The goals of this paper are threefold. First, we will present some previous formulations of normative consistency, in particular those from [23] and [13], and draw lessons from those for developing accounts of normative consistency. Second, using a framework based on xstit logic, we develop an account of normative consistency, show how it heeds those lessons, and how it relates to logical consistency. Finally, we discuss the technical and philosophical upshot of this account of normative consistency and draw some connections with ideas from [17]. Although there are some counterintuitive epiphenomena which arise from the formal analysis, it indicates important directions for further research to remove those issues.

In the following section we discuss the ideas of von Wright and Hamblin. What is needed to make the most of the ideas from those authors is

Received February 9, 2017. *Revised* May 22, 2017. *Published online* July 28, 2017

a language that can express norms and action. In Section 3.1 we present the logic xstit to represent action, and Section 3.2 is a discussion of how to extend xstit to represent norms and the account of normative systems that we are assuming. Section 4 contains our formal development of normative inconsistency in the xstit framework. Finally, we provide some interpretation of the formal results and possible variations of the formal concepts of Section 2. We conclude with a brief summary.

2. (In)Consistency of Codes: Some Background

Intuitively, normative inconsistency is different from regular logical inconsistency. While a normative code that is logically inconsistent should not also be normatively consistent, a normative code can be normatively inconsistent without being logically inconsistent. How this intuition is met differs depending on the framework. In this section we will present two.

2.1. Von Wright

Our discussion of formal accounts of normative inconsistency starts with [23].[1] Von Wright formulates a way of making sense of deontic logic as a logic of rational norm-making. When von Wright says 'norm' he means either a proposition $\mathrm{O}\,\varphi$ or a proposition $\mathrm{P}\,\varphi$ which are interpreted as 'φ is obligatory' and 'φ is permitted', respectively. The norms $\mathrm{O}\,\varphi$ are called O-norms and $\mathrm{P}\,\varphi$ are called P-norms. Also, von Wright says that *genuine* norms are those where the proposition φ is contingent.

A necessary condition for rational norm-making, on von Wright's view, is that norm-makers intend their norms to be followable. Normative consistency can then be defined in terms of followability: a set of norms is normatively consistent if and only if it is followable.

The concept of followability is explained formally in terms of the contents of sets of norms. The content of $\mathrm{O}\,\varphi$ is φ, and the content of $\mathrm{P}\,\varphi$ is φ. But there are different ways that a set of norms can fail to be followable. A set of obligations can be unfollowable when their contents conflict. That is, if Γ is a set of O-norms, then Γ is followable if and only if $\{\,\varphi : \mathrm{O}\,\varphi \in \Gamma\,\}$ is consistent.

But von Wright notes that a set of P-norms is *always* followable. Even when $\mathrm{P}\,\neg\varphi$ and $\mathrm{P}\,\varphi$ are both in the set, any agent can do one or the

[1] A similar system can be found in [2].

other, they need not do both. However, a set of mixed norms, i.e., both P- and O-norms, will be unfollowable if something is both obligatory and omissible, i.e., $\mathrm{O}\,\varphi$ and $\mathrm{P}\,\neg\varphi$ are both in the set. This leads von Wright to a definition of normative consistency as ($\textsc{con}_N$): for a set of mixed norms Γ, $\textsc{con}_N(\Gamma)$ iff for each $\mathrm{P}\,\varphi \in \Gamma$, $\{\,\psi : \mathrm{O}\,\psi \in \Gamma\,\} \cup \{\,\varphi\,\}$ is logically consistent.

This provides an interesting and intuitive sense of normative inconsistency. There are two kinds of normative inconsistency. First, requiring that something be done, and permitting that it not be done. That first kind of inconsistency is for the case of mixed norms. The other is having conflicting obligations. Of course the unfollowability of a set of O-norms, implies that the set is not consistent in the $\textsc{con}_N$ sense. Thus, both senses of normative inconsistency can be captured under the $\textsc{con}_N$ sense alone.

Although von Wright ultimately defines normative consistency in terms of logical consistency, he does not upset the basic intuition above: $\{\,\mathrm{O}\,\varphi, \mathrm{O}\,\neg\varphi\,\}$ is normatively inconsistent, while being logically consistent, classically speaking. His requirement that that set of norms be normatively inconsistent produces a logical relationship between $\mathrm{O}\,\varphi$ and $\mathrm{O}\,\neg\varphi$, thus defining basic deontic logic.

There are three lessons to draw from this approach. First, the language is not very expressive. We might do better if we thought about normative inconsistency in another framework, i.e., one that expresses action in some way. Second, as von Wright suggests we should not start with deontic logic, but rather take normative consistency as basic. Third, followability is a good intuitive conception of normative consistency, and what we need to explore is how that notion relates to logical consistency, if it does at all.

2.2. Hamblin

[13] offers a more subtle sense in which norms can conflict or be inconsistent.[2] If we think about actions as transitions from one state of the world to another, and the world as a whole tree that branches into the future, then we have Hamblin's formal model. Each path from node to node in the tree, i.e., a branch, is a sequence of actions an agent might take, and so displays a way the world might unfold as a result of those

[2] There are other accounts of normative inconsistency that take inspiration from Hamblin, but use different formal frameworks, see [6].

actions. Hamblin thinks of norms as sets of transitions between nodes in such trees. Each norm, i.e., set of transitions, represents all of the *illegal transitions* according to a set of norms. So if a transition is in the norm N, then that transition transgresses against the norm. A set of norms or *code* is then the union of the individual norms.

This model for thinking about action and norms allows Hamblin to identify a sense of normative inconsistency that departs from the technical definition von Wright gives in terms of logical consistency, but remains true to the sprit of von Wright's idea of being followable. Note that Hamblin was not responding to von Wright or vice versa. These seem to be independent and isolated discussions of normative consistency. For Hamblin, inconsistency in a set of norms, i.e., a set of transitions in a tree, arises when an agent is placed in a situation where all of the transitions available to her are illegal. This is what Hamblin calls a *quandary*. Note that in Hamblin's formalism logical consistency does not connect to normative consistency at all since there is not a language to represent the norms. Therefore, Hamblin's account conforms to the basic intuition.

The primary innovation in Hamblin's work is that there are different types of normative inconsistency which correspond to the degree of ease an agent has in avoiding a quandary. This is referred to as a normative code's level of *quandary freeness*. We will not discuss the various kinds of quandaries Hamblin distinguishes in detail since we are just using Hamblin's work to motivate our own. However, here are two types of quandary freeness defined by Hamblin to keep in mind for later discussion: there are norm sets where there is no way for an agent to end up in a quandary (total quandary freeness), and there are others where an agent has to pick a particular sequence of transitions to avoid a quandary (strategic quandary freeness).

There are three lessons to take on board from Hamblin's work. First, there is an issue to do with expressiveness. On Hamblin's account, norms and actions are simply sets of transitions, and no language is provided for representing those norms. We will endeavour to fix that gap. Second, Hamblin's formalism is ready made for his discussion; while it is enlightening, if a more common framework is used, his distinctions can be more widely applied. Lastly, the proliferation of kinds of inconsistency which correspond to how easy it is to avoid a quandary is something to keep in mind. Whereas for von Wright the two kinds of inconsistency could be captured in one easy condition, we can always ask whether that same simplification can be effected.

3. Representing Norms and Actions

One of our lessons from above was that we need a way to represent action. To do this we introduce xstit logic which allows us to represent the effects of choice. How to make sense of norms will come after.

3.1. The Xstit Formalism

The stit (sees-to-it-that) language that we use is called xstit in [8, 7], but the version here differs slightly.[3] Xstit is a logic like the standard stit logic from [5], but instead of seeing to something now/instantaneously, choices determine possible future states. The language is constructed from a *finite* set of agents $\mathbf{Ag}$ and an infinite set of atomic propositions from $\mathbf{At}$. Sets of agents are denoted by capital Roman letters $\mathbf{A}, \mathbf{B}, \mathbf{C}$. We can then define a language $\mathcal{L}_{\text{xstit}}$ as follows:

$$\varphi := \ \bot \mid \mathbf{p} \mid \neg\varphi \mid \Box\varphi \mid \ [\mathbf{A}\,\mathsf{xstit}]\,\varphi \mid X\varphi \mid \varphi \wedge \varphi \mid \varphi \supset \varphi$$

The notation $[\mathbf{A}\,\mathsf{xstit}]\,\varphi$ is more in-line with that from dynamic logic since it says that $\mathbf{A}$ sees to it that φ in the next state. $\Box\varphi$ means that φ is historically necessary. That just means, however, that φ is true relative to every history *at this moment*. This will become clearer when we introduce the semantics. $X\varphi$ means that in the next state relative to the history we are in, φ is true. We will define the other boolean connectives in the usual manner.

Like in [5], the semantics evaluates formulas relative to histories *and* moments. A history is a set of moments that is linearly ordered. This is modelled as follows:

DEFINITION 1. An xstit frame is a triple $\mathfrak{F} = \langle S, H, E\rangle$ such that:

1. $S \neq \varnothing$ are called the *static states*.
2. $H \neq \varnothing$ is a set of orders sets $\langle h, <_h\rangle$ such that for each $h \in H$
   (a) $h \subseteq S$ and $\langle h, <_h\rangle$ is isomorphic to $\mathbb{Z}$ with its usual order, and

[3] The Xstit formalism is definitely a *choice* of focus in formalism. There are other approaches to the logic of norms which have a different focus cf. [10]. There the author takes actions as the basic focus of norms in a dynamic logic setting, and the deontic operators apply directly to those actions. Agents, locations, times, etc., are all subsequently dealt with as parts of situations in which the norms are applied. This Xstit approach takes something of the reverse view by including agents and the effects of their actions and defining deontic concepts from those components.

(b) if $s \in h \cap h'$, then $\{ s' : s' <_h s \} = \{ s' : s' <_{h'} s \}$. Since each order is isomorphic with $\mathbb{Z}$, there is a unique successor and predecessor in h for each $s \in h$, we refer to these by $\mathrm{lub}(s, h)$ and $\mathrm{glb}(s, h)$, respectively. We generalize lub in the following way $\mathrm{lub}(s) = \{ s' : \exists h \mathrm{lub}(s, h) = s' \}$ to give the set of successors of s.[4]

3. $E\colon S \times H \times \mathcal{P}(\mathbf{Ag}) \to \mathcal{P}(S)$ is called an h-effectivity function. The effectivity function assigns a set of static states to each triple $(s, h, \mathbf{A})$. The function E must obey the following conditions:
   (a) if $s \notin h$, then $E(s, h, \mathbf{A}) = \varnothing$
   (b) if $s' \in E(s, h, \mathbf{A})$, then $s' \in \mathrm{lub}(s)$
   (c) if $s \in h$, $\mathrm{lub}(s, h) \in E(s, h, \mathbf{A})$
   (d) $E(s, h, \varnothing) = \mathrm{lub}(s)$, if $s \in h$
   (e) if $s \in h$, then $E(s, h, \mathbf{Ag}) = \{ \mathrm{lub}(s, h) \}$
   (f) if $\mathbf{A} \subseteq \mathbf{B}$, then $E(s, h, \mathbf{B}) \subseteq E(s, h, \mathbf{A})$
   (g) if $\mathbf{A} \cap \mathbf{B} = \varnothing$ and $s \in h \cap h'$, then there is h'' with $s \in h''$ and $E(s, h'', \mathbf{A})$ and $E(s, h'', \mathbf{B})$ are contained in $E(s, h, \mathbf{A})$ and $E(s, h', \mathbf{B})$, respectively.[5]
4. A *dynamic state* is a history-static state pair (s, h), and the domain of $\mathfrak{F}$, $|\mathfrak{F}|$, is the set of dynamic states such that $s \in h$.

We will pause to explain the conditions. The first condition is standard in modal logic. Condition 2a says that we can order each history like $\mathbb{Z}$: $\ldots s_{-2}, s_{-1}, s_0, s_1, s_2, \ldots$ the set of integers. 2b says that these histories as a collection form a tree: there is a common trunk and once two histories diverge, they will not join other histories. Actually, it says that the histories form a forest; the histories can be a collection of non intersecting trees. However, we will require that the histories form just one tree that is $\bigcap H \neq \varnothing$. Such a frame/model will be called a *regular*, universal frame/model. This should conjure images of S5 being incapable of distinguishing the class of universal relations from the class of equivalence relations. The logic of xstit similarly cannot distinguish the regular universal models from the larger class of regular models.

The notion of effectivity function comes from game/social choice theory, see [1], and was brought into logic in coalition logic (cf. [18]). In

[4] The rationale behind lub and glb is to put the terminology in line with that from order theory: least upper bound and greatest lower bound, respectively.

[5] The condition g in [8] and [7] is different in that it only requires that $E(s, h, \mathbf{A}) \cap E(s, h', \mathbf{B}) \neq \varnothing$, but that does not seem to suffice for completeness of the system presented here.

general, an effectivity function represents what a group of agents is capable of bringing about: what results they can be *effective* in achieving. Since $E(s, h, \mathbf{A}) \subseteq \text{lub}(s)$, by condition 3b, for each $s \in S$, the effectivity function selects a set of states from the possible continuations given the current static state. When $s' \in E(s, h, \mathbf{A})$, $\mathbf{A}$'s action/choice relative to h guarantees that s' is one of the possible next states. Moreover, if $s' \notin E(s, h, \mathbf{A})$ but $s' \in \text{lub}(s)$, then s' is not one of the next states given $\mathbf{A}$s collective choice indexed to h.

Each effectivity function is evaluated at a dynamic state (s, h): a history-static state pair. Condition 3a states that it only makes sense for agents to be effective for anything at dynamic states where the static state is in the history, i.e., $s \in h$. This condition is mainly to make E *a function*, rather than a partial function. 3b says that the only states that agents can be effective for bringing about are those that follow in some history running through the current static state. Essentially, agents can only constrain the possible outcomes, not create new ones.

3c says that any set of agents is effective to constrain the outcomes to at least the immediate successor — relative to h — of the current static state. 3d says that the effectivity of the empty set of agents is all of the possible continuations from a static state. The empty set of agents is considered to be Nature's effectiveness; Nature sets the range of possible outcomes.

3e requires that the total set of agents, $\mathbf{Ag}$, determines the successor state of s at h for each $(s, h) \in |\mathfrak{F}|$. The next state in a history can be completely determined by the whole set of agents. Broersen and Meyer point out that although the next static state is determined by the set of all agents, static states are only half of the auxiliary parameters in the evaluation of formulas. The set of agents does not determine the next *dynamic state*.

3f states that the more choices that are made, the more the outcomes are constrained. That results in the anti-monotonicity of effectivity functions. Finally, 3g states that the choices of agents never eliminate what other agents are effective for. This is referred to as *independence of agency*.

The models of xstit are given as follows:

Definition 2. An xstit model $\mathfrak{M}$ is an xstit frame $\mathfrak{F}$ with a valuation $v \colon \mathbf{At} \to \mathcal{P}(S)$.

We can then give the semantics for the language $\mathcal{L}_{\text{xstit}}$:

DEFINITION 3. Truth or satisfaction of a formula in $\mathcal{L}_{\text{xstit}}$ relative to a model $\mathfrak{M}$ and $(s,h) \in |\mathfrak{M}|$ is defined by:

- $(s,h) \vDash \mathbf{p}$ iff $s \in v(\mathbf{p})$
- $(s,h) \nvDash \bot$
- $(s,h) \vDash \neg\varphi$ iff $(s,h) \nvDash \varphi$
- $(s,h) \vDash \varphi \wedge \psi$ iff $(s,h) \vDash \varphi$ and $(s,h) \vDash \psi$
- $(s,h) \vDash \Box\varphi$ iff for all h' with $s \in h'$, $(s,h') \vDash \varphi$
- $(s,h) \vDash X\varphi$ iff $(\text{lub}(s,h),h) \vDash \varphi$
- $(s,h) \vDash \varphi \supset \psi$ iff $(s,h) \nvDash \varphi$ or $(s,h) \vDash \psi$
- $(s,h) \vDash [\mathbf{A}\,\mathsf{xstit}]\,\varphi$ iff for all $s' \in E(s,h,\mathbf{A})$ and $h' \ni s'$, $(s',h') \vDash \varphi$

Satisfiability of a set of formulas Γ is defined as: there is some model $\mathfrak{M} = \langle \mathfrak{F}, v\rangle$ and (s,h) in the domain of $\mathfrak{M}$, $|\mathfrak{M}| = |\mathfrak{F}|$, such that $\mathfrak{M},(s,h) \vDash \varphi$ for each $\varphi \in \Gamma$. As usual we simplify that by writing $\mathfrak{M},(s,h) \vDash \Gamma$. A set Γ *xstit entails* a formula φ ($\Gamma \vDash_X \varphi$) iff for each xstit model $\mathfrak{M}$ and $(s,h) \in |\mathfrak{M}|$ such that $\mathfrak{M},(s,h) \vDash \Gamma$, $\mathfrak{M},(s,h) \vDash \varphi$ also. We denote the set dynamic states at which φ is true/satisfied in the model $\mathfrak{M}$ by $[\![\varphi]\!]_{\mathfrak{M}}$, and as usual reference to the model is often left implicit.

Note that in this semantics when a non-historical formula φ, that is a formula without xstit or X operators, is true, then it is true relative to all the coincident dynamic states. Formally, this means if $\mathfrak{M},(s,h) \vDash \varphi$, and φ is a non-historical formula, then $\mathfrak{M},(s,h') \vDash \varphi$ for all h' with $s \in h'$.

There are two things that we should look at and put to rest before moving on, both of which have to do with agent responsibility in xstit logic. There is something intuitively unusual about the standard xstit. Note two aspects about the satisfaction condition of $[\mathbf{A}\,\mathsf{xstit}]$ (what could be called the Chellas xstit corresponding to $[\mathbf{A}\ \mathit{cstit} : \varphi]$ in regular stit logic). First is that each agent sees-to-it-that all tautologies are true. Since tautologies hold at every moment-history pair, they will hold at all the (s,h) such that $s \in \text{lub}(s')$ and $s \in h$, i.e., all the state/history pairs that follow s'. Of course $E(s',h,\mathbf{A}) \subseteq \text{lub}(s)$, so any tautology will be true at all history/state pairs made from static states in $E(s',h,\mathbf{A})$. This might strike some as odd since no agent is responsible for making a tautology true. The second thing to notice is that what choice is made depends on the history and moment (static state), not just the moment. Relative to a moment, an agent might be able to see to many different sets of outcomes, but relative to a dynamic state, since the xstit operators are evaluated relative to a history, there is only one set of outcomes the agent(s) can see to. Each choice corresponds to or is indexed to a history.

Now we can deal with the oddity of seeing to tautologies. The Chellas xstit does not seem like it has much to do with agency, but when we are discussing ways that normative systems are inconsistent, that will not involve a requirement of agency. We will see more about this when we have discussed normative inconsistency in the xstit framework. The way taken to represent responsibility is via a deliberative xstit operator $[\mathbf{A}\ \mathsf{xdstit}]\,\varphi$ which is true just when $[\mathbf{A}\,\mathsf{xstit}]\,\varphi$ and φ is not true in all possible next states. The reason there is responsibility in the xdstit case is because it was possible for φ not to be the case in the next states. So $\mathbf{A}$'s choice was the cause of φ; $\mathbf{A}$ is agentive in the truth of φ.

A related issue is how responsibility is distributed between the members of $\mathbf{A}$ in xstit logic. In regular stit logic, so-called 'group stit' distributes effectiveness in an even manner. The set of outcomes that a group can determine is simply the intersection of the sets of outcomes that each of the members of $\mathbf{A}$ can determine. That assumption is not made in xstit logic. The outcome of a group choice may not be analyzable as the intersection of the choices of the group's members. That assumption may strike some as odd, but assigning individual responsibility from a group choice/decision is a philosophically difficult problem, cf. [22]. With this assumption we are simply not taking a stand on the matter. Even if we did analyze group choice as the intersection of individual choices, however, it would result in a smaller class of effectivity functions as defined in Definition 1. Now we will turn to the proof theory for xstit.

A Hibert-style proof theory for Xstit can be given as follows:

Definition 4. Assume that $\mathbf{A}, \mathbf{B} \subseteq \mathbf{Ag}$, $\mathbf{p} \in \mathbf{At}$ and $\varphi, \psi \in \mathcal{L}_{\text{xstit}}$,

CL Some axioms for classical logic

$(\mathbf{p})$ $\mathbf{p} \supset \Box\mathbf{p}$

S5 for $\Box$:

- $\Box(\varphi \supset \psi) \supset (\Box\varphi \supset \Box\psi)$
- $\Box\varphi \supset \varphi$
- $\Box\varphi \supset \Box\Box\varphi$
- $\varphi \supset \Box\neg\Box\neg\varphi$

KD for each of $[\mathbf{A}\,\mathsf{xstit}]\,\varphi$ and X:

- $[\mathbf{A}\,\mathsf{xstit}](\varphi \supset \psi) \supset ([\mathbf{A}\,\mathsf{xstit}]\,\varphi \supset [\mathbf{A}\,\mathsf{xstit}]\,\psi)$
- $[\mathbf{A}\,\mathsf{xstit}]\,\varphi \supset \neg\,[\mathbf{A}\,\mathsf{xstit}]\,\neg\varphi$

- $X(\varphi \supset \psi) \supset (X\varphi \supset X\psi)$
- $X\varphi \supset \neg X \neg\varphi$

[DetX] $\neg X\neg\varphi \supset X\varphi$
[NA] $[\mathbf{A}\,\mathsf{xstit}]\,\varphi \supset [\mathbf{A}\,\mathsf{xstit}]\,\Box\varphi$
[SettX] $[\varnothing\,\mathsf{xstit}]\,\varphi \equiv \Box X\varphi$
[XSett] $[\mathbf{Ag}\,\mathsf{xstit}]\,\varphi \equiv X\Box\varphi$
[C-mon] $[\mathbf{A}\,\mathsf{xstit}]\,\varphi \supset [\mathbf{A} \cup \mathbf{B}\,\mathsf{xstit}]\,\varphi$
[Indep-G] $\Diamond[\mathbf{A}\,\mathsf{xstit}]\,\varphi \wedge \Diamond[\mathbf{B}\,\mathsf{xstit}]\,\psi \;\supset\; \Diamond([\mathbf{A}\,\mathsf{xstit}]\,\varphi \wedge [\mathbf{B}\,\mathsf{xstit}]\,\psi)$, where $\mathbf{A} \cap \mathbf{B} = \varnothing$

For the rules we have modus ponens (MP) and the necessitation rule: If $\vdash \varphi$, then $\vdash \clubsuit\varphi$ for $\clubsuit \in \{\,\Box, X\,\} \cup \{\,[\mathbf{A}\,\mathsf{xstit}] : \mathbf{A} \subseteq \mathbf{Ag}\,\}$. Note that $\Diamond = \neg\Box\neg$. Since all of the axioms are in the Salquvist class, the axiom system is complete with respect to its Kripke models. By an unravelling argument it can be shown that the axioms are complete with respect to the class of regular models, i.e., those of Definition 2, and regular universal models. Now we will discuss how to represent norms with xstit logic.

3.2. Norms

To represent norms, we will make an assumption about the nature of norms. The assumption we make is that norms are constructed in a more or less Searlean manner as in [19, 20]. This is to say that a normative system is made up of a set of status function declarations. A status function declaration creates social items by imposing special roles on objects or persons that they could not perform or hold in virtue of their physical constitution. A favourite example of Searle's is money/currency. A paper banknote cannot simply play the role it does based on its physical constitution. It plays the role it does because we treat it as having that function; we impose that function on the piece of paper.

A system of norms on Searle's account can then be represented by a collection of constitutive rules, these are the rules that specify the status functions. These norms are represented as so-called *count-as conditionals*, and take the form 'x counts as y in context C'. An example of a count-as conditional that we might formulate is: Emma seeing to it that the dishes are washed counts as seeing to it that her weekly chores are completed in the context of our household rules. The collection of count-as conditionals that would be needed to specify any actual norm system would be enormous and very complex, but the assumption is that it would be possible to analyze such norm systems in terms of them.

However, we are working with a simplified formal system, particularly one that is propositional, albeit it can represent certain modalities of action. But to represent the count-as conditional, we have to extend the language with a special conditional: $\rightarrow$. The count-as conditional, however, has been studied by various authors, cf. [11, 15]. What is clear from those studies is that there is no consensus on how the count-as conditional behaves. But there is consensus that it does not behave like a material conditional. If we follow the approach in [15], then we can extend the logic in such a manner that incorporates a conditional $\rightarrow$ with a semantics that can be manipulated very easily and give $\rightarrow$ different properties.

To interpret $\rightarrow$, we add to models $\mathfrak{M} = \langle S, H, E, v\rangle = \langle \mathfrak{F}, v\rangle$ a function $f\colon |\mathfrak{F}| \times \mathcal{P}(|\mathfrak{F}|) \rightarrow \mathcal{P}(|\mathfrak{F}|)$. Alternatively, the function f is such that for each proposition X (set of dynamic states), and dynamic state (s, h), $f((s, h), X)$ is another proposition. The way we interpret $\rightarrow$, then is:

$$\mathfrak{M}, (s, h) \vDash \varphi \rightarrow \psi \Longleftrightarrow [\![\psi]\!] \subseteq f((s, h), [\![\varphi]\!]).$$

Having said all of this, we do not want to commit to a particular set of logical properties for $\rightarrow$, apart from being able to detach the consequent when the antecedent is true. But what we say in what follows is not affected as long the logic that is used when extended by $\rightarrow$, is complete with respect to the class of models that is extended with the function f. Giving such a completeness proof would take us too far off track, and we can see how it would proceed using the usual canonical model-type constructions in [9] anyway. Thus we will suppress reference to $\rightarrow$, assuming that there is such a logic in place. There is, however, another aspect of representing norms that we must include in our later discussions.

One of the fundamental parts of a normative system is to define what duties and rights its subjects have. This can be accomplished via constitutive, counts-as norms by using Anderson's reduction of deontic logic to alethic modal logic allows use a violation constant V, see [4]. This method is suggested by [11] and in a different manner in [12]. Anderson's reduction simply adds a new symbol V to the language of alethic modal logic, and interprets it as 'There is a violation'. Thus, it is not a personal violation, i.e., one attributed to a particular individual.

In Anderson's original work he assumed that V can not be true at every possible world: $\neg\Box V$ was an axiom. What that means from an intuitive understanding is that it would always be possible to avoid violations. However, we do not share that assumption because that assumption implies that one could not create an incoherent sets of norms. Indeed, the whole point of this discussion is to investigate that possibility.

In this extended formalism, norms are represented by sentences that stipulate what counts as what, and in particular what counts-as a violation. For example, when we are trying to specify the illegality of speeding while driving a car in a school zone we can do it as follows: driving over 60 km/h in a car on particular roadways in a certain vicinity of a school counts as a violation in the context of the traffic code. Formally, this would be represented as $\varphi \to V$, where φ encodes the antecedent of the English sentence above.

We can also specify other legal/deontic relations particularly those from [16] which are representations of the legal relations in [14]. The central deontic/legal concepts that we should look at representing are: duties, prohibitions, rights, permissions, and powers.

To say that an agent or group has a power is simply expressed via the xstit operator: $\mathbf{A}$ has the power to bring about φ iff $\Diamond\,[\mathbf{A}\,\mathsf{xstit}]\,\varphi$ is true, i.e., it is possible for $\mathbf{A}$ to see to φ. The way that duties are specified in this reduction is as: $\mathbf{A}$ has a duty to bring about φ iff $\mathbf{A}$ failing to bring about φ counts as a violation.

Prohibitions can be defined as well as $\mathbf{A}$ bringing about φ counts as bringing about a violation. Permissions are tricky since we are talking about permission relative to a code, and there is a long standing debate regarding internal and external permissions (see [3, 21]). An internal permission is an explicit permission in the code that says something is permitted. An external permission is one that that is simply not prohibited by the code. But following [24], we look at explicit/internal permissions in this setting as playing no genuine role since an explicit permission to bring about φ does not provide any new options unless the permission is granted as an *exception* to a prohibition or duty in some particular circumstance. Thus permissions, when meaningful, are written into the rules as exceptions to duties or prohibitions. Thus if something is not prohibited, it is permitted, but that is just to say: bringing about φ is not prohibited.

Rights are represented as kinds of prohibitions. $\mathbf{A}$ would have a right to φ when all of the other agents would be in violation if they block $\mathbf{A}$ from achieving φ. Formally,

$$\bigvee_{\mathbf{B}\subseteq\mathbf{Ag}\setminus\mathbf{A}}([\mathbf{B}\,\mathsf{xstit}]\,\neg\Diamond\,[\mathbf{A}\,\mathsf{xstit}]\,\varphi) \to V\,.$$

That is, some group other than **A** making **A** unable to see to φ counts as a violation.

Formally, we have added the violation constant V to the already extended xstit language with $\to$. But the violation constant is treated as any other atomic sentence. Thus $v(V)$ can be any subset of S. In what follows the violation constant will play a crucial role since it will be central to defining when a code is unfollowable at a dynamic state (s, h).

As a brief aside, note that there is an interesting expression in this new language: $[\mathbf{A}\,\mathsf{xstit}](\varphi \to \psi)$. This expresses that **A** brings it about that φ counts as ψ. Thus we can express that agents bring about norms, as they often do. However, we will leave investigation of that and norm change for another paper.

We are now in a position to indicate how we make good on some of the goals derived from the discussion of von Wright and Hamblin. We have a language in which we can express action — in a certain manner. Most importantly in the extended language which we will dub $\mathcal{L}^I_{\text{xstit}}$, that is $\mathcal{L}_{\text{xstit}}$ extended with $\to$ and V, we get a language in which we can represent norms. The notions of satisfiability, consequence, provability, and consistency will all carry over to this extended language. The formulation also assumes a certain conception of norms that is fairly popular. Searle's account may not be popular in all of its details, but specifying norm systems as count-as conditionals is fairly common. Thus norm systems or codes can simply be specified formally as a set of $\mathcal{L}^I_{\text{xstit}}$ sentences, $\Gamma, \Delta, \ldots$. Note that although we have specified deontic relations, we have not assumed a deontic logic of any sort. Whatever deontic logic that might be derived is one that would arise from the logic of the language $\mathcal{L}^I_{\text{xstit}}$, which we will refer to as $\vdash_{ix}$ — ix is an xstit logic for institutions. Institutions, after all, are specified by their norms. Now that we have laid down a rough philosophical and formal framework for representing norms, we can consider normative inconsistency.

4. Normative Inconsistency

We will interpret the notion of normative inconsistency in terms of norms being transgressed which connects to our discussion of Hamblin in Section 2. In Hamblin's case the content of norms are represented as sets of transitions; they are the transitions that disobey the norms. Consistency corresponds to quandary freeness, i.e., avoiding situations where all of the transitions are transgressions. To translate this idea into the xstit framework presented above, we have to discuss some aspects of norms not touched on yet.

The metaphysics of norms presented above takes norms to be constitutive rules. We say that a norm is *in force* when it is in effect, valid, recognized, ...; in other words a norm is in force when it is *a norm*. Driving on the righthand side of the road is a norm in force in North America, it is not in force in Britain. The account of Section 3.2 analyzes norms as status function declarations. In our formalism, the formulas correspond to the content of the status functions which have been declared. It is the act of declaration that would bring a norm into force. However, we will not represent that act in the formalism. We represent the effect of such acts of declaration by stipulating that a norm is in force when the formula corresponding to its content is *true*.

In our formal model, then, any formula can represent a norm. That may seem odd, but it is supposed to permit us to model the fact that certain social facts are brought about by a declaration. For example, suppose a certain group is granted a certain legal power φ. That would correspond to the truth of $\Diamond\,[\mathbf{A}\,\mathsf{xstit}]\,\varphi$. But we also want to allow the possibility that certain basic social facts be declared true by fiat: that would correspond to the truth of an atomic sentence. Of course we do not think that every basic fact, if there are such things, would be under the control of institutional powers. But this formal system is an abstract language, so any atomic sentence could be interpreted as something that is under institutional control.

A normative system or code is represented by a set of formulas Δ; as an abuse of terminology we will identify the norms with formulas. A code need not be a theory since it is supposed to represent the collection of norms explicitly promulgated. Therefore, a code is in force when each of its member formulas is true. Thus, in the formal models, we say that a code Δ is in force at a dynamic state (s, h) in a model $\mathfrak{M}$, when $\mathfrak{M}, (s, h) \vDash \Delta$.

Consider the theory of a dynamic state, i.e., $\{\, \varphi : \mathfrak{M}, (s, h) \vDash \varphi \,\}$. In our framework that could be a code. Indeed, which formulas are distinguished as norms of a code is something that is decided external to the formal model. Nonetheless, we would like to distinguish certain things simply happening, i.e., certain formulas being true at some states, from being a norm that is in force. To accomplish this we introduce a technical concept for codes, that of being *sustained in force*.

Definition 5. Let

$$\mathrm{AFT}(s,h) = \{\, (s',h') \in |\mathfrak{M}| : \ s \in h' \ \&\ s \leq_{h'} s' \,\} \,.$$

A code Δ is sustained in force at a point (s,h) in an $\mathcal{L}^I_{\text{xstit}}$-model $\mathfrak{M}$, in symbols $\mathfrak{M},(s,h) \lessdot \Delta$, iff for each $(s',h') \in \mathrm{AFT}(s,h)$, $\mathfrak{M},(s',h') \vDash \Delta$.[6]

Note that $|\mathfrak{M}| = \{\, (s,h) \in S \times H : s \in h \,\}$. Also, if $\mathfrak{M},(s,h) \lessdot \Delta$, then $\mathfrak{M},(s,h) \vDash \Delta$. When Δ is sustained in force at (s,h), it is satisfied at (s,h), and it is satisfied at every state "after" that one. The set of states in $\mathrm{AFT}(s,h)$ are all states that are potentially after (s,h), they are all ways the world could turn out after (s,h). This definition requires any code that can be sustained in force to be $\mathcal{L}^I_{\text{xstit}}$-consistent (by completeness). As an abuse of notation will we say $s' \in \mathrm{AFT}(s,h)$ iff there is h' with $s' \in h'$ and $(s',h') \in \mathrm{AFT}(s,h)$. We can now return to our development of normative consistency.

When we consider the consistency of a code Δ we will look at characteristics of how it is sustained in force. This dovetails with Hamblin's account of looking at how a code is followed over time, and it also allows us to distinguish which formulas are parts of the code from those which are true on particular occasions. That is particularly important for representing how a code is transgressed.

In Hamblin's framework, a transgression of a code is making one of the transitions that is in its representation; a code is represented by all of its illegal transitions. In the xstit framework, we can represent that a code has been transgressed by the truth of the violation constant V. But there is a difficulty that will arise in representing transgressions in terms of the violation constant.

The problem is as follows: Suppose that $\mathfrak{M},(s,h) \vDash V$. Now ask 'which code has been violated?' As we stated before, which code is in force at a dynamic state is a matter of perspective; it could be any set of formulas satisfied at (s,h). Thus, whether the truth of the violation constant is considered a violation of a particular code is also dependant on that perspective. This is a technical decision, and our reason for setting the problem aside will be dealt with in Section 5.1.

---

[6] 'AFT' is an abbreviation of 'after'.

As an abuse of notation that we will use in the following discussion we shorten 'for any $s' \in E(s, h, \mathbf{A})$, and $s' \in h'$, $s', h' \vDash V$', to '$E(s, h, \mathbf{A}) \subseteq [\![V]\!]$'. Although this is an abuse of notation, it is not too misleading because V does not contain X or $[\mathbf{A}\,\mathsf{xstit}]$ operators, so when it is true at a dynamic state it is true at all coincident dynamic states.

We want to characterize normative consistency as the followability of a code. But, as we learned from Hamblin, how followable a code is, is a matter of degree. We should also be careful to distinguish two senses in which a code can be followable.

One sense in which a code is followable is that of being *implementable*, and another sense is when it is *obeyable*. When a code is logically ($\mathcal{L}^I_{\text{xstit}}$-)inconsistent then it is not implementable since it could never be in force, let alone sustained in force. But, as the reader might have guessed, the formal meaning we will give to a code being implementable is that of being sustainable in force.

Definition 6. Δ is implementable iff there is an $\mathcal{L}^I_{\text{xstit}}$-model $\mathfrak{M}$, and $(s, h) \in |\mathfrak{M}|$ such that $\mathfrak{M}, (s, h) \lessdot \Delta$.

This is a generalization of logical consistency: generalizing consistency when $\vdash_{ix}$ is the notion of logical consequence, and the set has to be satisfiable continuously. Since it is a generalization of logical consistency, we will require that every code which is not implementable, must also be normatively *inconsistent*. A code can fail to be non-implementable in a number of ways. $\{\, \mathbf{p}, \neg\mathbf{p} \,\}$ is not implementable since it is not consistent. But $\{\, X\mathbf{p}, \neg\mathbf{p} \,\}$ is not implementable since it could not be satisfied at two consecutive dynamic states. In this investigation we will not worry further about the syntactic make up of implementable codes.

Next we have to consider how obeyable a code can be; that is the central sense of normative inconsistency that we will focus on for the rest of the paper. Following Hamblin's ideas, a code is bad when it puts its subjects in a quandary, i.e., leaves them with no legal transitions. An agent is in a quandary when the code is unobeyable.

4.1. Norm Inconsistency, Semantically

In our framework a code has been obeyed when V is false or no transgression has occurred. Thus, the code *has not* been obeyed when V is true. A group of agents $\mathbf{A}$ would be in a quandary, then, when all of $\mathbf{A}$'s choices lead to violation states. But the mere existence of a situation where an agent/group is in a quandary does not seem damning for a code; it may be possible to make a series of bad decisions and end up in a quandary. But an isolated, independent, and free series of decisions by an agent that leads to a quandary may not be the fault of the code.

If quandaries are too easy to come by, then we can say that the code is not obeyable. Thus we have to provide a precise and formal way of characterizing what it means to say that a quandary is "too easy to come by".

To be in a quandary is to be in a position to only xstit a violation, i.e., $E(s, h, \mathbf{A}) \subseteq [\![V]\!]$ for all h such that $s \in h$. A more limited version of this is simply when $\mathbf{A}$ can xstit V, i.e., $E(s, h, \mathbf{A}) \subseteq [\![V]\!]$ for some h. Let's say that $\mathbf{A}$ is in a 'bad situation' at (s, h) when $E(s, h, \mathbf{A}) \subseteq [\![V]\!]$. A code that could never lead to a bad situation would be a utopian code. So clearly isolated incidents of xstit-ing V is a kind of unproblematic situation, although not ideal. What we should be worried about are codes that give rise to bad situations in too many cases, and now our task is to quantify what constitutes 'too many cases'.

We are attempting to treat normative consistency within logic, independent from deontic logic, and in a formal manner. We take a cue for normative inconsistency from logical inconsistency. For logic, inconsistency means no models make all the formulas true. Thus, 'too many' is interpreted as the extreme or limit case of *all models*. In treating normative inconsistency as a logical property it makes sense to deal with similar extreme cases: all or some.

Thus, normative inconsistency arises when bad situations arise in all models. But in our framework there are two other parameters that should be included when considering how often bad situations arise in all models when a code is sustained in force.

(1) in all models after a code is sustained in force, how frequently do bad situations arise?
(2) in all models after a code is sustained in force, how many groups frequently end up in bad situations?

These questions can be interpreted as asking about properties of classes of xstit models that sustain a code Δ in force. Each of the questions corresponds to a restricted quantifier. The quantifier which corresponds to the 'all models after the code is sustained in force' is (A) $\forall\, \mathfrak{M}, (s, h) \lessdot \Delta$, of course. But the other two questions can be represented as asking about (B) $\forall/\exists(s', h') \in \mathrm{AFT}(s, h)$, and (C) $\forall/\exists \mathbf{A} \subseteq \mathbf{Ag}$. The

(B) quantifier allows us to ask question 1, i.e., after Δ is sustained in force, do bad situations arise at all or only some dynamic states after? The last quantifier (C) allows us to ask whether bad situations arise for all or only some groups? Using these quantifiers we can get a well defined collection of putative quandaries by looking at all of the ways to arrange these quantifiers.

Formally, we can arrange the quantifiers as follows: The (s, h)-values in the kind (A) quantifier depend on the model $\mathfrak{M}$, so we cannot alternate the (A) and (B) quantifiers as in: $\forall(s, h)\forall\mathfrak{M}$. However, the $\mathbf{A}$ variables are not model dependent; they are part of the language. So alternating the type (A) and (C) quantifiers as in $\forall\mathbf{A}\forall\mathfrak{M}$ *is* intelligible. Given the intelligibility of these arrangements of quantifiers, we have the full range of possibilities for quandaries. Each potential quandary will take the form $Q_1, Q_2, Q_3, E(s', h', \mathbf{A}) \subseteq [\![V]\!]$, where Q_1 is either $\forall\mathfrak{M}$ or $\forall\mathbf{A} \subset \mathbf{Ag}$, Q_2 is either $\forall\mathfrak{M}$ or $\forall\mathbf{A} \subset \mathbf{Ag}$ or $\forall(s', h') \in \mathrm{AFT}(s, h)$, and Q_3 is either $\forall\mathbf{A} \subset \mathbf{Ag}$ or $\forall(s', h') \in \mathrm{AFT}(s, h)$. There is a complete, enumerated list — up to obvious logical equivalences obtained by switching adjacent, matching quantifiers — in the appendix.

Each of these conditions represents a potential problem for a code Δ, if the models of Δ satisfy that condition. Each condition offers an interpretation of 'frequently'. The conception of 'frequently' is determined by the different arrangements of quantifiers. However, we can do some pruning to cut away unproblematic conditions, conditions that are arguably too strict. We do this by interpreting the model theoretic conditions in intuitive ways.

We have already excluded any condition that has '$\exists\mathfrak{M}(s, h) \lessdot \Delta$' by analogy with logical inconsistency. But we can see an independent reason to reject conditions starting with '$\exists\mathfrak{M}(s, h) \lessdot \Delta$'. We ask: what if there is a model in which everyone is in a bad situation all of the time? Such a condition is not really problematic, it should be expected. Our intuition is that it should be possible for everyone to be bad all of the time; it should not be an act of logic or of legislation that the code is sometimes obeyed, regardless of the way the world turns out. Agents should be free to disobey the rules imposed on them. Let's consider some other conditions.

Consider the condition:

$$\forall\mathfrak{M}, (s, h) \lessdot \Delta\exists\mathbf{A} \subseteq \mathbf{Ag}, \exists(s', h') \in \mathrm{AFT}(s, h), E(s', h', \mathbf{A}) \subseteq [\![V]\!]$$

we suggest that this condition can be true of the class of models that sustain Δ in force without Δ being problematic. The condition says that in any model, there will be a group who *could* end up in a bad situation. That kind of possibility is to be expected of a set of rules. A code's subjects should run the risk of getting into trouble, if they make illegal decisions. Indeed, this should be the case for *every* group in every model — except $\varnothing$. That means

$$\forall \mathfrak{M}, (s,h) \lessdot \Delta \forall (\varnothing \neq) \mathbf{A} \subseteq \mathbf{Ag}, \exists (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$$

is unproblematic as well.

Also consider the following condition:

$$\exists \mathbf{A} \subseteq \mathbf{Ag}, \forall \mathfrak{M}, (s,h) \lessdot \Delta, \exists (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!] \,.$$

It says that there is a group that in any model can be in trouble. This condition is a bit more worrisome since it is kind of discriminatory. It implies that one particular group must be extra careful and not step out of line. Of course we have groups like that, e.g., police. So a particular group running the risk of getting into trouble should not be a worry.

What makes for worrisome conditions are those which do not allow a code's subjects any choice in getting into trouble. That makes the really worrisome conditions those in which someone is *always* in trouble in every model after the code is sustained in force. The conditions that instantiate that kind of worry are[7]

1. $\forall \mathfrak{M}, (s,h) \forall \mathbf{A} \subseteq \mathbf{Ag}, \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$
3. $\forall \mathfrak{M}, (s,h) \exists \mathbf{A} \subseteq \mathbf{Ag}, \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$
9. $\forall \mathfrak{M}, (s,h) \forall (s',h') \in \mathrm{AFT}(s,h), \exists \mathbf{A} \subseteq \mathbf{Ag}, E(s',h',\mathbf{A}) \subseteq [\![V]\!]$
12. $\exists \mathbf{A} \subseteq \mathbf{Ag}, \forall \mathfrak{M}, (s,h) \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$

In the case of 1, everyone is always in a bad situation. That's bad when it is forced by legislation. Since in every model everyone is in a bad situation, we can say that that condition is due to the elements of the code and not the actions of the code's subjects. Similarly with condition 3. In each model, there is a group that is always in trouble. It is only a completely misanthropic code that would require that somebody must be persecuted. In condition 9 someone/group is always in trouble, although who is in trouble may depend on the dynamic state/situation.

[7] The numbers refer to those found in the appendix.

Again it is a misanthropic code that would force that condition. Finally, condition 12 is also a problem: it is discrimination. One particular group is always in trouble in every model, regardless of the model. That kind of issue can certainly be blamed on the code rather than its subjects.

There are two other conditions that instantiate the same worry as above, but in a very direct manner. Consider:

T1 $\forall\mathfrak{M},(s,h)\forall(s',h')\in\mathrm{AFT}(s,h),\mathrm{lub}(s')\subseteq[\![V]\!]$
T2 $\forall\mathfrak{M},(s,h)\exists(s',h')\in\mathrm{AFT}(s,h),\mathrm{lub}(s')\subseteq[\![V]\!]$

T1 says that in every model, and every situation after the code comes into force is a bad situation. That is clearly undesirable for similar reasons as above. Condition T2 is also, in a sense, undesirable. It says that in *every model*, after the code is in force we are guaranteed some sequence of choices that leads us into a quandary, i.e., leads us into a situation where there are no decisions that allow us to obey the code. Note that T2 is equivalent to

$$\forall\mathfrak{M},(s,h)\lessdot\Delta\forall\mathbf{A}\subseteq\mathbf{Ag},\exists(s',h')\in\mathrm{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$$

Which is condition 2 in the appendix and was considered above, but above we required that the group be non-empty. If condition 2 holds, then for any model that sustains Δ in force at (s,h), every group can, at some point in $\mathrm{AFT}(s,h)$, xstit a violation. So that is true of $\varnothing$, and $\mathrm{lub}(s')=E(s',h',\varnothing)\subseteq[\![V]\!]$ so T2 is true. And if T2 is true, then $E(s',h',\varnothing)=\mathrm{lub}(s')\subseteq[\![V]\!]$, so for any $\mathbf{A}$, $\mathbf{A}$ will xstit a violation at (s',h'). So condition 2 is true as well. Although condition 2 is problematic, it can be folded into T2.

We give these special conditions some names to make them easy to refer to. We have already introduced the T-quandary terminology — total quandaries — for quandaries that result from V being true everywhere. What we will do now is refine this a bit. There is another kind of total quandary that might arise where a particular state s is such that $\mathrm{lub}(s)\subseteq[\![V]\!]$ *after* the code comes into force, whenever it comes into force. This means that there is no way to proceed without everyone possibly ending up in a quandary, and if that follows by legislation that is a problem. We call that a 'TE-quandary' (Total Existential quandary). A T-quandary is condition T1 and a TE-quandary is condition T2. Note that if there is a T-quandary, then there is a TE-quandary.

We introduce another kind of quandary called a global quandary or G-quandary for short. We also introduce a second kind of G-quandary that we call a GE-quandary to parallel the T-types. The G-types are made up of conditions 1 and 9. In condition 9, there is always someone in trouble, potentially anyone can be in trouble, so it is a problem for everyone, i.e., a global problem. Finally, we will call conditions 3 and 12 Discriminatory Quandaries (DE- and D-quandary, respectively). This is because they are similar in that they discriminate against at least one particular group. In 3 that group may depend on the model, in 12 it does not. Again we note that D-quandaries imply DE-quandaries.

So we can finally arrive at a definition for the types of quandaries.

DEFINITION 7. We say that an implementable code Δ has a

- T-quandary iff every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that $\forall s' \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$
- TE-quandary iff every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that $\exists s' \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$
- G-quandary iff every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that $\forall \mathbf{A} \subseteq \mathbf{Ag}, \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$
- GE-quandary iff every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that $\forall (s',h') \in \mathrm{AFT}(s,h), \exists \mathbf{A} \subseteq \mathbf{Ag}, E(s',h',\mathbf{A}) \subseteq [\![V]\!]$
- D-quandary iff $\exists \mathbf{A} \subseteq \mathbf{Ag}$ such that for every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$, $\forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$
- DE-quandary iff every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that $\exists \mathbf{A} \subseteq \mathbf{Ag}, \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$

Now we can raise two questions about what kinds of codes lead to G, T, or D-type quandaries. The definitions of the quandaries were given in terms of conditions on models, so we may want to attempt to characterize those conditions by what those codes can prove. In Section 4.2 we ask: what conditions on the consequences of codes correspond, if any, to the quandaries? The second question to raise is whether there are connections between the kinds of quandaries.

As it turns out, the conditions on quandaries are very similar given the formalism that we are working with. We can connect them according to the following theorem.

THEOREM 1. *Given an implementable code Δ, the following conditions are equivalent*:

- Δ *has a G-quandary*
- Δ *has a GE-quandary*
- Δ *has a D-quandary*

- Δ *has a DE-quandary*
- Δ *has a T-quandary.*

This is proved in the appendix. The upshot of this theorem is that the only difference between the quandaries, really, is 5/6 of them on one side and TE on the other. But the other five all imply TE.

The big, counterintuitive jumps in proving Theorem 1, like from D to T or GE to G, that cause the collapse are forced because of the anti-monotonicity of the effectivity function, the fact that $E(s, h, \mathbf{Ag}) = \{\operatorname{lub}(s, h)\}$ for any $s \in h$, and that if $\mathfrak{M}, (s, h) \vDash V$, then $\mathfrak{M}, (s, h') \vDash V$ for all h' with $s \in h'$. We will discuss more about the technical issues and what they mean in Section 5.1.

From the semantic side of things we can define the consistency of a normative system as quandary freeness, like Hamblin, but all that we need to ensure quandary freeness is that there is no TE-quandary. Thus we are in a position much like that of von Wright's condition. However, it does not reduce simply to regular inconsistency since to have a TE-quandary the code must be $\mathcal{L}^{I}_{\mathrm{xstit}}$-consistent.

4.2. Normative Inconsistency, Syntactically

We can move on to our question about characterizing the quandaries in terms of provability. Consequence can usually be related to consistency, i.e., $\Gamma \vdash \varphi$ iff $\Gamma \cup \{\neg\varphi\}$ is inconsistent. With normative inconsistency and consequence, it is different. The normative consistency of a set Δ connects to $\mathcal{L}^{I}_{\mathrm{xstit}}$-consequence indirectly. But we can show a tight relationship between the $\mathcal{L}^{I}_{\mathrm{xstit}}$-consequences of a set related to a code and the quandaries. We first have to build up some technical results to support our investigation. The proofs for the results are in the appendix.

We first define a set Δ_{if} for 'Δ-in force'. Δ_{if} is constructed as follows:

$$\Delta_{if} = \{\Box X^n \delta \mid \delta \in \Delta \;\&\; n \in \mathbb{N}\}.$$

So for any $\delta \in \Delta$, we have $\Box\delta \in \Delta_{if}$, $\Box XXXX\delta \in \Delta_{if}$, $XXXX\delta \in \Delta_{if}$ and $\delta \in \Delta_{if}$, for instance. Now we can make an observation:

OBSERVATION 1. $\mathfrak{M}, (s, h) \lessdot \Delta$ *if and only if* $\mathfrak{M}, (s, h) \vDash \Delta_{if}$.

To continue we require one more lemma.

LEMMA 1. *Suppose that there is a* (s, h) *in* $\mathfrak{M}$*, such that* $\mathfrak{M}, (s, h) \vDash \Delta_{if}$*, and* $(s', h') \in \mathrm{AFT}(s, h)$*. Then* $\mathfrak{M}, (s', h') \vDash \Delta_{if}$

Now that we have these results we can show how to define conditions on provability that will translate into quandaries and back. Because we have Theorem 1, we really only need two results, one for T and one for TE. However, we can provide more specific results for most of the types of quandaries. We collect the conditions together in Theorem 2. Recall that the set $\mathbf{Ag}$ is finite, and so $\mathcal{P}(\mathbf{Ag})$ is finite. As another bit of notational convenience, let $\Delta_V = \{ \neg\Diamond X^n \Box XV \mid n \in \mathbb{N} \}$.

Theorem 2. *If Δ is a code, $\vdash = \vdash_{ix}$, and $\Delta_{if} \nvdash \bot$, then*

- *Δ has a T quandary iff $\Delta_{if} \vdash \Box XV$*
- *Δ has a TE quandary iff $\Delta_{if} \cup \Delta_V \vdash \bot$*
- *Δ has a G quandary iff $\Delta_{if} \vdash \bigwedge_{\mathbf{A} \in \mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\, V$*
- *Δ has a GE quandary iff $\Delta_{if} \vdash \bigvee_{\mathbf{A} \in \mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\, V$*
- *Δ has a D quandary iff $\Delta_{if} \vdash [\mathbf{A}\,\mathsf{xstit}]\, V$ for some $\mathbf{A} \in \mathcal{P}(\mathbf{Ag})$*
- *Δ has a DE quandary iff ?*

There is not a clear condition that matches up for the DE case. But because of the equivalence of DE to D, we do not have to worry about representing the condition here. Now we are finally in a position to provide a definition of normative consistency for our system.

Clearly normative consistency CON_N will have something to do with the existence of quandaries. Fortunately, in the current situation, if there is a quandary of any type (T,G,D), then there is a TE-quandary. But that means if we say that a code does not have a TE-quandary, then it will not have any of the other problematic quandaries either. But there is also the situation of standard, or $\mathcal{L}^I_{\text{xstit}}$-inconsistency, i.e., $\Delta \vdash_{ix} \bot$. A code that is simply inconsistent is rather problematic, at least from a logical standpoint.

A code Δ that is inconsistent can not be satisfied at all, so it can not do its job in our setting, by completeness. But it is also a problem if a code is not implementable, i.e., $\Delta_{if} \vdash_{ix} \bot$. As we noticed in Observation 1, the consistency of Δ_{if} is necessary and sufficient for Δ to be implementable. So if $\Delta_{if} \vdash_{ix} \bot$, then it is no good, normatively speaking. By the monotonicity of $\vdash_{ix}$, if $\Delta \vdash_{ix} \bot$, then $\Delta_{if} \vdash_{ix} \bot$. That means the consistency of Δ_{if} implies the usability/implemtability of Δ. Thus part of being a normatively consistent code is for Δ_{if} to be consistent. The other part, as you might have guessed, is for there to be no TE-quandary.

When there is a TE quandary, it follows that $\Delta_{if} \cup \Delta_V \vdash_{ix} \bot$. If $\Delta_{if} \vdash_{ix} \bot$, then by monotonicity of $\vdash_{ix}$, $\Delta_{if} \cup \Delta_V \vdash_{ix} \bot$. So if there is no TE-quandary, then $\Delta_{if} \nvdash_{ix} \perp$. Thus, if no TE-quandary, Δ is normatively consistent. Ultimately, we have the following definition.

Definition 8. A code Δ is normatively inconsistent iff it has a TE-quandary. So we can define normative inconsistency $\overline{\text{CON}_N}(\Delta)$ formally as follows

$$\overline{\text{CON}_N}(\Delta) \Longleftrightarrow [\Delta_{if} \cup \Delta_V \vdash_{ix} \perp]$$

5. Discussion

There are a number of questions any formal work in philosophy must address. In the case of $\mathcal{L}^I_{\text{xstit}}$, we need to ask two of particular interest. 1) Given that there is a counterintuitive collapse of the quandary conditions, which relies on the formal framework, how appropriate is that formal framework? Or put another way: how realistic are the assumptions of the xstit framework? We will deal with that in Section 5.1. 2) Are the characterizations of a quandary and normative consistency any good? Are they reasonably presented? We will deal with that largely in Section 5.2.

5.1. Technical Issues

In answering question 1 above we will deal with four topics. First, we will consider the semantics of the xstit operator, particularly the conditions that are placed on the effectivity functions. Second, we will consider variations of the quandary conditions, focusing on how groups are portrayed. Third, we will consider how responsibility for action is represented, and how it interacts with quandaries. Finally, we will explain some technical difficulties with representing when a code is violated at a dynamic state.

The semantics of xstit is supposed to represent the effects of choices in an evolving world. However, some people may take issue with the assumptions of the model. In the xstit models, time is discrete. Although mathematical models of time tend to be continuous, the intended range of applications for Xstit models are systems that evolve by discrete steps. Indeed, it is basic in one of the most commonly used mathematical models in economics: game theory.

Of real interest are the the assumptions 3e, 3f, and 3g on the effectivity function. Condition 3e requires that all of the members of **Ag** can direct the course of events very specifically. It says that when **Ag** gets together at (s, h), and decides to make the decision indexed to h, they can guarantee that the next state is $\text{lub}(s, h)$. Usually people are not so powerful, things could go awry.

We must pause to note where these assumptions come from. Effectivity functions come from cooperative game theory, and they are a representation of the power that coalitions have in a game. These conditions guarantee what is called a 'playable effectivity function' for the game, see [18]. In a game, the outcome is completely decided by all of the decisions of the players. But what we want is a slightly more general formalism: discrete, but still leaves room for indeterminacy. Thus we may want to drop that assumption.

In the proof of Theorem 1, however, all that is needed when using condition 3e is condition 3c which provides us with $\text{lub}(s, h) \in E(s, h, \mathbf{Ag})$. Thus, dropping 3e leaves the results intact.

Condition 3f, the anti-monotonicity condition, is tricky. It is tricky because we have to ask what $E(s, h, \mathbf{A})$ really means. As we have discussed above in Section 3.2 we do not deal with how responsibility is distributed over a collective action, but what collective choice is does play a role here. $E(s, h, \mathbf{A})$ represents $\mathbf{A}$'s choice, the choice indexed to h, at the state s. But it is odd to think of choices in this way since, in the models, there are choices indexed to h at lots of states. What intuitive aspect of choice would that correspond to? The only thing that makes sense in this case is if h were some sort of plan, or making the choice relative to h would be to adhere to a specific plan. But then 3f requires a substantive assumption about the nature of cooperation. It requires that larger groups be more effective in achieving their plans.

Is that assumption reasonable? Maybe not. Remember, too many cooks spoil the broth. The lesson of that idiom is that it is difficult to get people to cooperate. But the difficulty involved in getting people to cooperate is not part of this model. All that is represented is successful cooperation. Intuitively, the more people who successfully cooperate, it seems, the better chance they have at achieving their goal. So maybe 3f is not unreasonable.

Condition 3g:

- if $\mathbf{A} \cap \mathbf{B} = \varnothing$ and $s \in h \cap h'$, then there is h'' with $s \in h''$ and $E(s, h'', \mathbf{A})$ and $E(s, h'', \mathbf{B})$ are contained in $E(s, h, \mathbf{A})$ and $E(s, h', \mathbf{B})$, respectively.

says that if disjoint groups are effective for some outcomes, they can not interfere with each other's effectiveness for those outcomes: 'outcome' in this case means 'proposition'. The way that is represented in this case is that both disjoint groups have choices indexed to the same history that have similar outcomes to their original choices. This condition also raises questions about how to interpret $E(s,h,\mathbf{A})$. If we interpret $E(s,h,\mathbf{A})$ as $\mathbf{A}$ acting according to h as a plan, then 3g says that there is a common plan for any two disjoint groups to achieve similar outcomes as their original choices. But that seems too strong of an assumption to make.

The intuitive idea behind 3g, which comes from regular stit theory, seems reasonable. If **a** is really effective to bring about a proposition, then no one else should be able undermine that. How that is represented in this model is by 3g, and if we think Indep-G, is an acceptable representation of the intuitive idea, 3g is the reasonable condition — the canonical model demonstrates that fact. At this point, it does not seem reasonable to dismiss the formalism as flawed.

Next we will deal with one aspect of the quandary conditions. The quandary conditions refer to groups simply as $\mathbf{A} \subseteq \mathbf{Ag}$. In the proofs of the results, it seems like cheating to allow $\varnothing$ to be a group. One might object that such a group is degenerate: it does not have any members! It may also seem like cheating to allow $\mathbf{A} = \mathbf{Ag}$. The, perhaps, surprising thing is that Theorem 1 is still provable when the group quantifier is replaced with $QA : \varnothing \neq A \neq \mathbf{Ag}$.

More explicitly:

Theorem 3. *For an implementable code Δ, the following are equivalent:*

G2 *Every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that* $\forall \mathbf{A} \subseteq \mathbf{Ag}(\varnothing \neq A \neq \mathbf{Ag}), \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$

GE2 *Every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that* $\forall (s',h') \in \mathrm{AFT}(s,h), \exists \mathbf{A} \subseteq \mathbf{Ag}(\varnothing \neq A \neq \mathbf{Ag}), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$

D2 $\exists \mathbf{A}(\varnothing \neq A \neq \mathbf{Ag}) \subseteq \mathbf{Ag}$ *such that for every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$,* $\forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$

DE2 *Every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that* $\exists \mathbf{A} \subseteq \mathbf{Ag}(\varnothing \neq A \neq \mathbf{Ag}), \forall (s',h') \in \mathrm{AFT}(s,h), E(s',h',\mathbf{A}) \subseteq [\![V]\!]$

Proof. Here we will provide a proof sketch. The reason is that all of these are equivalent to their corresponding original quandaries. Immediately we have the following relationships: G implies G2, D2 implies D, GE2 implies GE, and DE2 implies DE. For the other directions, any time one of the original quandaries is assumed to hold, there is a T-quandary. That means for any group, at any dynamic state after Δ is sustained in force, $E(s', h', \mathbf{A}) \subseteq [\![V]\!]$. To show that G2 implies G, assume G2 and suppose that $\mathfrak{M}, (s, h) \lessdot \Delta$ and $(s', h') \in \mathrm{AFT}(s, h)$. From G2 there is $\mathbf{A} \neq \varnothing$ and $\mathbf{A} \neq \mathbf{Ag}$, such that $E(s', h', \mathbf{A}) \subseteq [\![V]\!]$. And by condition 3f $E(s', h', \mathbf{Ag}) \subseteq E(s', h', \mathbf{A})$. But that is the case for every h'' such that $s' \in h''$ and any set of agents. But,

$$\bigcup_{h'' :\ s' \in h''} E(s', h'', \mathbf{Ag}) = \mathrm{lub}(s') \subseteq [\![V]\!].$$

Therefore, there is a T-quandary, and by Theorem 1, a GE-quandary. $\square$

Technically, that means there was no cheating going on in the original definitions of the quandaries in order to get the result. The counterintuitive collapse of all these distinctions is not found in the definitions. The problem, if we can call it that, is with the formal framework itself. Where we place the blame is with the Anderson-style reduction. Although we have offered particular ways of representing certain kinds of norms using the reduction, the results do not hang on those representations, nor on the conditional used (provided the logic is complete). What matters is that the violation constant is impersonal. The obvious solution would be to use violation constants for each agent: $V_{\mathbf{a}}$. These would say '**a** is in violation'. We will return to why this is dubious in the next section. Next we will discuss the technical characterization of responsibility in relation to the quandaries, and how violations need not be indexed to a code.

What about responsibility? If a code has a T or TE-quandary (a T-type quandary), then no one/group is properly responsible for the violations. Indeed, no one need be responsible in the sense of xdstit for any violation. If a code has a T-type quandary, there is definitely a problem, but we can not hold anyone responsible for the violation. There is no way for any group, or individual, to have done otherwise, or even to have the violation not occur.

Does that mean these quandaries are meaningless? No. The thought that responsibility is important for wrongdoing is a conceptual one to do with the law or morality. Indeed, problems with codes *do not* concern people's actions or choices, they concern problems with how the code is put together. The code does not provide opportunities for agents to act that will avoid violations when there is a T-quandary: the code is completely problematic. The fault lies with the code or — at best — the code's makers; the fault does not lie with the code's patients.

As promised in Section 2, we wanted to explain why we did not index the violations to a code. First, we only wanted to consider one code in this framework to keep things simple. But as we asked before: how might we attribute a violation as arising from one set of formulas true at a dynamic state from another?

The obvious solution is to say that V arises from Δ at (s, h), if there is a set $\Gamma \subset \{\, \varphi : \mathfrak{M}, (s, h) \vDash \varphi \,\}$, and $\Delta \cup \Gamma \vdash_{ix} V$. But that will not do since either $V \in \{\, \varphi : \mathfrak{M}, (s, h) \vDash \varphi \,\}$ or not, so this definition does not give us any assistance. But we could make a refinement.

What we want is that the violation results because of some set of, social and/or natural, facts. Of course, facts are represented by formulas. Thus as long as the set of facts added to Δ does not mention V, we might get something that we want. Denote by $at(\Gamma)$ the set of atoms used in formulas in Γ. Then we can define when Δ has been violated as follows:

Definition 9. Δ is violated in the model $\mathfrak{M}$, at $(s, h) \in |\mathfrak{M}|$ iff there is a set $\Gamma \subseteq \{\, \varphi : \mathfrak{M}, (s, h) \vDash \varphi \,\}$ such that $V \notin at(\Gamma)$ and $\Delta \cup \Gamma \vDash_{ix} V$.

There are two potential problems with this definition. First, and least seriously, is that it requires information about all models to be used to evaluate something within one model. Second, it is not clear how using this definition will affect the technical results.

In future work we will include this as part of the definition of what it means for a code to be sustained in force. What is clear about such an adjustment is that it will not affect Theorem 1, but it will affect our syntactic characterization. That requires a representation of when a code can be sustained in force, and when the definition is refined in this way, the connection seems to break down.

5.2. Philosophical Issues

We next turn to the question of whether this formulation of normative consistency is any good. We have already dealt with that question some, at least whether it is technically sound, in the last section. But now we can ask whether it is philosophically sound. One way to approach this is to see if anyone else has had a similar idea. Another is to consider other representations of quandaries. It turns out that a similar notion is contained in [17]. But first we will look at an alternative formulation of quandaries suggested by an anonymous referee to a previous iteration of this paper.

In our version of a quandary, we represent the failure to obey by $E(s, h, \mathbf{A}) \subseteq [\![V]\!]$. But a failure to obey could also be represented as a failure to guarantee non-violation. In similar notation that would be $E(s, h, \mathbf{A}) \cap [\![V]\!] \neq \varnothing$; that condition is what we now mean by 'bad situation'. Thus in each of the original quandaries we simply replace $E(s, h, \mathbf{A}) \subseteq [\![V]\!]$ with $E(s, h, \mathbf{A}) \cap [\![V]\!] \neq \varnothing$, except in the case of the T-quandaries, $\text{lub}(s) \subseteq [\![V]\!]$ is replaced by $\text{lub}(s) \cap [\![V]\!] \neq \varnothing$. If X is the name of an original quandary, X3 will denote the new quandary. Since $E(s, h, \mathbf{A})$ is never empty when $s \in h$, any of the original quandaries implies any of the new quandaries. Thus D will imply D3, for example.

The T-quandaries in this case cease to be problematic. Each of them now says that a violation follows every state where the code is in force. But that is not worrisome; it is commonplace for there to be *some* way to break the rules. We can give similar justifications for keeping the other 3-versions of the quandaries as was done in Section 4.1. What is more interesting, however, is how the remaining quandaries relate to one another. It is left as an exercise to demonstrate the following:

THEOREM 4. *For an implementable code* Δ*, the following relationships hold*:

- D3 *is equivalent to* DE3*,*
- DE3 *implies* GE3*,*
- G3 *implies* GE3*,*
- G3 *implies* D3.

In this case, there is no collapse of the quandaries, at least not entirely. And from Theorem 4 it follows that as long as there is no GE3-quandary, none of the other quandaries obtain. So we could define normative consistency in terms of the non-existence of a GE3-quandary. But we should not be overzealous. Let's consider that condition. The GE3 quandary is:

GE3 Every model $\mathfrak{M}$ with (s, h) such that $\mathfrak{M}, (s, h) \lessdot \Delta$ is such that $\forall (s', h') \in \text{AFT}(s, h), \exists \mathbf{A} \subseteq \mathbf{Ag}, E(s', h', \mathbf{A}) \cap [\![V]\!] \neq \varnothing$, and so consistency would be its negation: either not implementable or

∼GE3 Some model $\mathfrak{M}$ and (s, h) such that $\mathfrak{M}, (s, h) \lessdot \Delta$, and a $(s', h') \in \text{AFT}(s, h)$ where for every $A \subseteq \mathbf{Ag}, E(s', h', \mathbf{A}) \cap [\![V]\!] = \varnothing$.

But ∼GE3 is far too weak of a condition for consistency, intuitively speaking. If there is merely some state in some model where all of the groups can guarantee non-violation, it could be that it is only in that particular circumstance that the code can be obeyed. We should expect better from a consistent code. Thus we reject that characterization of normative consistency. But there is an alternative reason for sticking with our original conception of normative consistency, at least for the moment.

In her [17], Marcus introduces a notion of consistency for moral rules. She says "we can define a set of rules as consistent if there is some possible world in which they are all obeyable in all circumstances in *that* world" (p. 128). Marcus makes clear that what she is after is not a sufficient condition for a set of rules to be moral or even morally binding. At most it is a necessary condition on a morally binding set of rules/norms. We have taken an even more liberal approach by detaching questions of consistency from those of whether a general set of norms, not just moral norms, is *in force*.

Why does Marcus choose this definition? She is drawing an analogy with logical consistency. A set of sentences Γ is logically consistent iff there is a possible world in which each member of Γ is true. For Marcus, then, the conditions of a norm being obeyable is analogous to a declarative sentence being true. She also gives a reason for using *obeyable* rather than *obeyed*:

> I want to allow for the partition of cases where a rule-governed action fails to be done between those cases where the failure is a personal failure of the agent — an imperfect will in Kant's terms — and those cases where "external" circumstances prevent the agent from meeting conflicting obligations. To define consistency relative to a kingdom of ends, a deontically perfect world in which all actions that ought to be done are done, would be too strong; for that would require both perfection of will and the absence of circumstances that generate moral conflict. [17]

Marcus wants to define moral consistency in a way that may allow moral dilemmas to arise even if a system of moral norms is consistent. What is crucial is that consistency does not require the possibility of avoiding moral conflict. A set of moral norms can be consistent even if people do not follow the norms. Even if people a long long time ago in a galaxy far far away are able to obey the norms, the norms are consistent. This notion of consistency may not satisfy many moral philosophers since the notions of consistency they might prefer are those to do with *this* world, and *these* people. However, that is beside the point for the moment; logical investigations should start with general investigations and try to find the lowest common denominator, then try to capture what specific philosophers want.

Now that Marcus' definition has been introduced and her reasons behind the definition's formulation been mentioned, it is time to analyze more closely the ideas in the definition. First, we will stipulate that the concern is with rules that can be represented within this formal framework. Does that include moral norms? That is an argument we will set aside. Recall Marcus' definition: Γ is normatively consistent iff there is some possible world in which they(Γ) are all obeyable in all circumstances in *that* world. The concepts included in the definition that need explication are as follows:

- How should we understand rules/norms?
- How should we understand 'possible world'?
- How should we understand 'all circumstances'?
- How should we understand 'Obeyable'?

In our discussion, we have offered ways of understanding each of these questions. Rules/norms are represented by sets of $\mathcal{L}^I_{\text{xstit}}$-formulas, possible worlds are represented by $\mathcal{L}^I_{\text{xstit}}$-models, circumstances are represented by static states in the models, and so we are just left with representing how to obey the norms.

But representing 'obeyable' requires a choice of perspective. Our version of obeying from above can be characterized as follows: $\mathbf{A}$ has obeyed the code at (s,h) when $E(s,h,\mathbf{A}) \not\subseteq [\![V]\!]$, let's call this version A of obeying. So $\mathbf{A}$'s choice does not guarantee a violation state. This does not rule out that a violation state could occur, it might, but we need not say that $\mathbf{A}$ did not obey. It was out of $\mathbf{A}$'s control; the world just did not turn out in $\mathbf{A}$'s favour. However, someone might suggests that merely not guaranteeing a violation is not sufficient to obey; in order to obey a code violation must be avoided completely. Thus $\mathbf{A}$ obeys at (s,h) when $E(s,h,\mathbf{A}) \subseteq [\![\neg V]\!]$ or, equivalently, $E(s,h,\mathbf{A}) \cap [\![V]\!] = \varnothing$. Let's call this version B of obeying.

Our version of normative consistency is the negation of aTE-quandary

$\sim$TE There is a model $\mathfrak{M}$ and (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ in which for all $s' \in \text{AFT}(s,h)$, $\text{lub}(s') \cap [\![\neg V]\!] \neq \varnothing$

i.e., there is a model in which there is always a non-violation state that follows any state at which Δ is in force. With $\sim$TE we can accommodate both versions of obeying, but in different ways. Suppose that $\mathfrak{M},(s,h) \lessdot \Delta$ is a model which is a witness to the normative consistency of Δ. At any state $s' \in \text{AFT}(s,h)$, there is $s'' \in \text{lub}(s')$ such that $s'' \notin v(V)$. Thus, by condition 3f on effectivity functions, $E(s',h'',\mathbf{Ag}) \subseteq [\![\neg V]\!]$

where $\mathrm{lub}(s', h'') = s''$. That means $\mathbf{Ag}$ is able to obey Δ at any state according to version B.

On the other hand, in the same model as above, given any group $\mathbf{A}$, $s'' \in E(s', h'', \mathbf{A})$ by condition 3c, so $E(s', h'', \mathbf{A}) \not\subseteq [\![V]\!]$. Therefore, according to version A, every group can obey at each state. It might be objected that $\mathbf{Ag}$'s ability to avoid violation should not confer obeyability since it may take the whole group to avoid violation and that is too restrictive. Indeed, legislators may have designed a code in a communistic manner. The legislators may think that it is our responsibility to cooperate in our choices to such a degree that it require the cooperation of all of society to obey the rules they lay down. Thus our formal framework can make sense of Marcus' semi-formal account of normative/moral consistency. Within our framework, however, we can see a number of ambiguities which can be parsed apart. Formalism that allows us to unambiguously see hidden distinctions has been held up as a value, if not *the* value, of formal work since the time of Russell's "On Denoting".

Finally, as promised, we return to the issue of individualised violation constants. The frameworks that employ Anderson-style reductions provide a formal language in which one could formulate the rules for a system. If the rules have to be specified in such a way that they include who is in violation, that seems odd. Whether someone is in violation of a code depends on what they did; packing that into one sentence removes, we think, the interest in presenting a model which includes action. Including specific violation constants would beg questions against distributions of responsibility for collective effects. We could introduce violation constants for each group of agents as $V_{\mathbf{A}}$ in an effort to avoid begging those questions, but then the relationship between group and subgroup violation is also at issue. We prefer to take the more philosophically neutral position of using an impersonal violation constant. What this analysis demonstrates is that, although many distinctions can be explicated, more refinement is needed to get things right. But that is the case with most formal analyses.

6. Conclusion

In this paper we have combined intuitive ideas about normative inconsistency and previous formal representations of them, then shown how to represent these informal/formal notions in an xstit framework. We have also characterized the notion of normative inconsistency in both the semantic and syntactic setting of the extended xstit language $\mathcal{L}^I_{\text{xstit}}$. This characterization leads to a simple definition of normative inconsistency making good on one of the lessons learned from von Wright: try to combine separate notions of normative consistency into one. But we have also been able to distinguish normative consistency from logical consistency, at least in relation to $\mathcal{L}^I_{\text{xstit}}$-consistency through Definition 8 and Theorem 2. This maintains our basic intuition that they are separate notions, and we can see their exact relationship once we have fixed a formal framework.

Any intuitive notion will take on certain characteristics of the formal framework in which it is explicated, such is the nature of explication. This framework offers, we believe, interesting further avenues of research. Particularly around notions of obeyability. Perhaps other versions of quandaries/inconsistency can be uncovered.

Technical Appendix

Below is a list of the conditions. But we have omitted one of two equivalent conditions where two universal quantifiers or two existential quantifiers are adjacent. For example: if the condition is of the form $\forall\mathfrak{M}\forall\mathbf{A}\varphi$, we keep that formulation, and leave out the equivalent $\forall\mathbf{A}\forall\mathfrak{M}\varphi$.

1. $\forall\mathfrak{M},(s,h)\forall\mathbf{A}\subset\mathbf{Ag},\forall(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
2. $\forall\mathfrak{M},(s,h)\forall\mathbf{A}\subset\mathbf{Ag},\exists(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
3. $\forall\mathfrak{M},(s,h)\exists\mathbf{A}\subset\mathbf{Ag},\forall(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
4. $\forall\mathfrak{M},(s,h)\exists\mathbf{A}\subset\mathbf{Ag},\exists(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
5. $\exists\mathfrak{M},(s,h)\forall\mathbf{A}\subset\mathbf{Ag},\forall(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
6. $\exists\mathfrak{M},(s,h)\forall\mathbf{A}\subset\mathbf{Ag},\exists(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
7. $\exists\mathfrak{M},(s,h)\exists\mathbf{A}\subset\mathbf{Ag},\forall(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
8. $\exists\mathfrak{M},(s,h)\exists\mathbf{A}\subset\mathbf{Ag},\exists(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
9. $\forall\mathfrak{M},(s,h)\forall(s',h')\in\text{AFT}(s,h),\exists\mathbf{A}\subset\mathbf{Ag},E(s',h',\mathbf{A})\subseteq[\![V]\!]$
10. $\exists\mathfrak{M},(s,h)\forall(s',h')\in\text{AFT}(s,h),\exists\mathbf{A}\subset\mathbf{Ag},E(s',h',\mathbf{A})\subseteq[\![V]\!]$
11. $\exists\mathbf{A}\subset\mathbf{Ag},\forall\mathfrak{M},(s,h)\exists(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$
12. $\exists\mathbf{A}\subset\mathbf{Ag},\forall\mathfrak{M},(s,h)\forall(s',h')\in\text{AFT}(s,h),E(s',h',\mathbf{A})\subseteq[\![V]\!]$

The possible T-quandaries are

T1 $\forall\mathfrak{M},(s,h)\forall(s',h')\in\text{AFT}(s,h),\text{lub}(s')\subseteq[\![V]\!]$

T2 $\forall\mathfrak{M}, (s,h)\exists(s',h') \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$
T3 $\exists\mathfrak{M}, (s,h)\forall(s',h') \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$
T4 $\exists\mathfrak{M}, (s,h)\exists(s',h') \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$

PROOF OF THEOREM 1. We will proceed by showing that T-quandaries implies G-quandaries and D-quandaries, then that G-quandaries imply T-quandaries and D-quandaries imply T-quandaries. Since it is trivial that in each type of quandary X, X implies XE for X=T,G,D, we complete the proof by showing that DE-quandaries imply GE-quandaries, and GE-quandaries imply G-quandaries.

It is easy to see that if Δ has a T-quandary, then Δ has both G and D-quandaries. Let $\mathfrak{M}$ be any model of Δ such that $(s,h) \in |\mathfrak{M}|$, and $\mathfrak{M}, (s,h) \lessdot \Delta$. If Δ has T quandary, all states after the code is sustained in force only have successor states that are V states, i.e., for all $(s',h') \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$. So whatever any group is effective for is going to be a set of V states, i.e., for all $A \subseteq \mathbf{Ag}$, and $(s',h') \in \mathrm{AFT}(s,h)$, $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$ since $E(s',h',\mathbf{A}) \subseteq \mathrm{lub}(s')$. Since $\mathfrak{M}$ was arbitrarily chosen, Δ has a G-quandary. Under the same condition of Δ having a T-quandary, for any $\mathfrak{M}, (s,h) \lessdot \Delta$, and $(s',h') \in \mathrm{AFT}(s,h)$, $E(s',h',\varnothing) = \mathrm{lub}(s') \subseteq [\![V]\!]$. But that will be true for $\mathbf{A} = \varnothing \subseteq \mathbf{Ag}$ for any model, so Δ has a D-quandary.

Now we show that G implies T. Suppose Δ has a G-quandary. Then in any model with (s,h) such that $\mathfrak{M}, (s,h) \lessdot \Delta$, every group is such that it is always effective for violations, i.e., for all $\mathbf{A} \subseteq \mathbf{Ag}$, and $(s',h') \in \mathrm{AFT}(s,h)$, $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$. But that means $\varnothing$ is effective for violations, i.e., $E(s',h',\varnothing) \subseteq [\![V]\!]$ for all $(s',h') \in \mathrm{AFT}(s,h)$. But that is just to say that $\mathrm{lub}(s') \subseteq [\![V]\!]$ by condition d on effectivity functions. Since $\mathfrak{M}$ was arbitrarily, chosen Δ has a T-quandary.

To show D implies T suppose Δ has a D-quandary, i.e., there is $\mathbf{A}$ such that for any model with (s,h) such that $\mathfrak{M}, (s,h) \lessdot \Delta$, for all $(s',h') \in \mathrm{AFT}(s,h)$, $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$. But then by the anti-monotonicity of E and because $\mathbf{A} \subseteq \mathbf{Ag}$, $E(s',h',\mathbf{Ag}) \subseteq [\![V]\!]$. But that holds for every $s' \in h'$. Let $s'' \in \mathrm{lub}(s')$, arbitrary s' from $\mathrm{AFT}(s,h)$. Then there is h'' such that $s'' \in h''$ and (s'',h'') is in $\mathrm{AFT}(s,h)$ since s' is. But also $\mathrm{lub}(s',h'') = (s'',h'')$, and $E(s',h'',\mathbf{Ag}) \subseteq [\![V]\!]$ (because it holds for all elements of $\mathrm{AFT}(s,h)$). But then $(s'',h'') \vDash V$, so $s'' \in [\![V]\!]$. Since s'' was arbitrary, $\mathrm{lub}(s') \subseteq [\![V]\!]$, i.e., Δ has a T-quandary.

From DE to GE is simple since if for each model with (s,h) such that $\mathfrak{M}, (s,h) \lessdot \Delta$, there is an $\mathbf{A}$, call it $\mathbf{B}$, such that for any $(s',h') \in$

$\mathrm{AFT}(s,h)$, $E(s',h',\mathbf{B}) \subseteq [\![V]\!]$, then for any $(s',h') \in \mathrm{AFT}(s,h)$ there is an $\mathbf{A}$, viz. $\mathbf{B}$, such that $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$, i.e., Δ has a GE-quandary

So we are left with showing that GE implies G. Suppose that for each model with (s,h) that $\mathfrak{M},(s,h) \lessdot \Delta$, and any $(s',h') \in \mathrm{AFT}(s,h)$, there is $\mathbf{A} \subseteq \mathbf{Ag}$, such that $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$, i.e., Δ has a GE-quandary. We want to show that Δ has a G-quandary, that is for any $\mathfrak{M},(s,h) \lessdot \Delta$, and $\mathbf{A} \subseteq \mathbf{Ag}$, and $(s',h') \in \mathrm{AFT}(s,h)$, $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$. So let $\mathfrak{M},(s,h) \lessdot \Delta$, and $\mathbf{B} \subseteq \mathbf{Ag}$ and $(s',h') \in \mathrm{AFT}(s,h)$. We will show that $\mathrm{lub}(s') \subseteq [\![V]\!]$. That will imply $E(s',h',\mathbf{B}) \subseteq [\![V]\!]$, and since $\mathbf{B}$ and (s',h') were chosen arbitrarily, Δ will have a G-quandary.

Suppose that $h'' \ni s'$. Then $(s',h'') \in \mathrm{AFT}(s,h)$, so there is $\mathbf{A}'$ such that $E(s',h'',\mathbf{A}') \subseteq [\![V]\!]$. By the anti-monotonicity of E, $E(s',h'',\mathbf{Ag}) \subseteq [\![V]\!]$ so $\mathrm{lub}(s',h'') \in [\![V]\!]$. Since h'' was arbitrarily chosen, $\mathrm{lub}(s') \subseteq [\![V]\!]$ as we wanted. Thus Δ has a G-quandary.

Just for fun we can show that GE implies DE. If GE is true, then for each model with (s,h) that $\mathfrak{M},(s,h) \lessdot \Delta$, and any $(s',h') \in \mathrm{AFT}(s,h)$, there is $\mathbf{A} \subseteq \mathbf{Ag}$, such that $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$. As we noticed, that means $E(s',h',\mathbf{Ag}) \subseteq [\![V]\!]$, for each $(s',h') \in \mathrm{AFT}(s,h)$. But then T is true, and so D is true and D implies DE. $\square$

Proof of Observation 1. Suppose that $\mathfrak{M},(s,h) \vDash \Delta_{if}$. Note that $\mathfrak{M},(s,h) \vDash \Delta$ since $\Box[\Delta] \subseteq \Delta_{if}$ and $\Box[\Delta] \vdash_{ix} \delta$ for each $\delta \in \Delta$, and $\vdash_{ix}$ is sound. Then let (s',h') be "after" (s,h), so $s,s' \in h'$, and let $\delta \in \Delta$. Suppose, without loss of generality, that s' is the n-successor of h' from s. Then $\Box X^n\delta \in \Delta_{if}$, by definition. But that means $(s,h) \vDash \Box X^n\delta$, $(s,h') \vDash X^n\delta$ and $(s',h') \vDash \delta$, also by definition. Since δ was arbitrarily chosen, $(s',h') \vDash \Delta$. (s',h') was also arbitrary, so $\mathfrak{M},(s,h) \lessdot \Delta$.

For the other direction suppose $\mathfrak{M},(s,h) \lessdot \Delta$. Then let $\delta \in \Delta$ (i.e., $\Box X^n\delta \in \Delta_{if}$). Clearly, $(s,h) \vDash \Delta$ since $\Box[\Delta] \subseteq \Delta_{if}$. Let h' be such that $s \in h'$. Take the n-th h'-successor of s, call it s', then by assumption $(s',h') \vDash \delta$. But that means that $(s,h') \vDash X^n\delta$, and since h' was arbitrarily chosen, $(s,h) \vDash \Box X^n\delta$, and since n was also arbitrary this holds for all $n \in \mathbb{N}$. That means Δ_{if} is satisfied at (s,h). $\square$

Proof of Theorem 2. Here we will prove the T case, the TE case and the GE case, all of the other cases proceed in a similar manner.

(T case) [$\Rightarrow$] Suppose that Δ has a T-quandary. So every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M},(s,h) \lessdot \Delta$ is such that $\forall s' \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$, and there is $\mathfrak{M}' \lessdot \Delta$. That means $\mathfrak{M}',(s'',h'') \vDash \Delta_{if}$, for some (s'',h'') and so by completeness $\Delta_{if} \nvdash_{ix} \perp$. Let $\mathfrak{M},(s,h) \vDash \Delta_{if}$. By

Observation 1 $\mathfrak{M}, (s,h) \lessdot \Delta$. But that means $\mathrm{lub}(s') \subseteq [\![V]\!]$ for any s' in $\mathrm{AFT}(s,h)$. Suppose $s \in h'$. Then $(s,h') \in \mathrm{AFT}(s,h)$, so $\mathrm{lub}(s) \subseteq [\![V]\!]$. But that means $\mathfrak{M}, (s,h') \vDash XV$, and since h' was arbitrary $\mathfrak{M}, (s,h) \vDash \Box XV$. Since $\mathfrak{M}, (s,h)$ was arbitrary, $\Delta_{if} \vDash_{ix} \Box XV$, and by completeness $\Delta_{if} \vdash_{ix} \Box XV$.

(T case) [$\Leftarrow$] Suppose that $\Delta_{if} \vdash_{\mathrm{xp}} \Box XV$ and $\Delta_{if} \nvdash_{ix} \bot$, then by soundness $\Delta_{if} \vDash_{ix} \Box XV$, and by completeness and Observation 1 there is $\mathfrak{M}'$ such that $\mathfrak{M}' \lessdot \Delta$. Suppose $\mathfrak{M}, (s,h) \lessdot \Delta$, and let $s' \in \mathrm{AFT}(s,h)$. Then $(s',h') \vDash_{ix} \Delta_{if}$ by lemma 1 for any h' with $s \in h'$, so $(s',h') \vDash_{ix} \Box XV$. But that happens only when $\mathrm{lub}(s') \subseteq [\![V]\!]$. Since s' was arbitrary it holds for any s', and since $\mathfrak{M}, (s,h)$ was arbitrary, T holds.

(TE case) [$\Rightarrow$] Suppose that Δ has a TE-quandary. So every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M}, (s,h) \lessdot \Delta$ is such that $\exists s' \in \mathrm{AFT}(s,h), \mathrm{lub}(s') \subseteq [\![V]\!]$, and there is an implementation of Δ. The latter assumption means $\mathfrak{M}', (s'',h'') \vDash \Delta_{if}$ for some (s'',h''), and so by completeness, and Observation 1, $\Delta_{if} \nvdash_{ix} \bot$. Let $\mathfrak{M}, (s,h) \vDash \Delta_{if}$. Then, $\mathfrak{M}, (s,h) \lessdot \Delta$, by Observation 1, and that means there is $s' \in \mathrm{AFT}(s,h)$ such that $\mathrm{lub}(s') \subseteq [\![V]\!]$. And there is a history h' with $s' \in h'$. s' must be the nth h'-successor from s for some $n \in \mathbb{N}$, and $(s',h') \vDash \Box XV$. But that means, $(s,h') \vDash X^n \Box XV$. And so $(s,h) \vDash \Diamond X^n \Box XV$. Since $\mathfrak{M}$ and (s,h) were arbitrary, there are no models of $\Delta_{if} \cup \{\, \neg\Diamond X^n \Box XV : n \in \mathbb{N} \,\}$. That means, by completeness of $\vdash_{ix}$, $\Delta_{if} \cup \{\, \neg\Diamond X^n \Box XV : n \in \mathbb{N} \,\} \vdash_{ix} \bot$.

(TE case) [$\Leftarrow$] Suppose that $\Delta_{if} \cup \Delta_V \vdash_{ix} \bot$, and $\Delta_{if} \nvdash_{ix} \bot$, then — by completeness and Observation 1 — there is $\mathfrak{M}'$ and $(s',h') \in |\mathfrak{M}|$ such that $\mathfrak{M}', (s',h') \lessdot \Delta$ from the latter assumption. Suppose $\mathfrak{M}, (s,h) \lessdot \Delta$. Then $(s,h) \vDash \Delta_{if}$, by Observation 1. By compactness of $\vdash_{ix}$ there must be $n_1, \ldots, n_k \in \mathbb{N}$ such that $\Delta_{if} \cup \{\, \neg\Diamond X^{n_i} \Box XV : 1 \leq i \leq k \,\} \vdash_{ix} \bot$. But that means $\Delta_{if} \vdash_{ix} \bigvee_{1 \leq i \leq k} \Diamond X^{n_i} \Box XV$ by classical logic, and by soundness of $\vdash_{ix}$, $\Delta_{if} \vDash_{ix} \bigvee_{1 \leq i \leq k} \Diamond X^{n_i} \Box XV$. What this means is that $\Diamond X^{n_i} \Box XV$ is true at (s,h) for some $1 \leq i \leq k$. Let it be n_i, i.e., $(s,h) \vDash \Diamond X^{n_i} \Box XV$. Then there is h' with $s \in h'$ and $(s,h') \vDash X^{n_i} \Box XV$. Let s' be the n_ith h'-successor from s, then $(s',h') \vDash \Box XV$. But that means $\mathrm{lub}(s') \subseteq [\![V]\!]$. Since $\mathfrak{M}$ and (s,h) were arbitrary, there is a TE-quandary.

(GE case) [$\Rightarrow$] Suppose that Δ has a GE quandary. So every model $\mathfrak{M}$ with (s,h) such that $\mathfrak{M}, (s,h) \lessdot \Delta$ is such that $\forall (s',h') \in \mathrm{AFT}(s,h)$, $\exists \mathbf{A} \subseteq \mathbf{Ag}$, such that $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$, and there is $\mathfrak{M}' \lessdot \Delta$. That means $\mathfrak{M}', (s'',h'') \vDash \Delta_{if}$ for some (s'',h''). Suppose that $\mathfrak{M}, (s,h) \vDash \Delta_{if}$. Then there is $\mathbf{A} \subseteq \mathbf{Ag}$ such that $E(s',h',\mathbf{A}) \subseteq [\![V]\!]$ by assumption for any $(s', h') \in \mathrm{AFT}(s,h)$. That means, since $(s,h) \in \mathrm{AFT}(s,h)$, $(s,h) \vDash [\mathbf{A}\,\mathsf{xstit}]\,V$, and so $(s,h) \vDash \bigvee_{\mathbf{A}\in\mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\,V$. Since $\mathfrak{M}, (s,h)$ was arbitrary, $\Delta_{if} \vDash_{ix} \bigvee_{\mathbf{A}\in\mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\,V$. By completeness, $\Delta_{if} \vdash_{ix} \bigvee_{\mathbf{A}\in\mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\,V$.

(GE case) [$\Leftarrow$] Suppose $\Delta_{if} \vdash_{ix} \bigvee_{\mathbf{A}\in\mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\,V$ and $\Delta_{if} \nvdash_{ix} \bot$. Then, by soundness, $\Delta_{if} \vDash_{ix} \bigvee_{\mathbf{A}\in\mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\,V$, and by completeness and Observation 1, there is $\mathfrak{M}'$ and $(s', h') \in |\mathfrak{M}|$ such that $\mathfrak{M}', (s', h') \lessdot \Delta$. Suppose for some $\mathfrak{M}$, $\mathfrak{M}, (s,h) \lessdot \Delta$, and let $(s', h') \in \mathrm{AFT}(s,h)$. Then $(s', h') \vDash \Delta_{if}$ by lemma 1, and so $(s', h') \vDash \bigvee_{\mathbf{A}\in\mathcal{P}(\mathbf{Ag})} [\mathbf{A}\,\mathsf{xstit}]\,V$ from our assumption. Thus for some $\mathbf{A} \subseteq \mathbf{Ag}$, $E(s', h', \mathbf{A}) \subseteq [\![V]\!]$. So there must be $\mathbf{A}$ such that $E(s', h', \mathbf{A}) \subseteq [\![V]\!]$ for any (s', h') since it was arbitrary. And because $\mathfrak{M}$ and (s,h) were arbitrary, GE holds.

The other cases follow similar patterns. □

Acknowledgments. The author would like to acknowledge the support of the Killam Foundation and the Social Science and Humanities Research Council of Canada through a Killam Postdoctoral Fellowship and a Banting Postdoctoral Fellowship, respectively. The author would also like to thank the anonymous referee for comments on this paper.

References

[1] Abdou, J., and H. Keiding, *Effectivity Functions in Social Choice*, volume 8 of "Theory and Decision Lbrary. Series C, Game Theory, Mathematical Programmlng, and Ooperat1ons Research", Springer, Dordrecht, 1991. DOI: 10.1007/978-94-011-3448-4

[2] Ågotnes, T., W. van der Hoek, J. A. Rodríguez-Aguilar, C. Sierra, and M. Wooldridge, "On the logic of normative systems", in G. Boella, L. W. N. van der Torre, and H. Verhagen (eds.), *Normative Multi-agent Systems*, volume 07122 of *Dagstuhl Seminar Proceedings*, Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2007.

[3] Alchourrón, C. E., "Conflicts of norms and the revision of normative systems", *Law and Philosophy* 10 (1991): 413–25. DOI: 10.2307/3504843

[4] Anderson, A. R., "A reduction of deontic logic to alethic modal logic", *Mind*, 67, 265 (1958): 100–103. DOI: 10.1093/mind/LXVII.265.100

[5] Belnap, N., and M. Perloff, "The way of the agent", *Studia Logica* 51, 3–4 (1992): 463–484. DOI: 10.1007/BF01028971

[6] Braybrooke, D., B. Brown, and P. Schotch, *Logic on the Track of Social Change*, Clarendon Library of Logic and Philosophy, Oxford: Clarendon Press, 1995. DOI: 10.1093/acprof:oso/9780198235309.001.0001

[7] Broersen, J., "On the reconciliation of logics of agency and logics of event types", chapter 3, pages 41–60, in R. Trypuz (ed.), *Krister Segerberg on the Logic of Actions*, volume 1 of "Outstanding Contributions to Logic", Springer, Dordrecht, 2014. DOI: 10.1007/978-94-007-7046-1_3

[8] Broersen, J., and J.-J. C. Meyer, "Action, failure and free will choice in stit logic", chapter 8, pages 141–168, in F. Lihoreau and M. Rebuschi (eds.), *Epistemology, Context and Formalism*, Synthese Library, Springer, 2013. DOI: 10.1007/978-3-319-02943-6_9

[9] Chellas, B. F., *Modal Logic: An introduction*, Cambridge University Press, Cambridge, UK, 1980. DOI: 10.1017/CBO9780511621192

[10] Czelakowski, J., *Freedom and Enforcement in Action: Elements of Formal Action Theory*, volume 42 of "Trends in Logic", Springer, Dordrecht, 2015. DOI: 10.1007/978-94-017-9855-6

[11] Grossi, D., "Desgining invisible handcuffs", PhD thesis, University of Utrecht, September 2007, SIKS Dissertation Series no. 2007-16.

[12] Grossi, D., "Norms as ascriptions of violations: An analysis in modal logic", *Journal of Applied Logic* 9, 2 (2011): 95–112. DOI: 10.1016/j.jal.2010.03.002

[13] Hamblin, C. L., "Quandries and the logic of rules", *Journal of Philosophical Logic* 1, 1 (1972): 74–85.

[14] Hohfeld, W. N., *Fundamentals of Legal Concepts*, Yale University Press, New Haven, Conn, 1920. Reprinted 1964.

[15] Jones, A. J. I., and M. Sergot, "A formal characterization of institutionalized power", *Logic Jnl IGPL* 4, 3 (1996): 427–443.

[16] Makinson, D., "On the formal representation of rights relations", *Journal of Philosophical Logic*, 15, 4 (1986): 403–425. DOI: 10.1007/BF00243391

[17] Marcus, R. B., "Moral dilemmas and consistency", *The Journal of Philosophy* 77, 3 (1980): 121–136. DOI: 10.2307/2025665

[18] Pauly, M., "Logics for social software", PhD thesis, University of Amsterdam, Amsterdam, 2001, Dissertation Series DS-2001-10.

[19] Searle, J., *The Construction of Social Reality*, The Free Press, 1995.

[20] Searle, J., *Making the Social World*, Oxford University Press, Oxford, UK, 2010. DOI: 10.1093/acprof:osobl/9780195396171.001.0001

[21] Stolpe, A., *Norms and Norm-System Dynamics*, Department of Philosophy, University of Bergen, Norway, 2008.

[22] van Hees, M., "Some general results on responsibility for outcomes", chapter 7, pages 99–109, in A. van Deeman and A. Rusinowska (eds.), *Collective Decision Making*, "Theory and Decision Library. Series C", Springer, Heidelberg, 2010. DOI: 10.1007/978-3-642-02865-6_7

[23] von Wright, G. H., "Is there a logic of norms?", *Ratio Juris* 4, 3 (1991): 265–283. DOI: 10.1111/j.1467-9337.1991.tb00099.x

[24] Weinberger, O., "The expressive conception of norms: An impasse for the logic of norms", *Law and Philosophy* 4, 2 (1985): 165–198. DOI: 10.2307/3504669

Gillman Payette
Department of Philosophy
University of British Columbia
Vancouver, Canada
gpayette@dal.ca