CONDITIONAL EXPRESSIVITY AND COLLECTIVE DEONTIC ADMISSIBILITY

FREDERIK VAN DE PUTTE
Erasmus School of Philosophy, Erasmus University Rotterdam
and
Centre for Logic and Philosophy of Science, Ghent University

HEIN DUIJF
Department of Philosophy and Religious Studies, Utrecht University
and

ALLARD TAMMINGA
Department of Philosophy, University of Greifswald

Abstract. This paper makes a twofold contribution to the study of expressivity. First, we introduce and study the novel concept of conditional expressivity. Taking a universal logic perspective, we characterize conditional expressivity both syntactically and semantically. We show that our concept of conditional expressivity is related to, but different from, the concept of explicit definability in Beth's definability theorem. Second, we use the concept to explore inferential relations between collective deontic admissibility statements for different groups. Negative results on conditional expressivity are stronger than standard (unconditional) inexpressivity results: we show that the well-known inexpressivity results from epistemic logic on distributed knowledge and on common knowledge only concern unconditional expressivity. By contrast, we prove negative results on conditional expressivity in the deontic logic of collective agency. In particular, we consider the full formal language of the deontic logic of collective agency, define a natural class of sublanguages of the full language, and prove that a collective deontic admissibility statement about a particular group is conditionally expressible in a sublanguage from the class if and only if that sublanguage includes a collective deontic admissibility statement about a supergroup of that group. Our negative results on conditional expressivity may serve as a proof of concept for future studies.

§1. Introduction. We introduce and study a novel variant of the standard concept of expressivity. Standardly, a statement ϕ from a given language is *expressible* in a sublanguage of that language if and only if there is a statement ψ in the sublanguage such that ϕ and ψ are logically equivalent.[1] We say that a statement ϕ from a given language is *conditionally expressible* in a sublanguage of that language if and only

Received: June 10, 2024.
2020 *Mathematics Subject Classification*: Primary 03B45.
Key words and phrases: expressivity, bisimulation, deontic logic, collective agency, collective admissibility.

[1] See, for instance, [35, definition 8.2]. In the literature, the terms 'expressible' and 'definable' are often used interchangeably. To distinguish the concept of expressibility from the related

doi:10.1017/S1755020325100798

if there are subsets Γ and Δ of the sublanguage such that Γ non-trivially implies that ϕ and Δ are logically equivalent (Definition 1 below).[2] In the first part of this paper, we take the perspective of universal logic and characterize our novel concept of conditional expressivity both syntactically and semantically. In particular, we show that if ϕ non-trivially implies or is non-trivially implied by any of the statements of the sublanguage, then ϕ is conditionally expressible in that sublanguage.[3] Consequently, if ϕ is not conditionally expressible in the sublanguage, then there is no statement in that sublanguage that non-trivially implies or is non-trivially implied by ϕ.

Although the universal logic perspective shows it to be widely applicable, our concept of conditional expressivity was originally motivated by the philosophical debate on collective agency, collective obligations, and collective responsibility.[4] That debate has been focused almost exclusively on the question whether and how statements about groups are inferentially related to statements about individuals. It has thereby prevented another question from being asked: whether and how statements about a given group are inferentially related to statements about *other groups*. In the second part of this paper, we rephrase the latter question using our concept of conditional expressivity, focusing on *collective deontic admissibility statements* of the form "Group $\mathcal{G}$ of agents performs a deontically admissible group action" (formalized as $\star_{\mathcal{G}}$). Such statements are a key component of a well-established *deontic logic of collective agency* that models actions, omissions, abilities, and obligations of finitely many individuals and groups of individuals.[5]

In the present study, we assess the inferential relations between, on the one hand, the collective deontic admissibility statement $\star_{\mathcal{G}}$ about a group $\mathcal{G}$ of agents and, on the other hand, collective deontic admissibility statements about $\mathcal{G}$'s subgroups, supergroups, outgroups, and partially overlapping groups. In particular, we define a natural class of sublanguages of the full language of the deontic logic of collective agency and assess, for each sublanguage in that class, whether $\star_{\mathcal{G}}$ is conditionally expressible in that sublanguage, that is, whether $\star_{\mathcal{G}}$ non-trivially implies or is non-trivially implied by any of the statements in that sublanguage.

---

but different concept of definability in Beth's definability theorem (see the discussion in §2), throughout this paper, we stick to 'expressible' and its variants.

2 In fact, there are at least four variants of expressivity: we could weaken the standard definition of (unconditional) expressivity by requiring that there be a possibly infinite set Δ of the sublanguage such that ϕ and Δ are logically equivalent. Likewise, we could strengthen our definition of conditional expressivity by requiring that there be statements χ and ψ in the sublanguage such that χ non-trivially implies that ϕ and ψ are logically equivalent. In this paper, we use the standard finitary variant of (unconditional) expressivity to discuss well-known examples from epistemic logic and the infinitary (and thus weaker) variant of conditional expressivity to state our central theorem, as the latter allows us to present our negative result in its strongest form.

3 A statement ϕ *trivially implies* a statement ψ if ψ is a logical truth. A statement ϕ *is trivially implied by* a statement ψ if ψ is a logical falsity.

4 Key contributors to the debate include Kutz [19], Isaacs [16], List and Pettit [20], Tollefsen [31], Collins [7], and Schwenkenbecher [25].

5 The deontic logic of collective agency [29] is a deontic logic in the tradition of *stit* ('sees to it that') logics of agency. See [2] and [14] for textbook presentations of *stit* logics and historical references. Collective actions and/or obligations have been studied using *stit*-like frameworks in [11], [32], [2, chap. 10], [14, chap. 6], [18], [5], [15], [9], [26], and [8].

Our central theorem on collective deontic admissibility establishes that the collective deontic admissibility statement $\star_{\mathcal{G}}$ is conditionally expressible in a given sublanguage from the natural class of sublanguages of the full language of the deontic logic of collective agency *if and only if* that sublanguage contains a collective deontic admissibility statement $\star_{\mathcal{H}}$ for some supergroup $\mathcal{H}$ of $\mathcal{G}$ (Theorem 10). The right-to-left direction of our central theorem is new. Its left-to-right direction is a considerable strengthening, in various ways, of an earlier result from [9]. Phrased differently, it says that if the sublanguage does not contain a collective deontic admissibility statement $\star_{\mathcal{H}}$ for some supergroup $\mathcal{H}$ of $\mathcal{G}$, then there are no non-trivial inferential relations between $\star_{\mathcal{G}}$ and any statement of that sublanguage.[6]

Our paper proceeds as follows. In §2, we first take the perspective of universal logic to define conditional expressivity and establish necessary and sufficient conditions for it. This is followed by a comparison of our concept of conditional expressivity with the concept of *explicit definability* in Beth's definability theorem. To illustrate the concept of conditional expressivity, we discuss two well-known inexpressivity results from epistemic logic and show that these impossibility results on (unconditional) expressivity do not generalize to conditional expressivity. We then study the concept of conditional expressivity semantically and give a semantic criterion for proving that a statement is not conditionally expressible in a given formal language. In §3, we recall the full formal language and the semantics of the deontic logic of collective agency. We define a natural class of sublanguages that differ with respect to (a) the included deontic admissibility statements and (b) the included *stit* operators for individual and group agency. We then give conditions under which two models are bisimilar with respect to one of the sublanguages in the class and prove a general Hennessy–Milner theorem that covers every sublanguage in the class. In §4, building on the previous steps, we generalize the impossibility result on conditional expressivity from [9] to much more expressive sublanguages. In §5, we prove that the collective deontic admissibility statement $\star_{\mathcal{G}}$ is conditionally expressible in a sublanguage from the natural class of sublanguages if and only if that sublanguage contains a collective deontic admissibility statement $\star_{\mathcal{H}}$ for some supergroup $\mathcal{H}$ of $\mathcal{G}$. Lastly, in §6, we compare our new impossibility result with the earlier result from [9] and discuss two straightforward applications.

§2. Conditional and unconditional expressivity. When can a particular statement be expressed in a particular formal language? To answer this question, we must do at least three things. First, we must specify the formal language $\mathfrak{L}_e$ that is supposed to do the expressing. Second, we must specify a semantics that gives truth-conditions for the statement to be expressed and for the statements of the language $\mathfrak{L}_e$.[7] Third—and this

---

[6] Favoring a particular type of *ontological* reductionism about groups, Collins [7, p. 75] argues that to believe otherwise "would be to suggest that groups are somehow free-floating, independent, or self-sustaining entities that are not in a closed causal system with the rest of the world." Nonetheless, when it comes to *explanatory* reductionism (where we should think of inference relations rather than of causal ones), these words aptly describe the predicament of collective deontic admissibility.

[7] Compare [30, p. 152]: "The question how a certain concept is to be defined is correctly formulated only if a list is given of the terms by means of which the required definition is to be constructed. If the definition is to fulfil its proper task, the sense of the terms in this list must admit of no doubt." and [22, p. 196]: "indefinability [...] is relative to a set of ideas;

is a point that seems to have gone unnoticed thus far—we must be fully explicit about what counts as an expression of a statement in the language $\mathfrak{L}_e$, because expressivity can be either conditional or unconditional.

We take the perspective of universal logic [3] to study expressivity. First, we give formal definitions of conditional and unconditional expressivity. Second, we state two conditions on the relation between the statement to be expressed and the language that is supposed to do the expressing. We show that a statement is not conditionally expressible in a given formal language if and only if both conditions are met. Third, we use the two conditions to illustrate the difference between standard (unconditional) expressivity and conditional expressivity with three varieties of group knowledge from epistemic logic. Lastly, we give semantic counterparts to the two conditions. These counterparts amount to a semantic criterion for proving that a statement is not conditionally expressible in a given formal language. This criterion will be at the basis of our negative and positive conditional expressivity results on collective deontic admissibility statements—see §4 and §5, respectively.

2.1. Expressivity and universal logic. Let $\mathfrak{L}$ and $\mathfrak{L}_e$ be two fixed non-empty formal languages such that $\mathfrak{L}_e \subset \mathfrak{L}$. (We think of the language $\mathfrak{L}_e$ as the language that is supposed to do the expressing.) We use ϕ and ψ as variables for statements in $\mathfrak{L}$, and Γ and Δ as variables for subsets of $\mathfrak{L}$. Let $\Vdash$ be a fixed consequence relation from subsets of $\mathfrak{L}$ to statements in $\mathfrak{L}$. (Note that $\Vdash$ may be specified proof-theoretically or model-theoretically.) We write $\Gamma \Vdash \Delta$ if for all ψ in Δ it holds that $\Gamma \Vdash \psi$.[8] We say that Γ is $\mathfrak{L}$-*trivial* if $\Gamma \Vdash \mathfrak{L}$. We assume the consequence relation $\Vdash$ to have two properties. Our first assumption is that if Δ and Σ are logically equivalent on the condition that Γ, then $\Gamma \cup \Delta$ is $\mathfrak{L}_e$-trivial if and only if $\Gamma \cup \Sigma$ is $\mathfrak{L}_e$-trivial. That is, for all $\Gamma, \Delta, \Sigma \subseteq \mathfrak{L}$, we have:[9]

UL_1: if both $\Gamma, \Delta \Vdash \Sigma$ and $\Gamma, \Sigma \Vdash \Delta$, then $\Gamma, \Delta \Vdash \mathfrak{L}_e$ if and only if $\Gamma, \Sigma \Vdash \mathfrak{L}_e$.

Our second assumption is that $\mathfrak{L}_e$ is $\mathfrak{L}$-trivial:

UL_2: $\mathfrak{L}_e \Vdash \mathfrak{L}$.

The properties UL_1 and UL_2 are very weak. UL_1 holds if the consequence relation $\Vdash$ is Tarskian, that is, if it is reflexive, transitive, and monotonic (note that the converse does not hold).

Given the consequence relation $\Vdash$, we can state the standard definition of (unconditional) expressivity as follows: a statement ϕ is *expressible* in a language $\mathfrak{L}_e$ if and only if there is a statement ψ in $\mathfrak{L}_e$ such that ψ and ϕ are logically equivalent, that is, if and only if there is a statement $\psi \in \mathfrak{L}_e$ such that $\psi \Vdash \phi$ and $\phi \Vdash \psi$. Accordingly, a statement ϕ is not expressible in a language $\mathfrak{L}_e$ if and only if for every $\psi \in \mathfrak{L}_e$ it holds that $\psi \nVdash \phi$ or $\phi \nVdash \psi$.

If a statement ϕ is not (unconditionally) expressible in a language $\mathfrak{L}_e$, then it is impossible to state necessary and sufficient conditions for ϕ in that language. This

and an idea which is indefinable relative to one set of primitives may be definable relative to another.”

8 Note that $\Gamma \Vdash \Delta$ if $\Delta = \emptyset$. An empty set on the right-hand side does not imply triviality.

9 We adopt standard notational conventions and write ‘Γ, Δ’ instead of ‘$\Gamma \cup \Delta$’, and ‘Γ, ϕ’ instead of ‘$\Gamma \cup \{\phi\}$’. Note that we do not use commas to the right of the consequence relation $\Vdash$.

does not, however, exclude that there are certain conditions that can be stated in $\mathfrak{L}_e$ under which ϕ is logically equivalent to a statement or set of statements in $\mathfrak{L}_e$. If that is indeed possible, ϕ is conditionally expressible in $\mathfrak{L}_e$. Formally, we say that a statement ϕ is *conditionally expressible* in a language $\mathfrak{L}_e$ if and only if there is a pair of subsets Γ and Δ of $\mathfrak{L}_e$ such that Γ is not $\mathfrak{L}_e$-trivial and ϕ and Δ are logically equivalent on the condition that Γ.[10]

DEFINITION 1 (Conditional Expressivity). *Let* $\phi \in \mathfrak{L} - \mathfrak{L}_e$. *Then* ϕ is conditionally expressible in $\mathfrak{L}_e$ *if and only if there are* $\Gamma, \Delta \subseteq \mathfrak{L}_e$ *such that*

(i) $\Gamma \not\Vdash \mathfrak{L}_e$;
(ii) $\Gamma, \Delta \Vdash \phi$;
(iii) $\Gamma, \phi \Vdash \Delta$.

Accordingly, a statement ϕ is not conditionally expressible in a language $\mathfrak{L}_e$ if and only if for every $\Gamma, \Delta \subseteq \mathfrak{L}_e$ it holds that if $\Gamma \not\Vdash \mathfrak{L}_e$, then $\Gamma, \Delta \not\Vdash \phi$ or $\Gamma, \phi \not\Vdash \Delta$. Or, equivalently, if and only if for every $\Gamma, \Delta \subseteq \mathfrak{L}_e$ it holds that if $\Gamma, \Delta \Vdash \phi$ and $\Gamma, \phi \Vdash \Delta$, then $\Gamma \Vdash \mathfrak{L}_e$.

Our concept of conditional expressivity is related to, but different from, the concept of *explicit definability* that is characterized in Beth's definability theorem.[11] Let us state Beth's theorem and compare both concepts. If $\mathfrak{L}_e$ and $\mathfrak{L}$ are first-order languages such that $\mathfrak{L}_e \subseteq \mathfrak{L}$ and M is an $\mathfrak{L}$-model, then $M|\mathfrak{L}_e$ is the $\mathfrak{L}_e$-submodel of M that only interprets the symbols that are in both $\mathfrak{L}_e$ and $\mathfrak{L}$. Using $\overline{a}$ for n-tuples of constants and $\overline{x}$ for n-tuples of variables, Beth's theorem can now be stated as follows.

THEOREM (Beth's Definability Theorem). *Let* $\mathfrak{L}_e$ *and* $\mathfrak{L}$ *be first-order languages such that* $\mathfrak{L}_e \subseteq \mathfrak{L}$. *Let* $\Gamma \subseteq \mathfrak{L}$ *and* $\phi(\overline{x}) \in \mathfrak{L}$. *Then the following are equivalent*:

(i) *If* $M_1 \models \Gamma$ *and* $M_2 \models \Gamma$ *and* $M_1|\mathfrak{L}_e = M_2|\mathfrak{L}_e$, *then for all* $\overline{a}$ *in* M_1 *it holds that* $M_1 \models \phi(\overline{a})$ *iff* $M_2 \models \phi(\overline{a})$;
(ii) *There is a* $\psi(\overline{x}) \in \mathfrak{L}_e$ *such that* $\Gamma \models \psi(\overline{x}) \leftrightarrow \phi(\overline{x})$.

To connect conditional expressivity and explicit definability, we assume a classical setting in which $\mathfrak{L}_e$ and $\mathfrak{L}$ are first-order languages such that $\mathfrak{L}_e \subseteq \mathfrak{L}$. Let $\Gamma \subseteq \mathfrak{L}$ and $\phi \in \mathfrak{L}$. We say that ϕ is *explicitly definable modulo* Γ *in terms of* $\mathfrak{L}_e$ if condition (ii) of Beth's definability theorem holds. Roughly,[12] we have that ϕ is conditionally expressible in $\mathfrak{L}_e$ if and only if there is a $\Gamma \subseteq \mathfrak{L}_e$ such that ϕ is explicitly definable modulo Γ in terms of $\mathfrak{L}_e$. Accordingly, if ϕ is not conditionally expressible in $\mathfrak{L}_e$, then for every $\Gamma \subseteq \mathfrak{L}_e$ that is not $\mathfrak{L}_e$-trivial it holds that ϕ is not explicity definable modulo Γ in terms of $\mathfrak{L}_e$. Note that conditional expressivity requires that Γ be a subset of $\mathfrak{L}_e$, not just of $\mathfrak{L}$.[13]

---

10 The concept of conditional expressivity can be seen as a generalization of Carnap's concept of a *bilateral reduction sentence* [6, §5–§10]. A bilateral reduction sentence for a term 'Cx' is of the form '$Sx \to (Cx \equiv Rx)$'. It "is not a full definition (which would have to be of the form '$Cx \equiv \ldots$', with 'Cx' constituting the definiendum); it specifies the meaning of 'Cx', not for all cases, but only for those that satisfy the condition S. In this sense, it constitutes only a partial, or conditional, definition for C" [12, pp. 129–130].

11 See, for instance, [13, Theorem 5.5.4]. We thank an anonymous reviewer for pointing out this relation.

12 To be precise, the equivalence holds if we require on the left-hand side that Δ be a singleton and on the right-hand side that Γ not be $\mathfrak{L}_e$-trivial.

13 Given the current state of the research, two differences between conditional expressivity and explicit definability come to mind. First, our definition of and our results on conditional

Our first theorem gives an analysis of conditional expressivity. We state two conditions, (C_1) and (C_2), on the relation between the statement ϕ that is to be expressed and the language $\mathfrak{L}_e$ that is supposed to do the expressing. Informally, (C_1) says that every subset of $\mathfrak{L}_e$ that proves ϕ is trivial, and (C_2) says that every subset of $\mathfrak{L}_e$ that disproves ϕ is trivial. It holds that a statement ϕ is not conditionally expressible in a language $\mathfrak{L}_e$ if and only if both conditions are met.

THEOREM 1. *Let $\phi \in \mathfrak{L} - \mathfrak{L}_e$ and let $\Vdash$ be a consequence relation that has the properties UL_1 and UL_2. Then ϕ is not conditionally expressible in $\mathfrak{L}_e$ if and only if both*

(C_1) *for every $\Gamma \subseteq \mathfrak{L}_e$, it holds that if $\Gamma \Vdash \phi$, then $\Gamma \Vdash \mathfrak{L}_e$, and*
(C_2) *for every $\Gamma \subseteq \mathfrak{L}_e$, it holds that if $\Gamma, \phi \Vdash \mathfrak{L}_e$, then $\Gamma \Vdash \mathfrak{L}_e$.*

Proof. $(\Rightarrow)$ Assume that (C_1) or (C_2) does not hold. If (C_1) does not hold, there must be a $\Gamma \subseteq \mathfrak{L}_e$ such that $\Gamma \Vdash \phi$ and $\Gamma \nVdash \mathfrak{L}_e$. Hence, (i) $\Gamma \nVdash \mathfrak{L}_e$, (ii) $\Gamma, \emptyset \Vdash \phi$, and (iii) $\Gamma, \phi \Vdash \emptyset$. Hence, ϕ is conditionally expressible in $\mathfrak{L}_e$. If (C_2) does not hold, there must be a $\Gamma \subseteq \mathfrak{L}_e$ such that $\Gamma, \phi \Vdash \mathfrak{L}_e$ and $\Gamma \nVdash \mathfrak{L}_e$. Because of UL_2, we have that $\Gamma, \mathfrak{L}_e \Vdash \mathfrak{L}$ and hence $\Gamma, \mathfrak{L}_e \Vdash \phi$. Hence, (i) $\Gamma \nVdash \mathfrak{L}_e$, (ii) $\Gamma, \mathfrak{L}_e \Vdash \phi$, and (iii) $\Gamma, \phi \Vdash \mathfrak{L}_e$. Hence, ϕ is conditionally expressible in $\mathfrak{L}_e$.

$(\Leftarrow)$ Assume that ϕ is conditionally expressible in $\mathfrak{L}_e$. Then there are $\Gamma, \Delta \subseteq \mathfrak{L}_e$ such that (i) $\Gamma \nVdash \mathfrak{L}_e$, (ii) $\Gamma, \Delta \Vdash \phi$, and (iii) $\Gamma, \phi \Vdash \Delta$. We consider two cases. Case (a): suppose that $\Gamma, \phi \Vdash \mathfrak{L}_e$. Then $\Gamma, \phi \Vdash \mathfrak{L}_e$ and $\Gamma \nVdash \mathfrak{L}_e$. Hence, (C_2) does not hold. Case (b): suppose that $\Gamma, \phi \nVdash \mathfrak{L}_e$. Because $\Gamma, \Delta \Vdash \phi$ and $\Gamma, \phi \Vdash \Delta$ and UL_1, we have that $\Gamma, \Delta \nVdash \mathfrak{L}_e$. Then, $\Gamma, \Delta \Vdash \phi$ and $\Gamma, \Delta \nVdash \mathfrak{L}_e$. Hence, (C_1) does not hold. Either way, (C_1) or (C_2) does not hold. $\square$

Generally, *positive* results on conditional expressivity are easy to obtain. From Theorem 1, it follows that if we wish to show that ϕ is conditionally expressible in $\mathfrak{L}_e$, it suffices to find a non-trivial $\Gamma \subseteq \mathfrak{L}_e$ such that $\Gamma \Vdash \phi$ or a non-trivial $\Gamma \subseteq \mathfrak{L}_e$ such that $\Gamma, \phi \Vdash \mathfrak{L}_e$. In the first case, by setting $\Delta = \emptyset$, we have $\Gamma \nVdash \mathfrak{L}_e$ and $\Gamma, \Delta \Vdash \phi$ and $\Gamma, \phi \Vdash \Delta$. In the second case, by setting $\Delta = \mathfrak{L}_e$, we have $\Gamma \nVdash \mathfrak{L}_e$ and $\Gamma, \Delta \Vdash \phi$ and $\Gamma, \phi \Vdash \Delta$.

2.2. Expressivity in epistemic logic. We illustrate the concept of conditional expressivity by way of three different varieties of group knowledge. These varieties have been studied in epistemic logic, in which epistemic statements about individuals like "Agent i knows that ϕ" are formalized as $K_i\phi$. The individualistic epistemic language $\mathfrak{L}_\varepsilon$ is given by the following Backus–Naur form:

$$\phi := p \mid \neg\phi \mid (\phi \wedge \phi) \mid K_i\phi,$$

where p ranges over a fixed countable set $\mathfrak{P}$ of atoms and i ranges over a fixed finite set $\mathcal{N}$ of individual agents.

Truth-conditions for the statements in the individualistic epistemic language $\mathfrak{L}_\varepsilon$ are specified in terms of epistemic models $M = \langle W, (R_i)_{i \in \mathcal{N}}, V \rangle$, where W is a non-empty set of possible worlds, every R_i is an equivalence relation on W that captures the pairs of worlds that individual agent i cannot distinguish epistemically, and V is a valuation function that assigns to each atom a set of possible worlds where that atom is true. The

expressivity are restricted neither to first-order languages nor to classical logic. Second, our present results on conditional expressivity only concern 0-ary predicates, whereas Beth's definability theorem concerns n-ary predicates.

truth-condition for individual knowledge that ϕ is as follows: $K_i\phi$ is true at a world w in an epistemic model M if and only if for all w' in W such that $R_i ww'$ it holds that ϕ is true at w'. In this subsection, let $\Vdash$ be the consequence relation associated with these epistemic models.

The individualistic epistemic language $\mathfrak{L}_\varepsilon$ has been extended by adding different modalities that characterize different varieties of group knowledge [10, §2.2]. We discuss three of these modalities to illustrate our concept of conditional expressivity. First, $\mathfrak{L}_\varepsilon$ can be extended by adding a new modality $E_\mathcal{G}$ to formalize statements of *general knowledge* among the members of $\mathcal{G}$. Let $R^E_\mathcal{G}$ be the union $\bigcup_{i\in\mathcal{G}} R_i$ of all of the group members' equivalence relations. The truth-condition for general knowledge among the members of $\mathcal{G}$ that ϕ is as follows: $E_\mathcal{G}\phi$ is true at a world w in an epistemic model M if and only if for all w' in W such that $R^E_\mathcal{G} ww'$ it holds that ϕ is true at w'. It is easy to see that the general knowledge statement $E_\mathcal{G} p$ is logically equivalent to the statement $\bigwedge_{i\in\mathcal{G}} K_i p$ from $\mathfrak{L}_\varepsilon$. We have:

$$E_\mathcal{G} p \Vdash \bigwedge_{i\in\mathcal{G}} K_i p \text{ and } \bigwedge_{i\in\mathcal{G}} K_i p \Vdash E_\mathcal{G} p.$$

Hence, $E_\mathcal{G} p$ is (unconditionally) expressible in $\mathfrak{L}_\varepsilon$.

Second, $\mathfrak{L}_\varepsilon$ can be extended by adding a new modality $D_\mathcal{G}$ to formalize statements of *distributed knowledge* among the members of $\mathcal{G}$. Let $R^D_\mathcal{G}$ be the intersection $\bigcap_{i\in\mathcal{G}} R_i$ of all of the group members' equivalence relations. The truth-condition for distributed knowledge among the members of $\mathcal{G}$ that ϕ is as follows: $D_\mathcal{G}\phi$ is true at w in M if and only if for all w' in W such that $R^D_\mathcal{G} ww'$ it holds that ϕ is true in w'. It is known that the statement $D_\mathcal{G} p$ is not (unconditionally) expressible in $\mathfrak{L}_\varepsilon$ [34]. Nonetheless, if $i \in \mathcal{G}$, then the distributed knowledge statement $D_\mathcal{G} p$ is implied by the statement $K_i p$ from $\mathfrak{L}_\varepsilon$. We have:

$$K_i p \Vdash D_\mathcal{G} p, \text{ if } i \in \mathcal{G}.$$

Because $K_i p \in \mathfrak{L}_\varepsilon$ and $K_i p \Vdash D_\mathcal{G} p$ and $K_i p \nVdash \mathfrak{L}_\varepsilon$, condition (C_1) does not hold for distributed knowledge. By Theorem 1, $D_\mathcal{G} p$ is conditionally expressible in $\mathfrak{L}_\varepsilon$. We have:[14]

$$K_i p, p \vee \neg p \Vdash D_\mathcal{G} p \text{ and } K_i p, D_\mathcal{G} p \Vdash p \vee \neg p, \text{ if } i \in \mathcal{G}.$$

Third, $\mathfrak{L}_\varepsilon$ can be extended by adding a new modality $C_\mathcal{G}$ to formalize statements of *common knowledge* among the members of $\mathcal{G}$. Let $R^C_\mathcal{G}$ be the transitive closure of the union $R^E_\mathcal{G}$ of all of the group members' equivalence relations.[15] The truth-condition for common knowledge among the members of $\mathcal{G}$ that ϕ is as follows: $C_\mathcal{G}\phi$ is true at

[14] Not all cases of conditional expressivity are 'degenerate' in the sense that they involve tautologies or contradictions. We have $p \to K_i p, K_i p \Vdash D_\mathcal{G} p$ and $p \to K_i p, D_\mathcal{G} p \Vdash K_i p$ if $i \in \mathcal{G}$, although $p \to K_i p \nVdash D_\mathcal{G} p$ and $D_\mathcal{G} p \nVdash p \to K_i p$. It is an open problem how to make a sharp and principled distinction between 'degenerate' and 'non-degenerate' cases of conditional expressivity. Such a distinction would presumably require additional conditions on conditional expressivity. Whatever such additional conditions might be, negative results on conditional expressivity—like the ones in §4 of this paper—imply negative results on 'non-degenerate' conditional expressivity.

[15] The *transitive closure* of a relation R is the smallest relation R^+ (in terms of set-theoretical inclusion) such that (i) $R \subseteq R^+$ and (ii) $\forall w \forall w' \forall w''((R^+ww' \wedge R^+w'w'') \to R^+ww'')$. Note that R^+ is reflexive if R is reflexive.

w in M if and only if for all w' in W such that $R^C_{\mathcal{G}} ww'$ it holds that ϕ is true in w'. Again, it is known that the statement $C_{\mathcal{G}}\, p$ is not (unconditionally) expressible in $\mathfrak{L}_\varepsilon$ [35, theorem 8.34].[16] Nonetheless, if i is in $\mathcal{G}$, the common knowledge statement $C_{\mathcal{G}}\, p$ implies the statement $K_i\, p$ from $\mathfrak{L}_\varepsilon$. We have:

$$C_{\mathcal{G}}\, p \Vdash K_i\, p, \text{ if } i \in \mathcal{G}.$$

Because $\neg K_i\, p \in \mathfrak{L}_\varepsilon$ and $\neg K_i\, p, C_{\mathcal{G}}\, p \Vdash \mathfrak{L}_\varepsilon$ and $\neg K_i\, p \nVdash \mathfrak{L}_\varepsilon$, condition (C_2) does not hold for common knowledge. By Theorem 1, $C_{\mathcal{G}}\, p$ is conditionally expressible in $\mathfrak{L}_\varepsilon$. We have:

$$\neg K_i\, p, p \wedge \neg p \Vdash C_{\mathcal{G}}\, p \text{ and } \neg K_i\, p, C_{\mathcal{G}}\, p \Vdash p \wedge \neg p, \text{ if } i \in \mathcal{G}.$$

In summary, the two inexpressivity results from epistemic logic on distributed knowledge and on common knowledge only concern *unconditional* expressivity. As we have seen, both varieties of group knowledge are conditionally expressible in the individualistic epistemic language $\mathfrak{L}_\varepsilon$. This observation might give the impression that any statement about groups is conditionally expressible in terms of statements about individuals. This is not the case. In §4, we study the conditional expressivity of collective deontic admissibility statements and prove that they are not even conditionally expressible in a wide range of sublanguages of the deontic logic of collective action. To do so, we first develop a semantic method for proving that a given statement is not conditionally expressible in a given language.

2.3. How to disprove conditional expressivity? To prove that a statement ϕ from a language $\mathfrak{L}$ is not conditionally expressible in a non-empty sublanguage $\mathfrak{L}_e$ of $\mathfrak{L}$, we must show that conditions (C_1) and (C_2) of Theorem 1 hold. In this subsection, we state a semantic counterpart to each of the two conditions and prove that the counterparts capture the conditions neatly. The two counterparts amount to a semantic criterion for proving that a statement is not conditionally expressible in a given formal language. In §4, we use that criterion to prove that collective deontic admissibility statements are not conditionally expressible in a range of sublanguages of the deontic logic of collective agency. Our semantic counterparts to the conditions rely on two assumptions.

First, we assume that the consequence relation $\Vdash$ is fixed by a model-theoretical semantics, that is, we assume that

Sem$_1$: There is a class $\mathcal{C}$ of *indices of evaluation* in terms of which the truth-conditions for the statements in $\mathfrak{L}$ are given. The consequence relation $\Vdash$ is defined from $\mathcal{C}$ in the standard, Tarskian way.

We write '$x \models \psi$' if the statement ψ from $\mathfrak{L}$ is true at the evaluation index x from $\mathcal{C}$. We write '$x \models \Gamma$' if for every ψ in the set of statements Γ from $\mathfrak{L}$ it holds that $x \models \psi$. The consequence relation $\Vdash$ from subsets of $\mathfrak{L}$ to statements of $\mathfrak{L}$ is standardly defined as follows: $\Gamma \Vdash \psi$ if and only if for every x in $\mathcal{C}$ it holds that if $x \models \Gamma$, then $x \models \psi$.

Second, we assume that every statement in the sublanguage $\mathfrak{L}_e$ of $\mathfrak{L}$ has a companion statement in $\mathfrak{L}_e$ that acts as its classical negation. More formally,

[16] Note that if Δ is the infinite set $\{p, \bigwedge_{i\in\mathcal{G}} K_i\, p, \bigwedge_{i\in\mathcal{G}} K_i(\bigwedge_{i\in\mathcal{G}} K_i\, p), \ldots\} \subseteq \mathfrak{L}_\varepsilon$, then it holds that $C_{\mathcal{G}}\, p \Vdash \Delta$ and $\Delta \Vdash C_{\mathcal{G}}\, p$. Accordingly, $C_{\mathcal{G}}\, p$ *is* (unconditionally) expressible in $\mathfrak{L}_\varepsilon$ if we opt for the *infinitary* variant—see footnote 1.

Sem$_2$: For every $\psi \in \mathfrak{L}_e$, there is a $\chi \in \mathfrak{L}_e$ such that for every x in $\mathcal{C}$, it holds that $x \models \psi$ if and only if $x \not\models \chi$.

The assumption *Sem*$_1$ implies that the consequence relation $\Vdash$ is Tarskian: it is reflexive, transitive, and monotonic. Hence, *Sem*$_1$ also implies UL_1. The assumptions *Sem*$_1$ and *Sem*$_2$ jointly imply UL_2.

We say that x and y in $\mathcal{C}$ are *equivalent on* $\mathfrak{L}_e$ (notation: $x \equiv_{\mathfrak{L}_e} y$) if for all $\psi \in \mathfrak{L}_e$ it holds that $x \models \psi$ if and only if $y \models \psi$. Note that *Sem*$_2$ ensures that the following property holds.

LEMMA 1. *Let* $(\mathcal{C}, \models)$ *be such that Sem*$_1$ *and Sem*$_2$ *are satisfied. Then*

$$\textit{If } \{\psi \in \mathfrak{L}_e : x \models \psi\} \subseteq \{\psi \in \mathfrak{L}_e : y \models \psi\}, \textit{ then } x \equiv_{\mathfrak{L}_e} y.$$

Proof. Assume $\{\psi \in \mathfrak{L}_e : x \models \psi\} \subseteq \{\psi \in \mathfrak{L}_e : y \models \psi\}$. Then for all $\psi \in \mathfrak{L}_e$ it holds that if $x \models \psi$, then $y \models \psi$. Suppose there is a $\psi \in \mathfrak{L}_e$ such that $y \models \psi$ and $x \not\models \psi$. By *Sem*$_2$, there is a $\chi \in \mathfrak{L}_e$ such that for every $z \in \mathcal{C}$, it holds that $z \models \psi$ if and only if $z \not\models \chi$. Hence, $x \models \chi$ and $y \not\models \chi$. From our assumption, it follows that $y \models \chi$. Contradiction. Hence, for all $\psi \in \mathfrak{L}_e$ it holds that if $y \models \psi$, then $x \models \psi$. Therefore, $x \equiv_{\mathfrak{L}_e} y$. □

Given a semantics for $\Vdash$ that satisfies *Sem*$_1$ and *Sem*$_2$, we can state semantic counterparts to conditions (C_1) and (C_2) of Theorem 1 and show that the counterparts characterize these conditions.

DEFINITION 2. *Let* $\phi \in \mathfrak{L} - \mathfrak{L}_e$ *and let* $(\mathcal{C}, \models)$ *be such that Sem*$_1$ *and Sem*$_2$ *are satisfied. Then the semantic conditions* (C_1') *and* (C_2') *are the following*:

$(C_1{}')$ *for every* $x \in \mathcal{C}$ *with* $x \models \phi$ *there is a* $y \in \mathcal{C}$ *with* $y \not\models \phi$ *and* $x \equiv_{\mathfrak{L}_e} y$;
$(C_2{}')$ *for every* $x \in \mathcal{C}$ *with* $x \not\models \phi$ *there is a* $y \in \mathcal{C}$ *with* $y \models \phi$ *and* $x \equiv_{\mathfrak{L}_e} y$.

Note that the semantic counterparts (C_1') and (C_2') are much stronger than what is required for disproving (unconditional) expressivity: to prove that a statement ϕ is not expressible in a language $\mathfrak{L}_e$, it suffices to show that there is *one* pair of evaluation indices x and y such that $x \models \phi$ and $y \not\models \phi$ and $x \equiv_{\mathfrak{L}_e} y$.

THEOREM 2. *Let* $\phi \in \mathfrak{L} - \mathfrak{L}_e$ *and let* $(\mathcal{C}, \models)$ *be such that Sem*$_1$ *and Sem*$_2$ *are satisfied. Then*

(C_1) *of Theorem 1 holds iff* (C_1') *of Definition 2 holds.*

Proof. ($\Rightarrow$) Assume that (C_1') does not hold. Then there is an $x \in \mathcal{C}$ such that (a) $x \models \phi$ and (b) for every $y \in \mathcal{C}$ it holds that if $x \equiv_{\mathfrak{L}_e} y$, then $y \models \phi$. Let $\Gamma = \{\psi \in \mathfrak{L}_e : x \models \psi\}$. It holds that $\Gamma \subseteq \mathfrak{L}_e$ and $x \models \Gamma$. By *Sem*$_2$, it must be that $\Gamma \nVdash \mathfrak{L}_e$. Suppose that $\Gamma \nVdash \phi$. By *Sem*$_1$, there is a $y \in \mathcal{C}$ such that $y \models \Gamma$ and $y \not\models \phi$. Note that $\{\psi \in \mathfrak{L}_e : x \models \psi\} \subseteq \{\psi \in \mathfrak{L}_e : y \models \psi\}$. By Lemma 1, it must be that $x \equiv_{\mathfrak{L}_e} y$. By (b), it must be that $y \models \phi$. Contradiction. Hence, $\Gamma \Vdash \phi$. Therefore, (C_1) of Theorem 1 does not hold.

($\Leftarrow$) Assume that (C_1') holds. Suppose that (C_1) does not hold. Then there is a $\Gamma \subseteq \mathfrak{L}_e$ such that (a) $\Gamma \nVdash \mathfrak{L}_e$ and (b) $\Gamma \Vdash \phi$. By (a) and *Sem*$_1$, there is an $x \in \mathcal{C}$ such that $x \models \Gamma$. By (b) and *Sem*$_1$, it must be that $x \models \phi$. By (C_1'), there is a $y \in \mathcal{C}$ with $y \not\models \phi$ and $x \equiv_{\mathfrak{L}_e} y$. Because $\Gamma \subseteq \mathfrak{L}_e$ and $x \models \Gamma$ and $x \equiv_{\mathfrak{L}_e} y$, it must be that $y \models \Gamma$. By (b) and *Sem*$_1$, it must be that $y \models \phi$. Contradiction. Hence, (C_1) of Theorem 1 holds. □

THEOREM 3. *Let $\phi \in \mathfrak{L} - \mathfrak{L}_e$ and let $(\mathcal{C}, \models)$ be such that Sem_1 and Sem_2 are satisfied. Then*

(C_2) *of Theorem 1 holds iff (C_2') of Definition 2 holds.*

Proof. $(\Rightarrow)$ Assume that (C_2') does not hold. Then there is an $x \in \mathcal{C}$ such that (a) $x \not\models \phi$ and (b) for every $y \in \mathcal{C}$ it holds that if $x \equiv_{\mathfrak{L}_e} y$, then $y \not\models \phi$. Let $\Gamma = \{\psi \in \mathfrak{L}_e : x \models \psi\}$. By Sem_2 and because $\mathfrak{L}_e$ is non-empty, Γ is non-empty. It holds that $\Gamma \subseteq \mathfrak{L}_e$ and $x \models \Gamma$. By Sem_2, it must be that $\Gamma \not\Vdash \mathfrak{L}_e$. Suppose that $\Gamma, \phi \not\Vdash \mathfrak{L}_e$. Then there is a $y \in \mathcal{C}$ such that $y \models \Gamma$ and $y \models \phi$. Note that $\{\psi \in \mathfrak{L}_e : x \models \psi\} \subseteq \{\psi \in \mathfrak{L}_e : y \models \psi\}$. By Lemma 1, it must be that $x \equiv_{\mathfrak{L}_e} y$. By (b), it must be that $y \not\models \phi$. Contradiction. Hence, $\Gamma, \phi \Vdash \mathfrak{L}_e$. Therefore, (C_2) of Theorem 1 does not hold.

$(\Leftarrow)$ Assume that (C_2') holds. Suppose that (C_2) does not hold. Then there is a $\Gamma \subseteq \mathfrak{L}_e$ such that (a) $\Gamma \not\Vdash \mathfrak{L}_e$ and (b) $\Gamma, \phi \Vdash \mathfrak{L}_e$. By (a) and Sem_1, there is an $x \in \mathcal{C}$ such that $x \models \Gamma$ and $x \not\models \mathfrak{L}_e$. By (b) and Sem_1, it must be that $x \not\models \phi$. By (C_2'), there is a $y \in \mathcal{C}$ with $y \models \phi$ and $x \equiv_{\mathfrak{L}_e} y$. Because $\Gamma \subseteq \mathfrak{L}_e$ and $x \models \Gamma$ and $x \equiv_{\mathfrak{L}_e} y$, it must be that $y \models \Gamma$ and $y \not\models \mathfrak{L}_e$. By (b) and Sem_1, it must be that $y \not\models \phi$. Contradiction. Hence, (C_2) of Theorem 1 holds. □

§3. The deontic logic of collective agency. In the remainder of the paper, we study conditional expressivity within the context of the deontic logic of collective agency. First, we recall its full language, its sublanguages, and its semantics ([29] provides an accessible introduction to the semantics). Second, we give the structural conditions under which two deontic game models are bisimilar with respect to a particular sublanguage. Third, we prove a general Hennessy–Milner theorem for this concept of bisimulation: two deontic game models are bisimilar with respect to a particular sublanguage if and only if the two models validate exactly the same set of statements from that sublanguage.

3.1. Languages and semantics. We fix a countable set $\mathfrak{P}$ of atoms and a finite set $\mathcal{N}$ of individual agents. We use p, q, and r as variables for atoms, ϕ, ψ, and χ as variables for statements, i, j, and k as variables for individual agents, and $\mathcal{F}$, $\mathcal{G}$, and $\mathcal{H}$ as variables for non-empty sets of individual agents. We use $-\mathcal{G}$ to refer to the complement $\mathcal{N} - \mathcal{G}$. We use $\mathbb{N}$ to refer to the set of all non-empty subsets of $\mathcal{N}$, and we use $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{Z}$ as variables for subsets of $\mathbb{N}$. Lastly, we use $\mathbf{G}$, $\mathbf{N}$, and $\mathbf{I}$ to refer to particular subsets of $\mathbb{N}$ that will be defined along the way.

Every formal language $\mathfrak{L}_{\mathbb{Y}}^{\mathbb{X}}$ to be studied in this paper is given by the following Backus–Naur form:

$$\phi := p \mid \star_{\mathcal{F}} \mid \neg\phi \mid (\phi \wedge \phi) \mid \Box\phi \mid [\mathcal{H}]\phi,$$

where p ranges over $\mathfrak{P}$ and $\mathcal{F}$ ranges over $\mathbb{X}$ and $\mathcal{H}$ ranges over $\mathbb{Y}$.

The operators $\vee$, $\rightarrow$, $\leftrightarrow$, $\Diamond$, and $\langle\mathcal{G}\rangle$ abbreviate the usual constructions. Brackets and braces are omitted if no ambiguities arise by leaving them out.

The language $\mathfrak{L}_{\mathbb{N}}^{\mathbb{N}}$ is the *full language* of our deontic logic of collective agency. Note that every language $\mathfrak{L}_{\mathbb{Y}}^{\mathbb{X}}$ is a sublanguage of $\mathfrak{L}_{\mathbb{N}}^{\mathbb{N}}$, that is, for all $\mathbb{X}$ and $\mathbb{Y}$, it holds that $\mathfrak{L}_{\mathbb{Y}}^{\mathbb{X}} \subseteq \mathfrak{L}_{\mathbb{N}}^{\mathbb{N}}$.

We specify a semantics that gives truth-conditions for the statements of the full language $\mathfrak{L}_{\mathbb{N}}^{\mathbb{N}}$ (and hence for the statements of every sublanguage $\mathfrak{L}_{\mathbb{Y}}^{\mathbb{X}}$) in terms of *deontic game models* [29].

Definition 3 (Deontic Game Model). *A* deontic game model M *is a quadruple* $\langle \mathcal{N}, (A_i), d, v\rangle$ *such that for each agent i in* $\mathcal{N}$ *it holds that* A_i *is a non-empty and finite set of actions available to agent i, such that* $d : A \to \{0,1\}$ *is a deontic ideality function, where* $A = \times_{i\in\mathcal{N}} A_i$ *and where there is at least one a in A with* $d(a) = 1$*, and such that* $v : \mathfrak{P} \to \wp(A)$ *is a valuation function.*

The set $A_{\mathcal{G}}$ of group actions that are available to a non-empty set $\mathcal{G}$ of individual agents is given by $A_{\mathcal{G}} = \times_{i\in\mathcal{G}} A_i$. We use $a_{\mathcal{G}}$ and $b_{\mathcal{G}}$ to refer to elements of $A_{\mathcal{G}}$. (Given an action profile $a \in A$ and a non-empty set $\mathcal{G}$ of individual agents, we also use $a_{\mathcal{G}}$ to refer to the combination of individual actions of $\mathcal{G}$'s members in a.) We order the group actions that are available to any (possibly singleton) set $\mathcal{G}$ of individual agents by way of a dominance relation.

Definition 4 (Simple Dominance). *Let* $M = \langle \mathcal{N}, (A_i), d, v\rangle$ *be a deontic game model. Let* $\mathcal{G} \subseteq \mathcal{N}$ *be a non-empty set of individual agents. Let* $a_{\mathcal{G}}, b_{\mathcal{G}} \in A_{\mathcal{G}}$*. Then* $a_{\mathcal{G}} \succeq_M b_{\mathcal{G}}$ *if and only if for all* $c_{-\mathcal{G}} \in A_{-\mathcal{G}}$ *it holds that* $d(a_{\mathcal{G}}, c_{-\mathcal{G}}) \geq d(b_{\mathcal{G}}, c_{-\mathcal{G}})$.

As per usual, $a_{\mathcal{G}}$ weakly dominates $b_{\mathcal{G}}$ (notation: $a_{\mathcal{G}} \succ_M b_{\mathcal{G}}$) if and only if $a_{\mathcal{G}} \succeq_M b_{\mathcal{G}}$ and $b_{\mathcal{G}} \not\succeq_M a_{\mathcal{G}}$.

A group action that is available to any (possibly singleton) set $\mathcal{G}$ of individual agents is *deontically admissible* if and only if it is not weakly dominated by any of $\mathcal{G}$'s available group actions.[17]

Definition 5 (Deontic Admissibility). *Let* $M = \langle \mathcal{N}, (A_i), d, v\rangle$ *be a deontic game model. Let* $\mathcal{G} \subseteq \mathcal{N}$ *be a non-empty set of individual agents. Then the set of* $\mathcal{G}$*'s* deontically admissible actions *in M, denoted by* $Adm_M(\mathcal{G})$*, is given by*

$$\{a_{\mathcal{G}} \in A_{\mathcal{G}} : \textit{there is no } b_{\mathcal{G}} \in A_{\mathcal{G}} \textit{ such that } b_{\mathcal{G}} \succ_M a_{\mathcal{G}}\}.$$

We can now give the truth-conditions for the statements of $\mathfrak{L}^{\mathbb{N}}_{\mathbb{N}}$.

Definition 6 (Truth-Conditions). *Let* $M = \langle \mathcal{N}, (A_i), d, v\rangle$ *be a deontic game model. Let* $\mathcal{G} \subseteq \mathcal{N}$ *be a non-empty set of individual agents. Let* $a \in A$ *be an action profile. Let* $p \in \mathfrak{P}$ *be an atom and let* $\phi, \psi \in \mathfrak{L}^{\mathbb{N}}_{\mathbb{N}}$ *be arbitrary statements. Then*

$$\begin{array}{rcl}
(M,a) \models p & \textit{iff} & a \in v(p) \\
(M,a) \models \star_{\mathcal{G}} & \textit{iff} & a_{\mathcal{G}} \in Adm_M(\mathcal{G}) \\
(M,a) \models \neg\phi & \textit{iff} & (M,a) \not\models \phi \\
(M,a) \models \phi \wedge \psi & \textit{iff} & (M,a) \models \phi \textit{ and } (M,a) \models \psi \\
(M,a) \models \Box\phi & \textit{iff} & (M,b) \models \phi \textit{ for all } b \in A \\
(M,a) \models [\mathcal{G}]\phi & \textit{iff} & (M,b) \models \phi \textit{ for all } b \in A \textit{ with } b_{\mathcal{G}} = a_{\mathcal{G}}\,.
\end{array}$$

An argument from premises Γ to a conclusion ϕ is *valid* (notation $\Gamma \Vdash \phi$) if for all deontic game models M and for all action profiles a in M, it holds that if $(M,a) \models \Gamma$, then $(M,a) \models \phi$.

[17] On admissibility in decision and game theory, see [1], [21], [24], [17], [23], and [4]. Deontic admissibility is used in [14, p. 130] to define collective obligations and in [27, pp. 200–201 and 207–209] to analyse collective rationality and backward-looking collective moral responsibility.

3.2. $\mathbb{X}/\mathbb{Y}$*-Bisimulations and Hennessy–Milner theorems.*

Definition 7 ($\mathbb{X}/\mathbb{Y}$-Bisimulation). *Let* $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{N}$. *Let* $M = \langle \mathcal{N}, (A_i), d, v \rangle$ *and* $M' = \langle \mathcal{N}, (A'_i), d', v' \rangle$ *be deontic game models. A relation* $R \subseteq A \times A'$ *is an* $\mathbb{X}/\mathbb{Y}$ *-* bisimulation *between* M *and* M' *if for all* $b \in A$ *and* $b' \in A'$ *with* $(b, b') \in R$ *it holds that*

(i) *for all* $p \in \mathfrak{P}$ *it holds that* $b \in v(p)$ *iff* $b' \in v'(p)$;
(ii) *for all* $\mathcal{F} \in \mathbb{X}$ *it holds that* $b_{\mathcal{F}} \in Adm_M(\mathcal{F})$ *iff* $b'_{\mathcal{F}} \in Adm_{M'}(\mathcal{F})$;
(iii) *for all* $c \in A$ *there is a* $c' \in A'$ *such that* $(c, c') \in R$;
(iv) *for all* $c' \in A'$ *there is a* $c \in A$ *such that* $(c, c') \in R$;
(v) *for all* $\mathcal{H} \in \mathbb{Y}$ *and* $c \in A$ *it holds that if* $c_{\mathcal{H}} = b_{\mathcal{H}}$, *then there is a* $c' \in A'$ *such that* $c'_{\mathcal{H}} = b'_{\mathcal{H}}$ *and* $(c, c') \in R$;
(vi) *for all* $\mathcal{H} \in \mathbb{Y}$ *and* $c' \in A'$ *it holds that if* $c'_{\mathcal{H}} = b'_{\mathcal{H}}$, *then there is a* $c \in A$ *such that* $c_{\mathcal{H}} = b_{\mathcal{H}}$ *and* $(c, c') \in R$.

We write $(M, a) \rightleftarrows^{\mathbb{X}}_{\mathbb{Y}} (M', a')$ if there is an $\mathbb{X}/\mathbb{Y}$-bisimulation R between M and M' such that $(a, a') \in R$.

Definition 8 ($\mathbb{X}/\mathbb{Y}$-Equivalence). *Let* $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{N}$. *Let* $M = \langle \mathcal{N}, (A_i), d, v \rangle$ *and* $M' = \langle \mathcal{N}, (A'_i), d', v' \rangle$ *be deontic game models. Then* (M, a) *and* (M', a') *are* $\mathbb{X}/\mathbb{Y}$-equivalent (*notation*: $(M, a) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', a')$) *if for all* $\psi \in \mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$, *it holds that* $(M, a) \models \psi$ *if and only if* $(M', a') \models \psi$.

We first prove a lemma, making use of a technique from [33, §3].

Lemma 2. *Let* $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{N}$. *Let* $M = \langle \mathcal{N}, (A_i), d, v \rangle$ *and* $M' = \langle \mathcal{N}, (A'_i), d', v' \rangle$ *be deontic game models. Let* $b \in A$ *and* $b' \in A'$. *If* $(M, b) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', b')$, *then for every* $c \in A$, *there is a* $c' \in A'$ *such that* $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', c')$.

Proof. Assume $(M, b) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', b')$. Take an arbitrary $c \in A$. For every $d \in A$, let $\phi_{c,d} = p \vee \neg p$ if $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M, d)$; otherwise, let $\phi_{c,d} = \psi$ for some $\psi \in \mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$ for which it holds that $(M, c) \models \psi$ and $(M, d) \not\models \psi$. Let $\phi_c = \bigwedge_{d \in A} \phi_{c,d}$. The finiteness of A ensures that ϕ_c is well defined. Note that (†) for every $d \in A$, it holds that if $(M, d) \models \phi_c$, then $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M, d)$.

Because $(M, c) \models \phi_c$, it holds that $(M, b) \models \Diamond\phi_c$. By our assumption, $(M', b') \models \Diamond\phi_c$. Then there is a $c' \in A'$ such that $(M', c') \models \phi_c$. Suppose $(M, c) \not\equiv^{\mathbb{X}}_{\mathbb{Y}} (M', c')$. Then there is a $\chi \in \mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$ such that $(M, c) \models \chi$ and $(M', c') \not\models \chi$. Then $(M', c') \models \phi_c \wedge \neg\chi$ and hence $(M', b') \models \Diamond(\phi_c \wedge \neg\chi)$. By our assumption, $(M, b) \models \Diamond(\phi_c \wedge \neg\chi)$. Then there is a $d \in A$ such that $(M, d) \models \phi_c \wedge \neg\chi$. By (†), the first conjunct entails that $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M, d)$. However, we have $(M, c) \models \chi$ and $(M, d) \not\models \chi$. Contradiction. Hence, $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', c')$. Therefore, because $c \in A$ was arbitrary, for every $c \in A$ there is a $c' \in A'$ such that $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', c')$. □

With the above in place, we can now prove the promised, generic Hennessy–Milner Theorem.[18]

[18] In modal logic, Hennessy–Milner theorems are typically proved for *image-finite* models, that is, for models in which any world is related to at most finitely many worlds. In our deontic game models, the set of ‘worlds’ is given by the set $A = \times_{i \in \mathcal{N}} A_i$ of action profiles. The set of action profiles that are related to an action profile a for a group $\mathcal{G}$ is given by $R_{\mathcal{G}}(a) = \{b \in A : a_{\mathcal{G}} = b_{\mathcal{G}}\}$. Because we require that each A_i be finite, our models are finite and hence image-finite. Note that if we were to allow for only one infinite A_i, the image $R_{\mathcal{G}}(a)$ would be infinite for every $\mathcal{G}$ such that $i \notin \mathcal{G}$ and every $a \in A$.

THEOREM 4. *Let* $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{N}$. *Then for all pointed deontic game models* (M, a) *and* (M', a'), *it holds that*

$$(M, a) \rightleftarrows^{\mathbb{X}}_{\mathbb{Y}} (M', a') \text{ iff } (M, a) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', a').$$

Proof. The left-to-right implication is proved by a straightforward structural induction on ϕ. For the right-to-left implication, assume $(M, a) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', a')$. Let $R = \{(b, b') \in A \times A' : (M, b) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', b')\}$. Note that $(a, a') \in R$. We prove that R is an $\mathbb{X}/\mathbb{Y}$-bisimulation between M and M', and hence, $(M, a) \rightleftarrows^{\mathbb{X}}_{\mathbb{Y}} (M', a')$.

Clause (i) of Definition 7 follows from the definition of R. Clauses (iii) and (iv) of Definition 7 follow from Lemma 2. Next, we show that R satisfies clauses (ii), (v), and (vi).

- (ii) Suppose $(b, b') \in R$ and $\mathcal{F} \in \mathbb{X}$. Suppose $b_{\mathcal{F}} \in \mathrm{Adm}_M(\mathcal{F})$. Then $(M, b) \models \star_{\mathcal{F}}$. By the definition of R, we have $(M', b') \models \star_{\mathcal{F}}$. Hence, $b'_{\mathcal{F}} \in \mathrm{Adm}_{M'}(\mathcal{F})$. The proof of the converse is analogous.
- (v) Suppose $(b, b') \in R$ and $\mathcal{H} \in \mathbb{Y}$ and $c \in A$. Suppose $c_{\mathcal{H}} = b_{\mathcal{H}}$. Define ϕ_c as in the proof of Lemma 2. Then $(M, c) \models \phi_c$. Because $c_{\mathcal{H}} = b_{\mathcal{H}}$, we have $(M, b) \models \langle \mathcal{H} \rangle \phi_c$. By the supposition and the definition of R, it must be that $(M', b') \models \langle \mathcal{H} \rangle \phi_c$. Then there is a $c' \in A'$ such that $c'_{\mathcal{H}} = b'_{\mathcal{H}}$ and $M', c' \models \phi_c$. By the same reasoning as in the proof of Lemma 2, we have $(M, c) \equiv^{\mathbb{X}}_{\mathbb{Y}} (M', c')$. Hence $(c, c') \in R$. Because (b, b') and $\mathcal{H}$ and c were arbitrary, it follows that for all $(b, b') \in R$ and all $\mathcal{H} \in \mathbb{Y}$ and all $c \in A$, it holds that if $c_{\mathcal{H}} = b_{\mathcal{H}}$, then there is a $c' \in A'$ such that $c'_{\mathcal{H}} = b'_{\mathcal{H}}$ and $(c, c') \in R$.
- (vi) Analogous to the proof of clause (v).

Therefore, $(M, a) \rightleftarrows^{\mathbb{X}}_{\mathbb{Y}} (M', a')$. □

§4. A necessary condition for the expressibility of $\star_{\mathcal{G}}$. In this section, we give a *necessary* condition for the conditional expressibility of the deontic admissibility statement $\star_{\mathcal{G}}$. Throughout the section, we hold fixed a non-empty subset $\mathcal{G}$ of $\mathcal{N}$. Let $\mathbf{G} = \{\mathcal{F} \in \mathbb{N} : \mathcal{G} \not\subseteq \mathcal{F}\}$ be the set of all non-empty subsets of $\mathcal{N}$ that are not supersets of $\mathcal{G}$. We prove that $\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathbf{G}}_{\mathbb{N}}$. By Theorems 1–3, it suffices to show that (C'_1) and (C'_2) hold for $\mathfrak{L}_e = \mathfrak{L}^{\mathbf{G}}_{\mathbb{N}}$ and $\phi = \star_{\mathcal{G}}$. To do so, we tweak the two transformations of deontic game models from [9] and establish some of their key properties.

To define the two transformations, in the next two sections, we use x and y as variables for the elements of $\{+, -\}^{\mathcal{N}}$. Following the notational conventions for action profiles, we use x_i to denote the projection of x onto i, and we use $x_{\mathcal{F}}$ to denote the projection of x onto $\mathcal{F}$. Accordingly, we can write the $2n$-tuple $(a_1, x_1, \ldots, a_n, x_n)$ as an ordered pair (a, x) of two n-tuples, where $a = (a_1, \ldots, a_n)$ and $x = (x_1, \ldots, x_n)$. Given an $x \in \{+, -\}^{\mathcal{N}}$ and an $\mathcal{F} \in \mathbb{N}$, we say that $x_{\mathcal{F}}$ is *even* if the number of i's in $\mathcal{F}$ such that $x_i = +$ is even. Otherwise, we say that $x_{\mathcal{F}}$ is *odd*. We make the following modest observation that crucially depends on $\mathcal{G} \not\subseteq \mathcal{F}$.

OBSERVATION. *Let* $\mathcal{F} \in \mathbf{G}$ *and let* $x_{\mathcal{F}} \in \{+, -\}^{\mathcal{F}}$. *Then*

- (i) *there is a* $y_{-\mathcal{F}} \in \{+, -\}^{\mathcal{N}-\mathcal{F}}$ *such that* $(x_{\mathcal{F}}, y_{-\mathcal{F}})_{\mathcal{G}}$ *is even*;
- (ii) *there is a* $y_{-\mathcal{F}} \in \{+, -\}^{\mathcal{N}-\mathcal{F}}$ *such that* $(x_{\mathcal{F}}, y_{-\mathcal{F}})_{\mathcal{G}}$ *is odd.*

4.1. The unit $\mathcal{G}$-transform. Given our non-empty subset $\mathcal{G}$ of $\mathcal{N}$, we transform any given deontic game model M into a larger deontic game model $M^1_{\mathcal{G}}$. We first duplicate the individual actions that are available to the individual agents in the given model by indexing each individual action in the given model with either $+$ or $-$. This results in new sets A^*_i of available individual actions, one for each individual agent i. The new set A^* of action profiles is defined as the Cartesian product $\times_{i\in\mathcal{N}} A^*_i$. The new deontic ideality function d^1 copies the 0s of the given model but tinkers with its 1s (depending on whether $x_{\mathcal{G}}$ is even). The new valuation function v^* copies the valuation function of the given model.

DEFINITION 9. *Let $M = \langle \mathcal{N}, (A_i), d, v\rangle$ be a deontic game model and let $\mathcal{G} \in \mathbb{N}$. Then the* unit $\mathcal{G}$ -transform *of M is $M^1_{\mathcal{G}} = \langle \mathcal{N}, (A^*_i), d^1, v^*\rangle$, where*

$$\begin{array}{rcl} A^*_i & = & A_i \times \{+,-\} \textit{ for every individual agent } i \in \mathcal{N} \\ d^1(a,x) & = & \begin{cases} d(a), & \textit{if } x_{\mathcal{G}} \textit{ is even} \\ 0, & \textit{if } x_{\mathcal{G}} \textit{ is odd} \end{cases} \\ (a,x) \in v^*(p) & \textit{iff} & a \in v(p). \end{array}$$

It is easy to check that $M^1_{\mathcal{G}}$ is a deontic game model. The transformation of M into $M^1_{\mathcal{G}}$ preserves admissibility for every group $\mathcal{F}$ in $\mathbf{G}$, that is, for every group $\mathcal{F}$ in $\mathbf{G}$ we have that $a_{\mathcal{F}}$ is admissible for $\mathcal{F}$ in M if and only if for all $x_{\mathcal{F}} \in \{+,-\}^{\mathcal{F}}$ it holds that $(a_{\mathcal{F}}, x_{\mathcal{F}})$ is admissible for $\mathcal{F}$ in $M^1_{\mathcal{G}}$. This follows from the following lemma.

LEMMA 3. *Let $M = \langle \mathcal{N}, (A_i), d, v\rangle$ be a deontic game model. Let $\mathcal{F} \in \mathbf{G}$ and let $a_{\mathcal{F}}, b_{\mathcal{F}} \in A_{\mathcal{F}}$ and $x, y \in \{+,-\}^{\mathcal{N}}$. Then*

(i) *if both $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are even, then $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$ iff $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succeq_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$;*
(ii) *if both $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are odd, then $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$ iff $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succeq_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$;*
(iii) *if $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succ_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$, then $a_{\mathcal{F}} \succ_M b_{\mathcal{F}}$.*

Proof. (i) Assume that both $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are even.
($\Rightarrow$) Suppose $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$. Take an arbitrary $c^*_{-\mathcal{F}} \in A^*_{-\mathcal{F}}$. Then there is a $c_{-\mathcal{F}} \in A_{-\mathcal{F}}$ and a $z_{-\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $c^*_{-\mathcal{F}} = (c_{-\mathcal{F}}, z_{-\mathcal{F}})$. Note that $(x_{\mathcal{F}}, z_{-\mathcal{F}}) \in \{+,-\}^{\mathcal{N}}$ and $(y_{\mathcal{F}}, z_{-\mathcal{F}}) \in \{+,-\}^{\mathcal{N}}$. Moreover, $(x_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}} = (x_{\mathcal{F}\cap\mathcal{G}}, z_{\mathcal{G}-\mathcal{F}})$ and $(y_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}} = (y_{\mathcal{F}\cap\mathcal{G}}, z_{\mathcal{G}-\mathcal{F}})$. Because both $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are even, there are only two cases:

(a) Both $(x_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ and $(y_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ are even. By Definition 9, it must be that $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c_{-\mathcal{F}}, z_{-\mathcal{F}}) = d(a_{\mathcal{F}}, c_{-\mathcal{F}})$ and $d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c_{-\mathcal{F}}, z_{-\mathcal{F}}) = d(b_{\mathcal{F}}, c_{-\mathcal{F}})$. By supposition, $d(a_{\mathcal{F}}, c_{-\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{-\mathcal{F}})$. Hence, $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c^*_{-\mathcal{F}}) \geq d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c^*_{-\mathcal{F}})$.
(b) Both $(x_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ and $(y_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ are odd. By Definition 9, it must be that $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c_{-\mathcal{F}}, z_{-\mathcal{F}}) = 0$ and $d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c_{-\mathcal{F}}, z_{-\mathcal{F}}) = 0$. Hence, $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c^*_{-\mathcal{F}}) \geq d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c^*_{-\mathcal{F}})$.

Because $c^*_{-\mathcal{F}}$ was arbitrary, it holds that $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c^*_{-\mathcal{F}}) \geq d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c^*_{-\mathcal{F}})$ for all $c^*_{-\mathcal{F}} \in A^*_{-\mathcal{F}}$. Therefore, $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succeq_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$.

($\Leftarrow$) Suppose $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succeq_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$. Take an arbitrary $c_{-\mathcal{F}} \in A_{-\mathcal{F}}$. By assumption and by our Observation, there must be a $z_{-\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $(x_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ and$(y_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ are even. Note that $(c_{-\mathcal{F}}, z_{-\mathcal{F}}) \in A^*_{-\mathcal{F}}$. By

Definition 9, we have $d(a_{\mathcal{F}}, c_{\_\mathcal{F}}) = d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c_{\_\mathcal{F}}, z_{\_\mathcal{F}})$ and $d(b_{\mathcal{F}}, c_{\_\mathcal{F}}) = d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c_{\_\mathcal{F}}, z_{\_\mathcal{F}})$. By supposition, it holds that $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c_{\_\mathcal{F}}, z_{\_\mathcal{F}}) \geq d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c_{\_\mathcal{F}}, z_{\_\mathcal{F}})$ and hence $d(a_{\mathcal{F}}, c_{\_\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{\_\mathcal{F}})$. Since $c_{\_\mathcal{F}}$ was arbitrary, it holds that $d(a_{\mathcal{F}}, c_{\_\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{\_\mathcal{F}})$ for all $c_{\_\mathcal{F}} \in A_{\_\mathcal{F}}$. Therefore, $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$.

(ii) Analogous to the proof of clause (i).

(iii) Assume $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succ_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$. If $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are both even or both odd, then by (i) and (ii) of this lemma, it must be that $a_{\mathcal{F}} \succ_M b_{\mathcal{F}}$. Suppose, then, that only one of them is even.

First, we prove $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$. Take an arbitrary $c_{\_\mathcal{F}} \in A_{\_\mathcal{F}}$. By our Observation, there must be a $z_{\_\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $(x_{\mathcal{F}}, z_{\_\mathcal{F}})_{\mathcal{G}}$ is odd, and hence $(y_{\mathcal{F}}, z_{\_\mathcal{F}})_{\mathcal{G}}$ is even. By Definition 9, we have $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c_{\_\mathcal{F}}, z_{\_\mathcal{F}}) = 0$. By assumption, $d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c_{\_\mathcal{F}}, z_{\_\mathcal{F}}) = 0$. By Definition 9 and because $(y_{\mathcal{F}}, z_{\_\mathcal{F}})_{\mathcal{G}}$ is even, it must be that $d(b_{\mathcal{F}}, c_{\_\mathcal{F}}) = 0$. Hence, $d(a_{\mathcal{F}}, c_{\_\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{\_\mathcal{F}})$. Since $c_{\_\mathcal{F}}$ was arbitrary, it holds that $d(a_{\mathcal{F}}, c_{\_\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{\_\mathcal{F}})$ for all $c_{\_\mathcal{F}} \in A_{\_\mathcal{F}}$. Therefore, $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$.

Second, we prove $b_{\mathcal{F}} \not\succeq_M a_{\mathcal{F}}$. Because $(b_{\mathcal{F}}, y_{\mathcal{F}}) \not\succeq_{M^1_{\mathcal{G}}} (a_{\mathcal{F}}, x_{\mathcal{F}})$, there is a $c'_{\_\mathcal{F}} \in A_{\_\mathcal{F}}$ and a $z'_{\_\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c'_{\_\mathcal{F}}, z'_{\_\mathcal{F}}) = 1$ and $d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c'_{\_\mathcal{F}}, z'_{\_\mathcal{F}}) = 0$. By Definition 9, we have $d(a_{\mathcal{F}}, c'_{\_\mathcal{F}}) = 1$.

Suppose $d(b_{\mathcal{F}}, c'_{\_\mathcal{F}}) = 1$. By our Observation, there must be a $z''_{\_\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $(y_{\mathcal{F}}, z''_{\_\mathcal{F}})_{\mathcal{G}}$ is even, and hence $(x_{\mathcal{F}}, z''_{\_\mathcal{F}})_{\mathcal{G}}$ is odd. By Definition 9, $d^1(b_{\mathcal{F}}, y_{\mathcal{F}}, c'_{\_\mathcal{F}}, z''_{\_\mathcal{F}}) = 1$ and $d^1(a_{\mathcal{F}}, x_{\mathcal{F}}, c'_{\_\mathcal{F}}, z''_{\_\mathcal{F}}) = 0$. This contradicts the assumption. Hence, $d(b_{\mathcal{F}}, c'_{\_\mathcal{F}}) = 0$. We already showed that $d(a_{\mathcal{F}}, c'_{\_\mathcal{F}}) = 1$. Therefore, $b_{\mathcal{F}} \not\succeq a_{\mathcal{F}}$. □

THEOREM 5. *Let* $M = \langle \mathcal{N}, (A_i), d, v \rangle$ *be a deontic game model and let* $R = \{(a, (a, x)) : a \in A \text{ and } x \in \{+,-\}^{\mathcal{N}}\}$. *Then* R *is a* $\mathbf{G}/\mathbb{N}$*-bisimulation between* M *and* $M^1_{\mathcal{G}}$.

Proof. Assume that $b \in A$ and $b' \in A^*$ and $(b, b') \in R$. Then $b' = (b, y)$ for some $b \in A$ and $y \in \{+,-\}^{\mathcal{N}}$. Note that $b'_i = (b_i, y_i)$ for every $i \in \mathcal{N}$. We prove clauses (i) through (vi) of Definition 7.

(i) By Definition 9.

(ii) Suppose $\mathcal{F} \in \mathbf{G}$. ($\Rightarrow$) Suppose $b'_{\mathcal{F}} \notin \mathrm{Adm}_{M^1_{\mathcal{G}}}(\mathcal{F})$. Then $(b_{\mathcal{F}}, y_{\mathcal{F}}) \notin \mathrm{Adm}_{M^1_{\mathcal{G}}}(\mathcal{F})$. Then there is a $(c_{\mathcal{F}}, z_{\mathcal{F}}) \in A^*_{\mathcal{F}}$ such that $(c_{\mathcal{F}}, z_{\mathcal{F}}) \succ_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$. By Lemma 3(iii), it must be that $c_{\mathcal{F}} \succ_M b_{\mathcal{F}}$. Therefore, $b_{\mathcal{F}} \notin \mathrm{Adm}_M(\mathcal{F})$.

($\Leftarrow$) Suppose $b_{\mathcal{F}} \notin \mathrm{Adm}_M(\mathcal{F})$. Then there is a $c_{\mathcal{F}} \in A_{\mathcal{F}}$ such that $c_{\mathcal{F}} \succ b_{\mathcal{F}}$, that is, $c_{\mathcal{F}} \succeq b_{\mathcal{F}}$ and $b_{\mathcal{F}} \not\succeq c_{\mathcal{F}}$. By Lemma 3, it must be that $(c_{\mathcal{F}}, y_{\mathcal{F}}) \succeq_{M^1_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$ and $(b_{\mathcal{F}}, y_{\mathcal{F}}) \not\succeq_{M^1_{\mathcal{G}}} (c_{\mathcal{F}}, y_{\mathcal{F}})$. Hence, $(b_{\mathcal{F}}, y_{\mathcal{F}}) \notin \mathrm{Adm}_{M^1_{\mathcal{G}}}(\mathcal{F})$. Therefore, $b'_{\mathcal{F}} \notin \mathrm{Adm}_{M^1_{\mathcal{G}}}(\mathcal{F})$.

(iii) By definition of R.

(iv) By definition of R.

(v) Suppose $\mathcal{H} \in \mathbb{Y}$ and $c \in A$. Suppose $c_{\mathcal{H}} = b_{\mathcal{H}}$. Let $c' = (c, y)$. Then $c' \in A^*$ and $(c, c') \in R$. Because $c_{\mathcal{H}} = b_{\mathcal{H}}$, it holds that $c'_{\mathcal{H}} = b'_{\mathcal{H}}$. Therefore, there is a $c' \in A^*$ such that $c'_{\mathcal{H}} = b'_{\mathcal{H}}$ and $(c, c') \in R$.

(vi) Suppose $\mathcal{H} \in \mathbb{Y}$ and $c' \in A^*$. Suppose $c'_{\mathcal{H}} = b'_{\mathcal{H}}$. Then $c' = (c, z)$ for some $c \in A$ and $z \in \{+,-\}^{\mathcal{N}}$. Then $(c, c') \in R$. Because $c'_{\mathcal{H}} = (c_{\mathcal{H}}, z_{\mathcal{H}})$ and

$b'_{\mathcal{H}} = (b_{\mathcal{H}}, y_{\mathcal{H}})$, it must be that $c_{\mathcal{H}} = b_{\mathcal{H}}$ and $y_{\mathcal{H}} = z_{\mathcal{H}}$. Therefore, there is a $c \in A$ such that $c_{\mathcal{H}} = b_{\mathcal{H}}$ and $(c, c') \in R$.

Therefore, R is a $\mathbf{G}/\mathbb{N}$-bisimulation between M and $M^1_{\mathcal{G}}$. □

4.2. The zero $\mathcal{G}$-transform. Again, given our non-empty subset $\mathcal{G}$ of $\mathcal{N}$, we transform any given deontic game model M into a larger deontic game model $M^0_{\mathcal{G}}$. The model $M^0_{\mathcal{G}}$ is exactly like M's unit $\mathcal{G}$-transform, except that the deontic ideality function d^0 copies the 1s of the given model but tinkers with its 0s (depending on whether $x_{\mathcal{G}}$ is even).

DEFINITION 10. *Let $M = \langle \mathcal{N}, (A_i), d, v\rangle$ be a deontic game model and let $\mathcal{G} \in \mathbb{N}$. Then the* zero $\mathcal{G}$ -transform *of M is $M^0_{\mathcal{G}} = \langle \mathcal{N}, (A^*_i), d^0, v^*\rangle$, where*

$$\begin{array}{rcl} A^*_i & = & A_i \times \{+,-\} \text{ for every individual agent } i \in \mathcal{N} \\ d^0(a,x) & = & \begin{cases} d(a), & \textit{if } x_{\mathcal{G}} \textit{ is even} \\ 1, & \textit{if } x_{\mathcal{G}} \textit{ is odd} \end{cases} \\ (a,x) \in v^*(p) & \textit{iff} & a \in v(p). \end{array}$$

Again, it is easy to check that $M^0_{\mathcal{G}}$ is a deontic game model. Just as the unit $\mathcal{G}$-transform of M does, the transformation of M into $M^0_{\mathcal{G}}$ preserves admissibility for every group $\mathcal{F}$ in $\mathbf{G}$.

LEMMA 4. *Let $M = \langle \mathcal{N}, (A_i), d, v\rangle$ be a deontic game model. Let $\mathcal{F} \in \mathbf{G}$ and let $a_{\mathcal{F}}, b_{\mathcal{F}} \in A_{\mathcal{F}}$ and $x, y \in \{+,-\}^{\mathcal{N}}$. Then*

(i) *if both $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are even, then $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$ iff $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succeq_{M^0_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$;*
(ii) *if both $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are odd, then $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$ iff $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succeq_{M^0_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$;*
(iii) *if $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succ_{M^0_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$, then $a_{\mathcal{F}} \succ_M b_{\mathcal{F}}$.*

Proof. The proof of (i) and (ii) is analogous to the proof of Lemma 3(i) and Lemma 3(ii), respectively.

(iii) Assume $(a_{\mathcal{F}}, x_{\mathcal{F}}) \succ_{M^0_{\mathcal{G}}} (b_{\mathcal{F}}, y_{\mathcal{F}})$. If $x_{\mathcal{F}\cap\mathcal{G}}$ and $y_{\mathcal{F}\cap\mathcal{G}}$ are both even or both odd, then by (i) and (ii) of this lemma, it must be that $a_{\mathcal{F}} \succ_M b_{\mathcal{F}}$. Suppose, then, that only one of them is even.

First, we prove $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$. Take an arbitrary $c_{-\mathcal{F}} \in A_{-\mathcal{F}}$. By our Observation, there must be a $z_{-\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $(x_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ is even, and hence $(y_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ is odd. By Definition 10, we have $d^0(b_{\mathcal{F}}, y_{\mathcal{F}}, c_{-\mathcal{F}}, z_{-\mathcal{F}}) = 1$. By assumption, $d^0(a_{\mathcal{F}}, x_{\mathcal{F}}, c_{-\mathcal{F}}, z_{-\mathcal{F}}) = 1$. By Definition 10 and because $(y_{\mathcal{F}}, z_{-\mathcal{F}})_{\mathcal{G}}$ is odd, it must be that $d(a_{\mathcal{F}}, c_{-\mathcal{F}}) = 1$. Hence, $d(a_{\mathcal{F}}, c_{-\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{-\mathcal{F}})$. Since $c_{-\mathcal{F}}$ was arbitrary, it holds that $d(a_{\mathcal{F}}, c_{-\mathcal{F}}) \geq d(b_{\mathcal{F}}, c_{-\mathcal{F}})$ for all $c_{-\mathcal{F}} \in A_{-\mathcal{F}}$. Therefore, $a_{\mathcal{F}} \succeq_M b_{\mathcal{F}}$.

Second, we prove $b_{\mathcal{F}} \not\succeq_M a_{\mathcal{F}}$. Because $(b_{\mathcal{F}}, y_{\mathcal{F}}) \not\succeq_{M^0_{\mathcal{G}}} (a_{\mathcal{F}}, x_{\mathcal{F}})$, there is a $c'_{-\mathcal{F}} \in A_{-\mathcal{F}}$ and a $z'_{-\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $d^0(a_{\mathcal{F}}, x_{\mathcal{F}}, c'_{-\mathcal{F}}, z'_{-\mathcal{F}}) = 1$ and $d^0(b_{\mathcal{F}}, y_{\mathcal{F}}, c'_{-\mathcal{F}}, z'_{-\mathcal{F}}) = 0$. By Definition 10, we have $d(b_{\mathcal{F}}, c'_{-\mathcal{F}}) = 0$.

Suppose $d(a_{\mathcal{F}}, c'_{-\mathcal{F}}) = 0$. By our Observation, there must be a $z''_{-\mathcal{F}} \in \{+,-\}^{\mathcal{N}-\mathcal{F}}$ such that $(x_{\mathcal{F}}, z''_{-\mathcal{F}})_{\mathcal{G}}$ is even, and hence $(y_{\mathcal{F}}, z''_{-\mathcal{F}})_{\mathcal{G}}$ is odd. By Definition 10, $d^0(a_{\mathcal{F}}, x_{\mathcal{F}}, c'_{-\mathcal{F}}, z''_{-\mathcal{F}}) = 0$ and $d^0(b_{\mathcal{F}}, y_{\mathcal{F}}, c'_{-\mathcal{F}}, z''_{-\mathcal{F}}) = 1$.

This contradicts the assumption. Hence, $d(a_{\mathcal{F}}, c'_{-\mathcal{F}}) = 1$. We already showed that $d(b_{\mathcal{F}}, c'_{-\mathcal{F}}) = 0$. Therefore, $b_{\mathcal{F}} \not\succeq a_{\mathcal{F}}$. □

THEOREM 6. *Let $M = \langle \mathcal{N}, (A_i), d, v\rangle$ be a deontic game model and let $R = \{(a, (a, x)) : a \in A$ and $x \in \{+,-\}^{\mathcal{N}}\}$. Then R is a $\mathsf{G}/\mathbb{N}$-bisimulation between M and $M^0_{\mathcal{G}}$.*

Proof. Analogous to the proof of Theorem 5. Use Lemma 4. □

4.3. The semantic conditions (C'_1) and (C'_2) revisited. We are now in a position to prove that the collective deontic admissibility statement $\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathsf{G}}_{\mathbb{N}}$. Given the groundwork we did in §2.3, it suffices to show that both (C'_1) and (C'_2) of Definition 2 hold for $\mathfrak{L}_e = \mathfrak{L}^{\mathsf{G}}_{\mathbb{N}}$ and $\phi = \star_{\mathcal{G}}$. We do just that using the two model transformations $M^1_{\mathcal{G}}$ and $M^0_{\mathcal{G}}$.

THEOREM 7. *For every pointed deontic game model (M, a) with $(M, a) \models \star_{\mathcal{G}}$, there is a pointed deontic game model (M', a') such that $(M', a') \not\models \star_{\mathcal{G}}$ and $(M, a) \equiv^{\mathsf{G}}_{\mathbb{N}} (M', a')$.*

Proof. Let $M = \langle \mathcal{N}, (A_i), d, v\rangle$ and $a \in A$ be such that $(M, a) \models \star_{\mathcal{G}}$. Let $M' = M^1_{\mathcal{G}}$ and let $a' = (a, x)$ for some $x \in \{+,-\}^{\mathcal{N}}$ such that $x_{\mathcal{G}}$ is odd. By Theorem 5, we have $(M, a) \rightleftarrows^{\mathsf{G}}_{\mathbb{N}} (M', a')$. By Theorem 4, we have $(M, a) \equiv^{\mathsf{G}}_{\mathbb{N}} (M', a')$. By Definition 9 and because $x_{\mathcal{G}}$ is odd, for all $c'_{-\mathcal{G}} \in A^*_{-\mathcal{G}}$, it holds that $d^1(a_{\mathcal{G}}, x_{\mathcal{G}}, c'_{-\mathcal{G}}) = 0$. Because $d(b) = 1$ for some $b \in A$, there is a $b' \in A^*$ such that $d^1(b') = 1$. Hence, $(a_{\mathcal{G}}, x_{\mathcal{G}}) \notin Adm_{M^1_{\mathcal{G}}}(\mathcal{G})$, that is, $a' \notin Adm_{M'}(\mathcal{G})$. Therefore, $(M', a') \not\models \star_{\mathcal{G}}$. □

THEOREM 8. *For every pointed deontic game model (M, a) with $(M, a) \not\models \star_{\mathcal{G}}$, there is a pointed deontic game model (M', a') such that $(M', a') \models \star_{\mathcal{G}}$ and $(M, a) \equiv^{\mathsf{G}}_{\mathbb{N}} (M', a')$.*

Proof. Analogous to the proof of Theorem 7. Let $M' = M^0_{\mathcal{G}}$ and let $a' = (a, x)$ for some $x \in \{+,-\}^{\mathcal{N}}$ such that $x_{\mathcal{G}}$ is odd. Use Theorem 6 and Definition 10 to show that $a' \in Adm_{M'}(\mathcal{G})$. □

THEOREM 9. *$\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathsf{G}}_{\mathbb{N}}$.*

Proof. By Theorems 1, 2, 3, 7, and 8. □

COROLLARY. *Let $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{N}$. Let $\mathcal{G} \in \mathbb{N}$ be such that $\mathcal{G} \notin \mathbb{X}$. Then*

If $\star_{\mathcal{G}}$ is conditionally expressible in $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$, then there is an $\mathcal{H} \in \mathbb{X}$ such that $\mathcal{G} \subset \mathcal{H}$.

Proof. By contraposition. Suppose that there is no $\mathcal{H} \in \mathbb{X}$ such that $\mathcal{G} \subset \mathcal{H}$. Then $\mathbb{X} \subseteq \mathsf{G}$ and $\mathbb{Y} \subseteq \mathbb{N}$. Hence $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}} \subseteq \mathfrak{L}^{\mathsf{G}}_{\mathbb{N}}$. By Theorem 9, it holds that $\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$. □

§5. A sufficient condition for the expressibility of $\star_{\mathcal{G}}$.

We have just seen that Theorem 9 gives a *necessary* condition for the conditional expressibility of the deontic admissibility statement $\star_{\mathcal{G}}$. In this short section, we give a *sufficient* condition, starting from the observation that if $\mathcal{G}$ is a proper subset of $\mathcal{H}$, then the collective admissibility statement $\star_{\mathcal{G}}$ is implied by the statement $\Box\star_{\mathcal{H}}$. We have the following lemma.

LEMMA 5. *Let $\mathcal{G}, \mathcal{H} \in \mathbb{N}$. Then*

$$\Box\star_{\mathcal{H}} \Vdash \star_{\mathcal{G}}, \text{ if } \mathcal{G} \subset \mathcal{H}.$$

Proof. Let $\mathcal{G}, \mathcal{H} \in \mathbb{N}$ be such that $\mathcal{G} \subset \mathcal{H}$. Let $M = \langle \mathcal{N}, (A_i), d, v \rangle$ and $a \in A$ be such that $(M, a) \not\models \star_{\mathcal{G}}$. We prove that $(M, a) \not\models \Box\star_{\mathcal{H}}$. Because $a_{\mathcal{G}} \notin Adm_M(\mathcal{G})$ and by Definition 5, there is a $b \in A$ such that $b_{\mathcal{G}} \succ_M a_{\mathcal{G}}$. Then (a) for all $c_{\text{-}\mathcal{G}} \in A_{\text{-}\mathcal{G}}$ it holds that $d(b_{\mathcal{G}}, c_{\text{-}\mathcal{G}}) \geq d(a_{\mathcal{G}}, c_{\text{-}\mathcal{G}})$ and (b) there is a $c^*_{\text{-}\mathcal{G}} \in A_{\text{-}\mathcal{G}}$ such that $d(b_{\mathcal{G}}, c^*_{\text{-}\mathcal{G}}) = 1$ and $d(a_{\mathcal{G}}, c^*_{\text{-}\mathcal{G}}) = 0$. Let $a' = (a_{\mathcal{G}}, c^*_{\text{-}\mathcal{G}})$ and $b' = (b_{\mathcal{G}}, c^*_{\text{-}\mathcal{G}})$. Note that $a'_{\mathcal{H}} = (a_{\mathcal{G}}, c^*_{\mathcal{H}\text{-}\mathcal{G}})$ and $b'_{\mathcal{H}} = (b_{\mathcal{G}}, c^*_{\mathcal{H}\text{-}\mathcal{G}})$. We prove that $b'_{\mathcal{H}} \succ_M a'_{\mathcal{H}}$, that is, (i) $b'_{\mathcal{H}} \succeq_M a'_{\mathcal{H}}$ and (ii) $a'_{\mathcal{H}} \not\succeq_M b'_{\mathcal{H}}$.

(i) Take an arbitrary $c_{\text{-}\mathcal{H}} \in A_{\text{-}\mathcal{H}}$. Then $(c^*_{\mathcal{H}\text{-}\mathcal{G}}, c_{\text{-}\mathcal{H}}) \in A_{\text{-}\mathcal{G}}$. By (a), we have $d(b'_{\mathcal{H}}, c_{\text{-}\mathcal{H}}) = d(b_{\mathcal{G}}, c^*_{\mathcal{H}\text{-}\mathcal{G}}, c_{\text{-}\mathcal{H}}) \geq d(a_{\mathcal{G}}, c^*_{\mathcal{H}\text{-}\mathcal{G}}, c_{\text{-}\mathcal{H}}) = d(a'_{\mathcal{H}}, c_{\text{-}\mathcal{H}})$. Because $c_{\text{-}\mathcal{H}}$ was arbitrary, it holds that $d(b'_{\mathcal{H}}, c_{\text{-}\mathcal{H}}) \geq d(a'_{\mathcal{H}}, c_{\text{-}\mathcal{H}})$ for all $c_{\text{-}\mathcal{H}} \in A_{\text{-}\mathcal{H}}$. Hence, $b'_{\mathcal{H}} \succeq_M a'_{\mathcal{H}}$.

(ii) By (b) and the definitions of a' and b', $d(b'_{\mathcal{H}}, c^*_{\text{-}\mathcal{H}}) = d(b_{\mathcal{G}}, c^*_{\mathcal{H}\text{-}\mathcal{G}}, c^*_{\text{-}\mathcal{H}}) = d(b_{\mathcal{G}}, c^*_{\text{-}\mathcal{G}}) = 1$ and $d(a'_{\mathcal{H}}, c^*_{\text{-}\mathcal{H}}) = d(a_{\mathcal{G}}, c^*_{\mathcal{H}\text{-}\mathcal{G}}, c^*_{\text{-}\mathcal{H}}) = d(a_{\mathcal{G}}, c^*_{\text{-}\mathcal{G}}) = 0$. Hence, $a'_{\mathcal{H}} \not\succeq_M b'_{\mathcal{H}}$.

Hence, $a'_{\mathcal{H}} \notin Adm_M(\mathcal{H})$ and $(M, a') \not\models \star_{\mathcal{H}}$. Therefore, $(M, a) \not\models \Box\star_{\mathcal{G}}$. □

Finally, we can now fulfill the promise made in the introduction and give a necessary and sufficient condition for the conditional expressibility of $\star_{\mathcal{G}}$ in a sublanguage $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$ of the full language $\mathfrak{L}^{\mathbb{N}}_{\mathbb{N}}$.

THEOREM 10. *Let* $\mathbb{X}, \mathbb{Y} \subseteq \mathbb{N}$. *Let* $\mathcal{G} \in \mathbb{N}$ *be such that* $\mathcal{G} \notin \mathbb{X}$. *Then*

$\star_{\mathcal{G}}$ *is conditionally expressible in* $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$ *iff there is an* $\mathcal{H} \in \mathbb{X}$ *such that* $\mathcal{G} \subset \mathcal{H}$.

Proof. ($\Rightarrow$) This is the Corollary from Theorem 9.

($\Leftarrow$) Suppose that there is an $\mathcal{H} \in \mathbb{X}$ such that $\mathcal{G} \subset \mathcal{H}$. Then $\Box\star_{\mathcal{H}} \in \mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$. By Lemma 5, $\Box\star_{\mathcal{H}} \Vdash \star_{\mathcal{G}}$. Note that $\Box\star_{\mathcal{H}} \nVdash \mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$. Hence, condition ($C_1$) of Theorem 1 does not hold. Therefore, $\star_{\mathcal{G}}$ is conditionally expressible in $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$. □

§6. Comparison and two applications. Let us compare this new result with the result from [9] on the conditional expressibility of collective deontic admissibility statements. Let $\mathcal{G} \in \mathbb{N}$ be a group that consists of at least two agents and let $\mathbf{I} = \{\{i\} : i \in \mathcal{N}\}$. Then $\mathfrak{L}^{\mathbf{I}}_{\mathbf{I}}$ is the individualistic sublanguage that, next to the standard operators $\neg$, $\wedge$, and $\Box$, only contains, for every agent $i \in \mathcal{N}$, agentive modalities of the type $[i]$ and individual deontic admissibility statements of the type $\star_i$. Duijf et al. [9] showed that $\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathbf{I}}_{\mathbf{I}}$. With Theorem 10, we now know that this result holds in much stronger languages. First, it does not make a difference if we add to $\mathfrak{L}^{\mathbf{I}}_{\mathbf{I}}$ all agentive modalities $[\mathcal{H}]$ with $\mathcal{H} \in \mathbb{N}$: the statement $\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathbf{I}}_{\mathbb{N}}$. Second, it does not make a difference if we add to $\mathfrak{L}^{\mathbf{I}}_{\mathbb{N}}$ all collective deontic admissibility statements $\star_{\mathcal{F}}$ with $\mathcal{F} \in \mathbb{N}$ and either (a) $\mathcal{F} \subset \mathcal{G}$, or (b) $\mathcal{F} \cap \mathcal{G} = \emptyset$, or (c) $\mathcal{F} - \mathcal{G} \neq \emptyset$ and $\mathcal{G} - \mathcal{F} \neq \emptyset$: again, the statement $\star_{\mathcal{G}}$ is not conditionally expressible in $\mathfrak{L}^{\mathbf{G}}_{\mathbb{N}}$. Third, only if there is an $\mathcal{H} \in \mathbb{X}$ such that $\mathcal{G} \subset \mathcal{H}$ is the statement $\star_{\mathcal{G}}$ conditionally expressible in $\mathfrak{L}^{\mathbb{X}}_{\mathbb{Y}}$.

Lastly, we discuss two straightforward applications of our impossibility result on conditional expressivity. A proof that a statement ϕ is not conditionally expressible in a language $\mathfrak{L}_e$ can be used to show that a particular model-theoretic property cannot be expressed in $\mathfrak{L}_e$. The argument is as follows: suppose that the statement ϕ is not conditionally expressible in a language $\mathfrak{L}_e$ and suppose that there is a non-empty class of models $\mathcal{C} = \{M : M \text{ has property } F\}$ and a subset Δ of $\mathfrak{L}_e$ such that ϕ and Δ are logically equivalent in every model M in the class $\mathcal{C}$. Because ϕ is not conditionally expressible in $\mathfrak{L}_e$ it follows that no (non-trivial) sufficient condition for the model-theoretic property F can be given by any subset Γ of $\mathfrak{L}_e$. Therefore, F cannot be expressed in $\mathfrak{L}_e$. We give two examples.[19]

First, let $\mathfrak{L}_e$ be $\mathfrak{L}_{\mathbb{N}}^{\mathbf{G}}$ and let $\mathcal{C}'$ be the class of pointed deontic game models that have exactly one deontically ideal action profile. For every (M, a) in $\mathcal{C}'$ it holds that $(M, a) \models \star_{\mathcal{G}} \leftrightarrow \bigwedge_{i \in \mathcal{G}} \star_i$. If a sufficient condition for the model-theoretic property of having exactly one deontically ideal action profile could be given by some subset Γ of $\mathfrak{L}_{\mathbb{N}}^{\mathbf{G}}$, then $\star_{\mathcal{G}}$ would be conditionally expressible in $\mathfrak{L}_{\mathbb{N}}^{\mathbf{G}}$, contradicting Theorem 9. Therefore, the model-theoretic property of having exactly one deontically ideal action profile cannot be expressed by any subset Γ of $\mathfrak{L}_{\mathbb{N}}^{\mathbf{G}}$.

Second, let $\mathbf{N} = \{\mathcal{F} \in \mathbb{N} : \mathcal{F} \neq \mathcal{N}\}$. Let $\mathfrak{L}_e$ be $\mathfrak{L}_{\mathbb{N}}^{\mathbf{N}}$ and let $\mathcal{C}''$ be the class of pointed deontic game models that only have deontically ideal action profiles. For every (M, a) in $\mathcal{C}''$ it holds that $(M, a) \models \star_{\mathcal{N}} \leftrightarrow \bigwedge_{\mathcal{F} \in \mathbf{N}} \star_{\mathcal{F}}$. Again, if a sufficient condition for the model-theoretic property of having only deontically ideal action profiles could be given by some subset Γ of $\mathfrak{L}_{\mathbb{N}}^{\mathbf{N}}$, then $\star_{\mathcal{N}}$ would be conditionally expressible in $\mathfrak{L}_{\mathbb{N}}^{\mathbf{N}}$, contradicting Theorem 9. Therefore, the model-theoretic property of having only deontically ideal action profiles cannot be expressed by any subset Γ of $\mathfrak{L}_{\mathbb{N}}^{\mathbf{N}}$.

Nonetheless, the model-theoretic property of having only deontically ideal action profiles can be expressed by the statement $\Box \star_{\mathcal{N}}$, because for every pointed deontic game model (M, a), it holds that $(M, a) \in \mathcal{C}''$ if and only if $(M, a) \models \Box \star_{\mathcal{N}}$. But, of course, $\Box \star_{\mathcal{N}}$ is not in $\mathfrak{L}_{\mathbb{N}}^{\mathbf{N}}$.

§7. Conclusion. In this paper, we located collective deontic admissibility statements within the inferential network of the deontic logic of collective agency. We introduced the novel concept of conditional expressivity and gave necessary and sufficient conditions for it from a universal logic perspective and from a semantic perspective. We proved that the collective deontic admissibility statement $\star_{\mathcal{G}}$ about a particular group $\mathcal{G}$ is conditionally expressible in a given sublanguage of the full language of the deontic logic of collective agency if and only if that sublanguage includes a collective deontic admissibility statement $\star_{\mathcal{H}}$ for some supergroup $\mathcal{H}$ of $\mathcal{G}$. Accordingly, there are no non-trivial inferential relations between the statement $\star_{\mathcal{G}}$ and the statements in the sublanguage that not only contains all individual deontic admissibility statements for ingroup and outgroup *individuals* but also contains all collective deontic admissibility statements for $\mathcal{G}$'s sub*groups*, out*groups*, and partially overlapping *groups*. We submit that our results on conditional expressivity serve as a proof of concept for future studies of expressivity.

Acknowledgments. The authors thank Diderik Batens, Thijs De Coninck, Barteld Kooi, Louwe Kuijer, Joke Meheus, Pawel Pawlowski, the two reviewers of this journal, and the audiences who attended our presentations at the University of Hamburg and Utrecht University for their questions, comments, and suggestions.

Funding. Frederik Van De Putte's research was funded by a grant from the Dutch Research Council, no. VI.Vidi.191.105.

[19] Note that if we define a *deontic game frame* as a triple $\langle \mathcal{N}, (A_i), d \rangle$, including a deontic ideality function d but excluding a valuation function v, our two examples concern frame conditions.

BIBLIOGRAPHY

[1] Arrow, K. J. (1951). Alternative approaches to the theory of choice in risk-taking situations. *Econometrica*, **19**, 404–437.

[2] Belnap, N. D., Perloff, M., & Xu, M. (2001). *Facing the Future: Agents and Choice in Our Indeterminist World*. New York: Oxford University Press.

[3] Béziau, J.-Y. (1994). Universal logic. In Childers, T., and Majer, O., editors, *Logica '94: Proceedings of the 8th International Symposium*. Prague: Filosofia Publishing House, pp. 73–93.

[4] Brandenburger, A., Friedenberg, A., & Keisler, H. J. (2008). Admissibility in games. *Econometrica*, **76**, 307–352.

[5] Carmo, J. (2010). Collective agency, direct action and dynamic operators. *Logic Journal of the IGPL*, **18**, 66–98.

[6] Carnap, R. (1936/37). Testability and meaning. *Philosophy of Science*, **3**, 419–471 and **4**, 1–40.

[7] Collins, S. (2019). *Group Duties: Their Existence and Their Implications for Individuals*. Oxford: Oxford University Press.

[8] Duijf, H. (2022). *The Logic of Responsibility Voids*. Cham: Springer.

[9] Duijf, H., Tamminga, A., & Van De Putte, F. (2021). An impossibility result on methodological individualism. *Philosophical Studies*, **178**, 4165–4185.

[10] Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (2003). *Reasoning about Knowledge*. Cambridge: The MIT Press.

[11] Hansson, S. O. (1986). Individuals and collective actions. *Theoria*, **52**, 87–97.

[12] Hempel, C. G. (1964). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: The Free Press.

[13] Hodges, W. (1997). *A Shorter Model Theory*. Cambridge: Cambridge University Press.

[14] Horty, J. F. (2001). *Agency and Deontic Logic*. New York: Oxford University Press.

[15] ———. (2011). Perspectival act utilitarianism. In Girard, P., Roy, O., and Marion, M., editors. *Dynamic Formal Epistemology*. Dordrecht: Springer, pp. 197–221.

[16] Isaacs, T. (2011). *Moral Responsibility in Collective Contexts*. New York: Oxford University Press.

[17] Kohlberg, E., & Mertens, J.-F. (1986). On the strategic stability of equilibria. *Econometrica*, **54**, 1003–1037.

[18] Kooi, B., & Tamminga, A. (2008). Moral conflicts between groups of agents. *Journal of Philosophical Logic*, **37**, 1–21.

[19] Kutz, C. (2000). *Complicity: Ethics and Law for a Collective Age*. Cambridge: Cambridge University Press.

[20] List, C., & Pettit, P. (2011). *Group Agency*. Oxford: Oxford University Press.

[21] Luce, R. D., & Raiffa, H. (1957). *Games and Decisions*. New York: John Wiley and Sons.

[22] Richardson, R. (1954). *Definition*. Oxford: Clarendon Press.

[23] Samuelson, L. (1992). Dominated strategies and common knowledge. *Games and Economic Behavior*, **4**, 284–313.

[24] Savage, L. J. (1972). *The Foundations of Statistics* (second edition). New York: Dover Publications, Inc.

[25] Schwenkenbecher, A. (2021). *Getting Our Act Together: A Theory of Collective Moral Obligations*. New York: Routledge.

[26] Sergot, M. (2021). Some forms of collectively bringing about or 'seeing to it that'. *Journal of Philosophical Logic*, **50**, 249–293.

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

[28] Tamminga, A., Duijf, H., & Van De Putte, F. (2021). Expressivity results for deontic logics of collective agency. *Synthese*, **198**, 8733–8753.

[29] Tamminga, A., & Hindriks, F. (2020). The irreducibility of collective obligations. *Philosophical Studies*, **177**, 1085–1109.

[30] Tarski, A. (1956). *Logic, Semantics, Metamathematics*. Oxford: Clarendon Press.

[31] Tollefsen, D. P. (2015). *Groups as Agents*. Cambridge: Polity Press.

[32] Tuomela, R. (1989). Collective action, supervenience, and constitution. *Synthese*, **80**, 243–266.

[33] van Benthem, J., Bezhianishvili, N., Enqvist, S., & Yu, J. (2017). Instantial neighbourhood logic. *Review of Symbolic Logic*, **10**, 116–144.

[34] van der Hoek, W., & Meyer, J.-J. Ch. (1992). Making some issues of implicit knowledge explicit. *International Journal of Foundations of Computer Science*, **3**, 193–223.

[35] van Ditmarsch, H., van der Hoek, W., & Kooi, B. (2007). *Dynamic Epistemic Logic*. Dordrecht: Springer.

ERASMUS SCHOOL OF PHILOSOPHY
ERASMUS UNIVERSITY ROTTERDAM
BURGEMEESTER OUDLAAN 50
3062 PA ROTTERDAM
THE NETHERLANDS
and
CENTRE FOR LOGIC AND PHILOSOPHY OF SCIENCE
GHENT UNIVERSITY
BLANDIJNBERG 2
9000 GHENT
BELGIUM
E-mail: vandeputte@esphil.eur.nl

DEPARTMENT OF PHILOSOPHY AND RELIGIOUS STUDIES
UTRECHT UNIVERSITY
JANSKERKHOF 13
3512 BL UTRECHT
THE NETHERLANDS
E-mail: duijf.hein@gmail.com

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF GREIFSWALD
BADERSTRAßE 6
17489 GREIFSWALD
GERMANY
E-mail: allard.tamminga@uni-greifswald.de