Deontic Logic for Strategic Games

Allard Tamminga

Abstract We develop a multi-agent deontic action logic to study the logical behaviour of two types of deontic conditionals: (1) conditional obligations, having the form "If group $\mathcal{H}$ were to perform action $\alpha_{\mathcal{H}}$, then, in group $\mathcal{F}$'s interest, group $\mathcal{G}$ ought to perform action $\alpha_{\mathcal{G}}$" and (2) conditional permissions, having the form "If group $\mathcal{H}$ were to perform action $\alpha_{\mathcal{H}}$, then, in group $\mathcal{F}$'s interest, group $\mathcal{G}$ may perform action $\alpha_{\mathcal{G}}$". First, we define a formal language for multi-agent deontic action logic and a class of consequentialist models to interpret the formulas of the language. Second, we define a transformation that converts any strategic game into a consequentialist model. Third, we show that an outcome a^* is a Nash equilibrium of a strategic game if and only if a conjunction of certain conditional permissions is true in the consequentialist model that results from the transformation of that strategic game.

1 Introduction

Deontic logic concerns the formal study of obligations, permissions, and prohibitions. Since its inception in 1951 by G.H. von Wright, it has been a lively and fruitful branch of philosophical logic that has led to a wealth of technical results on formal aspects of normativity. It has long been largely confined, however, to the formal study of norms within single-agent or even agentless contexts, because deontic logicians primarily focussed on how things ought to be rather than on what agents ought to do. The development of a multi-agent logic of agency in the late 1990s has finally made it possible to transpose deontic logic from single-agent to multi-agent settings, and to use it to investigate normative aspects of strategic

A. Tamminga (✉)
Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52, 9712 GL Groningen,
The Netherlands
e-mail: A.M.Tamminga@rug.nl

interaction.[1] As a result, *multi-agent deontic logic* studies obligations, permissions, and prohibitions within the context of formal models of strategic interaction between (groups of) agents with different preferences.

Strategic interaction between (coalitions of) players with different preferences is, of course, also studied mathematically in *game theory*. Hence, multi-agent deontic logic and game theory both study multi-agent phenomena that are largely comparable, albeit that these phenomena are approached from widely diverging perspectives: whereas deontic logicians concentrate on the formal structure of moral obligations, game theorists focus on interactive decision making. Several authors have called for an integration of deontic logic and game theory (Apostel 1960; Åqvist 1974; van Hees 1995, 1996). They did so not only to pinpoint these differences in perspective, but also to give more structure to the obligations, permissions, and prohibitions they studied. Their formalisms, however, were not so fine-grained as is the multi-agent deontic logic that was set forth only recently by Horty (2001) and Kooi and Tamminga (2008b). Built upon on the *stit* logics of agency that were developed in Kanger (1957, 1972), Pörn (1970), von Kutschera (1986) and Horty and Belnap (1995),[2] their multi-agent deontic logic provides a unified framework for the formal interpretation of group actions, group abilities, and group obligations. The question of how to establish connections between this new multi-agent deontic logic and game theory is therefore both natural and pressing. The aim of the present paper is to give a partial answer to this question: we translate strategic games into the models that we use to interpret multi-agent deontic action logic[3] and then characterize Nash equilibria of strategic games in terms of conditional permissions from this logic.[4]

Game theorists may have reservations against any interpretation of game theory in terms of concepts from moral theory. We therefore briefly discuss the conditions under which a rapprochement between game theory and moral theory is conceptually defensible. We submit that under the double assumption that the preferences that figure in our deontic logic are *extrinsic preferences* and that

---

[1] See Belnap et al. (2001) for a textbook presentation of multi-agent logics of agency.

[2] For more on the early history of *stit* logics of agency, see Hansson (1986).

[3] Pauly (2001) also establishes formal relations between strategic games and models. He uses his coalition models, which are possible worlds-based neighbourhood models, to interpret Coalition Logic, a modal logic to reason about group abilities.

[4] Other logical characterizations of Nash equilibria have been provided by Harrenstein et al. (2003), van der Hoek et al. (2005), van Benthem et al. (2005), Bonzon et al. (2006), van Benthem (2007), Roy (2008) and Lorini (2010). Their approaches differ from ours in various respects: Bonzon et al. (2006) characterize Nash equilibria of *n-player Boolean games*, whereas Harrenstein et al. (2003), van Benthem et al. (2005), van Benthem (2007) and Roy (2008) all focus on Nash equilibria of *extensive games*. Only van der Hoek et al. (2005) and Lorini (2010) characterize, just as we do here, Nash equilibria of *strategic games*. The former expresses Nash equilibria in terms of Alternating-time Temporal Logic (ATL) extended with counterfactual commitment operators and atomic propositions that capture the agents' utilities, whereas the latter uses a combination of a variant of Propositional Dynamic Logic (PDL) and preference logic to formulate Nash equilibria. We shall characterize them in terms of conditional permissions, without making the agents' utilities or preferences explicit in the language. Moreover, unlike the models proposed by van der Hoek et al. (2005) and Lorini (2010) that can be used to represent sequences of group actions, our models represent group abilities, actions, obligations, and permissions at a single moment in time.

evaluative act consequentialism is our moral theory, the conceptual gap between game theory and moral theory is minimized. (Those who are primarily interested in our formal results may skip the rest of this introduction, except for the last paragraph.)

To minimize this conceptual gap, we have to be specific on (1) the type of moral theory that gives rise to the obligations and permissions we set out to formalize, and (2) the type of preferences that figure in our deontic logic as the evaluative basis for the moral rightness of actions. Let us address the latter point first. At first sight, obligatory actions and preferred actions are worlds apart: it is perfectly possible that I have the obligation to do *X*, but at the same time prefer not-*X* to *X*. Things begin to look different, however, as soon as we make a distinction between *extrinsic* preferences (which are the result of a previous judgment of betterness on the basis of reasons) and *intrinsic* preferences (which reflect the unreasoned subjective likings of the agents concerned)—see von Wright (1963, p. 14). Preferences in game theory typically are all-things-considered, extrinsic preferences. Now, given the distinction between extrinsic and intrinsic preferences, it still makes perfect sense that I have the obligation to do *X* and at the same time *intrinsically* prefer not-*X* to *X* ("I have the obligation to pay my debts, but I prefer not to do so: I just don't feel like it"). Some intellectual effort is needed, however, to imagine a situation where I have the obligation to do *X* and at the same time *extrinsically* prefer not-*X* to *X*. Hence, a first step in bringing deontic logic and game theory together, is to assume that the preferences that figure in our deontic logic, guiding agents in evaluating the moral rightness of their actions, are *extrinsic*.

To make the conceptual match between deontic logic and game theory even closer, we also have to be specific on the type of moral theory that gives rise to the obligations and permissions we aim to formalize. From a deontological perspective, it still might be that I have the obligation to do *X* and at the same time extrinsically prefer not-*X* to *X* ("I have the obligation to return the gun to you, because I promised that, but all-things-considered I prefer not to do so"). This possibility is minimized once we adopt an evaluative version of *act consequentialism* as the moral theory that tells us where obligations and permissions come from.[5] For the sake of the argument, this is what we shall do here. In evaluative act consequentialism, the moral rightness of an action only depends on the value of its consequences: the value of the consequences of returning the gun to you is weighed against the value of the consequences of not returning the gun to you. The action with the highest-valued consequences is the morally right one. (Unlike classical act utilitarianism, where the moral rightness of an action depends on the consequences it has for *all* agents, our brand of act consequentialism acknowledges different sorts of moral rightness of an action depending on the value of its consequences for specific interest groups. To formalize these different sorts of moral rightness, we assume that each group $\mathcal{F}$ of agents has its own extrinsic preference relation over the full set of possible consequences.) It will be seen below that the modelling of obligations by way of a formal framework inspired by evaluative act consequentialism using extrinsic preferences as the basis for the evaluation of the

---

[5] See Darwall (2003) for an overview of the main positions within consequentialism.

moral rightness of actions makes it plausible that in the interest of a group $\mathcal{F}$ of agents an action is obligatory if and only if that action has according to $\mathcal{F}$ the highest-valued consequences, and that an action has according to $\mathcal{F}$ the highest-valued consequences if and only if that action is best in the ordering induced by $\mathcal{F}$'s extrinsic preference relation, and that an action is best in the ordering induced by $\mathcal{F}$'s extrinsic preference relation if and only if the group $\mathcal{F}$ of agents extrinsically prefers that action to all the other actions. The consequentialist models we shall use to interpret the formulas of multi-agent deontic action logic thus establish a strong conceptual bond between (our particular version of) deontic logic and game theory. The main purpose of this paper, however, is to establish formal connections between multi-agent deontic action logic and game theory.

The set-up of the paper is as follows. First, we define a formal language for multi-agent deontic action logic and a class of consequentialist models to formally interpret the formulas of that language. Second, we give standard definitions of strategic games and Nash equilibria. Third, we define a transformation $\mathfrak{T}$ and a valuation function v that convert any strategic game G into a consequentialist model $\langle \mathfrak{T}(G), v \rangle$. Fourth, we show, as a benchmark case for establishing formal connections between deontic logic and game theory, that an outcome a^* is a Nash equilibrium of strategic game G if and only if a finite conjunction of certain conditional permissions is true in the consequentialist model $\langle \mathfrak{T}(G), v \rangle$.

2 Multi-Agent Deontic Action Logic

The multi-agent deontic action logic to be presented in this paper studies the logical behavior of four types of deontic statements: (1) *absolute obligations* of the form "In group $\mathcal{F}$'s interest, group $\mathcal{G}$ ought to perform action $\alpha_{\mathcal{G}}$" (abbreviated as $\mathsf{O}_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}}$), (2) *conditional obligations* of the form "If group $\mathcal{H}$ were to perform action $\alpha_{\mathcal{H}}$, then, in group $\mathcal{F}$'s interest, group $\mathcal{G}$ ought to perform action $\alpha_{\mathcal{G}}$" (abbreviated as $\mathsf{O}_{\mathcal{G}}^{\mathcal{F}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})$), (3) *absolute permissions* of the form "In group $\mathcal{F}$'s interest, group $\mathcal{G}$ may perform action $\alpha_{\mathcal{G}}$" (abbreviated as $\mathsf{P}_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}}$), and (4) *conditional permissions* of the form "If group $\mathcal{H}$ were to perform action $\alpha_{\mathcal{H}}$, then, in group $\mathcal{F}$'s interest, group $\mathcal{G}$ may perform action $\alpha_{\mathcal{G}}$" (abbreviated as $\mathsf{P}_{\mathcal{G}}^{\mathcal{F}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})$).[6] (Conditions under which conditional obligations reduce to absolute obligations, and under which conditional permissions reduce to absolute permissions are given in Lemmas 1 and 3.) Obligations and permissions are closely related. Starting from the fairly common assumption that obligations with respect to a certain moral code do not recommend incompatible actions, in the current framework an obligation singles out the unique course of action that best serves the interest of a particular group. Whereas an obligation is uniquely action-guiding, a permission might leave open several options for acting. As a consequence, obligations and permissions have different logical properties (see Lemma 3).

---

[6] Conditional obligations and conditional permissions were first introduced in deontic logic by von Wright (1956, p. 509).

We interpret obligations and permissions on *consequentialist models*. These models are identical to the ones that were first set forth by Kooi and Tamminga (2008b), except for some minor details: Kooi and Tamminga (1) use utility functions rather than preference relations, (2) conceive of group utility as the arithmetical mean of individual utilities, and (3) think of utility functions as part of the interpretation of consequentialist frames rather than as an integral part of the consequentialist frames themselves.[7] Kooi and Tamminga (2008b) use their consequentialist models to interpret absolute obligations indexed by an acting group $\mathcal{G}$ and an interest group $\mathcal{F}$. Their approach differs from ours in that we here study the actions a group ought to perform, rather than the states of affairs a group ought to ensure. Moreover, Kooi and Tamminga do not discuss deontic conditionals, which are central to the present paper.

Our deontic conditionals can be used to study properties of strategic games. For example, Lemma 2 tells us that the formula $(O_{\mathcal{H}}^{\mathcal{F}} \alpha_{\mathcal{H}} \wedge O_{\mathcal{G}}^{\mathcal{F}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})) \rightarrow O_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}}$ ('deontic detachment') is not true in all consequentialist models. Given our transformation that converts any strategic game into a consequentialist model, we can ask for the class of strategic games that is characterized by the property of deontic detachment.[8] In this way, our multi-agent deontic action logic enables us to pose new questions about strategic games. The answers, however, must be postponed to another paper. Let us address the basic formalities first.

2.1 Language

Throughout the paper, we use a propositional modal language $\mathfrak{L}$ built from (1) a countable set $\mathfrak{P} = \{p_1, p_2, \ldots\}$ of atomic propositions and (2) a countable set $\mathfrak{A} = \{\alpha_{\mathcal{G}}^n : \mathcal{G} \subseteq \mathcal{N} \text{ and } n \in \mathbb{N}\}$ of atomic action propositions, where $\mathcal{N}$ is an arbitrary but fixed finite set of agents and $\mathbb{N}$ is the set of natural numbers. Thus, for each group $\mathcal{G}$ of agents there is a countable set $\mathfrak{A}_{\mathcal{G}} = \{\alpha_{\mathcal{G}}^1, \alpha_{\mathcal{G}}^2, \ldots\}$ of atomic action propositions. We use p and q as variables for atomic propositions in $\mathfrak{P}$, and $\alpha_{\mathcal{G}}$ and $\alpha_{\mathcal{H}}$ as variables for atomic action propositions in $\mathfrak{A}$. The formal language $\mathfrak{L}$ is the smallest set (in terms of set-theoretical inclusion) satisfying the conditions (i) through (ix):

(i) $\mathfrak{P} \cup \mathfrak{A} \subseteq \mathfrak{L}$
(ii)   If $\varphi \in \mathfrak{L}$, then $\neg\varphi \in \mathfrak{L}$
(iii)   If $\varphi, \psi \in \mathfrak{L}$, then $\varphi \wedge \psi \in \mathfrak{L}$
(iv)   If $\varphi \in \mathfrak{L}$, then $\Diamond\varphi \in \mathfrak{L}$
(v)   If $\alpha_{\mathcal{G}} \in \mathfrak{A}$ and $\varphi \in \mathfrak{L}$, then $[\alpha_{\mathcal{G}}]\varphi \in \mathfrak{L}$
(vi)   If $\alpha_{\mathcal{G}} \in \mathfrak{A}$ and $\mathcal{F} \subseteq \mathcal{N}$, then $O_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}} \in \mathfrak{L}$
(vii)   If $\alpha_{\mathcal{G}} \in \mathfrak{A}$ and $\mathcal{F} \subseteq \mathcal{N}$, then $P_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}} \in \mathfrak{L}$

---

[7] Kooi and Tamminga's consequentialist models are a generalization of the utilitarian models Jeff Horty constructed to formally interpret statements of the form "Agent i ought to see it that φ" and "Agent i has the ability to see it that φ" (Horty 1996, 2001). Horty's utilitarian models, in turn, are grafted on the branching-time models for *stit* logics of agency that were developed by Nuel Belnap and others.

[8] A class $\mathfrak{C}$ of consequentialist models characterizes a formula φ, if for all consequentialist models $\mathfrak{M}$ it holds that $\mathfrak{M} \in \mathfrak{C}$ if and only if $\mathfrak{M} \models \varphi$.

(viii)   If $\alpha_{\mathcal{G}}, \alpha_{\mathcal{H}} \in \mathfrak{A}$ and $\mathcal{F} \subseteq \mathcal{N}$ and $\mathcal{H} \subseteq \mathcal{N} - \mathcal{G}$, then $\mathsf{O}_{\mathcal{G}}^{\mathcal{F}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}}) \in \mathfrak{L}$

(ix)   If $\alpha_{\mathcal{G}}, \alpha_{\mathcal{H}} \in \mathfrak{A}$ and $\mathcal{F} \subseteq \mathcal{N}$ and $\mathcal{H} \subseteq \mathcal{N} - \mathcal{G}$, then $\mathsf{P}_{\mathcal{G}}^{\mathcal{F}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}}) \in \mathfrak{L}$.

We leave out brackets and braces if the omission does not give rise to ambiguities.

$\mathfrak{L}$ is rich enough to express (1) *action statements* like "Group $\mathcal{G}$ performs action $\alpha_{\mathcal{G}}$" (formalized as $\alpha_{\mathcal{G}}$), "Group $\mathcal{G}$ has the ability to perform action $\alpha_{\mathcal{G}}$" (formalized as $\Diamond \alpha_{\mathcal{G}}$), and "Group $\mathcal{G}$ sees to it that φ by performing action $\alpha_{\mathcal{G}}$" (formalized as $[\alpha_{\mathcal{G}}]\varphi$) and (2) *deontic statements* like "In his own interest, agent i may perform action α_i" (formalized as $\mathsf{P}_i^i \alpha_i$) and "If agent i were to perform action α_i, then, in the grand coalition's interest, group $\mathcal{G}$ ought to perform action $\alpha_{\mathcal{G}}$" (formalized as $\mathsf{O}_{\mathcal{G}}^{\mathcal{N}}(\alpha_{\mathcal{G}}/\alpha_i)$). How diverse these statements may be, they can all be formally interpreted in terms of consequentialist models.

2.2 Consequentialist Models

Consequentialist models are Kripke-style possible worlds models. They are built from a non-empty set of possible worlds and a finite set of agents. The models are used to interpret group abilities, actions, obligations, and permissions at a single moment in time. Each group of agents is assigned its own choice set that consists of the group's options for acting. A group of agents performs an action by choosing an option from its choice set. Each choice set $\mathfrak{F}$ is modelled as a partition of the full set of possible worlds, and hence the performance of an action by a group is modelled as a restriction of the full set of possible worlds to those worlds that are elements of the option that corresponds to the action being performed. Each group of agents has its own preference relation over the full set of possible worlds.[9]

We make these ideas precise by first defining the consequentialist frames that model group actions and group preferences. On the basis of these frames, we then define the consequentialist models that interpret the atomic (action) propositions.

Definition 1 A *consequentialist frame* $\mathfrak{F}$ is a quadruple $\langle \mathcal{W}, \mathcal{N}, Choice, (\succeq_{\mathcal{F}}) \rangle$, where $\mathcal{W}$ is a non-empty set of possible worlds, $\mathcal{N}$ is a finite set of agents, *Choice* is a partition function, and $\succeq_{\mathcal{F}}$ is a reflexive, transitive, and complete relation on $\mathcal{W}$ for each $\mathcal{F} \subseteq \mathcal{N}$.

Choice sets of *individual agents* are given by a partition function $Choice : \mathcal{N} \to \wp(\wp(\mathcal{W}))$ that meets two conditions: (1) for each agent $i \in \mathcal{N}$ it holds that $Choice(i)$ is a partition of $\mathcal{W}$, and (2) for each selection function s assigning to each agent $i \in \mathcal{N}$ a set of possible worlds $s(i)$ such that $s(i) \in Choice(i)$ it holds that $\bigcap_{i \in \mathcal{N}} s(i)$ is non-empty.[10]

---

[9] Although it seems natural to think of a group's preference relation as being somehow reducible to the preference relations of the group's constituent members (but see List and Pettit 2011), we have chosen to keep matters general by not choosing for a particular aggregation function from a wide variety of options. This policy prevents such an aggregation function from interfering with the logic. A group's preference relation represents that group's 'interests', broadly conceived. It not only covers the group's own welfare, but may well include the welfare of others who are not in that group.

[10] The second condition is the requirement of *agent independence*. It guarantees that there is a possible world in which each individual agent performs the action of his choice, irrespective of the actions all

For example, let $\mathcal{W} = \{w_1, w_2, w_3, w_4\}$ and $\mathcal{N} = \{i, j\}$. Let $Choice(i) = \{\{w_1, w_2\}, \{w_3, w_4\}\}$ and $Choice(j) = \{\{w_1, w_3\}, \{w_2, w_4\}\}$. Then $Choice$ is a partition function, since it meets the two conditions. As for condition (1), note that both $Choice(i)$ and $Choice(j)$ are partitions of $\mathcal{W}$. As for condition (2), note that there are in our example four possible selection functions:

$$
\begin{aligned}
s_1(i) &= \{w_1, w_2\} & s_1(j) &= \{w_1, w_3\} \\
s_2(i) &= \{w_1, w_2\} & s_2(j) &= \{w_2, w_4\} \\
s_3(i) &= \{w_3, w_4\} & s_3(j) &= \{w_1, w_3\} \\
s_4(i) &= \{w_3, w_4\} & s_4(j) &= \{w_2, w_4\}.
\end{aligned}
$$

Hence, for every selection function s that assigns to agent i an option $s(i)$ in $Choice(i)$ and to agent j an option $s(j)$ in $Choice(j)$ it holds that $s(i) \cap s(j) \neq \emptyset$.

Next, we generalize the partition function for individual agents to a partition function $Choice : \wp(\mathcal{N}) \to \wp(\wp(\mathcal{W}))$ for *groups of agents*. Let $Select$ be the set of all selection functions s assigning to each individual agent $i \in \mathcal{N}$ an option $s(i) \in Choice(i)$. Then

$$
Choice(\mathcal{G}) = \left\{ \bigcap_{i \in \mathcal{G}} s(i) : s \in Select \right\},
$$

if $\mathcal{G}$ is non-empty. Otherwise, $Choice(\mathcal{G}) = \{\mathcal{W}\}$.

In our example, the set $Choice(i, j)$ of actions available to the group $\{i, j\}$ is given by $\{\{w_1\}, \{w_2\}, \{w_3\}, \{w_4\}\}$, since $Select = \{s_1, s_2, s_3, s_4\}$ and hence $Choice(i, j) = \{s_1(i) \cap s_1(j), s_2(i) \cap s_2(j), s_3(i) \cap s_3(j), s_4(i) \cap s_4(j)\}$.

Definition 2 A *consequentialist model* $\mathfrak{M}$ is an ordered pair $\langle \mathfrak{F}, v \rangle$, where $\mathfrak{F}$ is a consequentialist frame and v a valuation function that assigns to each atomic proposition $p \in \mathfrak{P}$ a set of worlds $v(p) \in \wp(\mathcal{W})$ and to each atomic action proposition $\alpha_\mathcal{G} \in \mathfrak{A}$ a set of worlds $v(\alpha_\mathcal{G}) \in Choice(\mathcal{G})$.

In a consequentialist model, each possible action of each group of agents is assumed to have a name, that is, for each $\mathcal{G} \subseteq \mathcal{N}$ and each $K \in Choice(\mathcal{G})$ there is an atomic action proposition $\alpha_\mathcal{G} \in \mathfrak{A}$ such that $v(\alpha_\mathcal{G}) = K$.

2.3 Absolute and Conditional $\mathcal{F}$-Dominance

As we are currently interested in building bridges between deontic logic and game theory, we follow the conceptual guidelines of the introduction and interpret obligations and permissions in consequentialist models that make use of extrinsic preferences as the basis for the evaluation of the moral rightness of actions. In search of such a formal interpretation of conditional obligations and permissions, we adopt the distinction made above that obligations are uniquely action-guiding,

Footnote 10 continued
other individual agents perform. Hence, the definition of *Choice* ensures that at a single moment in time no individual agent can prevent any other individual agent from performing an action. See for defenses of this requirement Belnap et al. (2001, pp. 217–218 and p. 283) and Horty (2001, pp. 30–31).

whereas permissions are not. Hence, we stipulate that an action is obligatory for a group of agents if it is the single best thing the group can do. Likewise, an action is permitted for a group of agents if it is among the best things the group can do.[11] To implement these stipulations within the context of a consequentialist frame $\mathfrak{F}$, this means that for each group $\mathcal{G}$ of agents we need to order its available actions in terms of the preference relations $\succeq_{\mathcal{F}}$. Hence, we have to transform each preference relation $\succeq_{\mathcal{F}}$ into an $\mathcal{F}$-ordering of the set $Choice(\mathcal{G})$. Obviously, there are several ways to do this. Here we adopt the notion of dominance and adapt it to the present situation.[12]

Definition 3 Let $\mathfrak{F}$ be a consequentialist frame. Let $\mathcal{F}, \mathcal{G} \subseteq \mathcal{N}$ and $\mathcal{H} \subseteq \mathcal{N} - \mathcal{G}$. Let $K, K' \in Choice(\mathcal{G})$ and $L \in Choice(\mathcal{H})$. Then

$K \geq_{\mathcal{G}}^{\mathcal{F}} K'$	iff	for all $S \in Choice(\mathcal{N} - \mathcal{G})$ and for all $w, w' \in \mathcal{W}$ it holds that
		if $w \in K \cap S$ and $w' \in K' \cap S$, then $w \succeq_{\mathcal{F}} w'$.
$K \geq_{(\mathcal{G}/\mathcal{H},L)}^{\mathcal{F}} K'$	iff	for all $S \in Choice((\mathcal{N} - \mathcal{G}) - \mathcal{H})$ and for all $w, w' \in \mathcal{W}$ it holds that
		if $w \in K \cap L \cap S$ and $w' \in K' \cap L \cap S$, then $w \succeq_{\mathcal{F}} w'$.

As usual, $K >_{\mathcal{G}}^{\mathcal{F}} K'$ if and only if $K \geq_{\mathcal{G}}^{\mathcal{F}} K'$ and $K' \not\geq_{\mathcal{G}}^{\mathcal{F}} K$. In the same fashion, $K >_{(\mathcal{G}/\mathcal{H},L)}^{\mathcal{F}} K'$ if and only if $K \geq_{(\mathcal{G}/\mathcal{H},L)}^{\mathcal{F}} K'$ and $K' \not\geq_{(\mathcal{G}/\mathcal{H},L)}^{\mathcal{F}} K$.

Intuitively, it holds that $K \geq_{\mathcal{G}}^{\mathcal{F}} K'$ if and only if group $\mathcal{G}$'s action K furthers the interests of group $\mathcal{F}$ at least as well as group $\mathcal{G}$'s action K', regardless of the collective action that is performed by the group $\mathcal{N} - \mathcal{G}$ of remaining agents.

Likewise, it holds that $K \geq_{(\mathcal{G}/\mathcal{H},L)}^{\mathcal{F}} K'$ if and only if given that group $\mathcal{H}$ performs action L, group $\mathcal{G}$'s action K furthers the interests of group $\mathcal{F}$ at least as well as group $\mathcal{G}$'s action K', regardless of the collective action that is performed by the group $(\mathcal{N} - \mathcal{G}) - \mathcal{H}$ of remaining agents.

To illustrate the notions of absolute and conditional $\mathcal{F}$-dominance, consider the following (partial) description of a consequentialist frame built from three agents and eight possible worlds, where $Choice(i) = \{K_1, K_2\}$ and $Choice(j) = \{L_1, L_2\}$ and $Choice(k) = \{M_1, M_2\}$. The preference relations $\succeq_j$ and $\succeq_{i,j,k}$ are given in Fig. 1 by the left- and right-hand side utilities, respectively (the other six preference relations are left unspecified).

---

[11] Leo Apostel writes: "an act is obligatory, if it is the only act such that there is no other act equally good or better" and "an act is permissible if it can be considered as the application of a strategy such that there is no better one (there may be many equally good)" (Apostel 1960, p. 75). Similarly, in an interview with Alex Voorhoeve, Ken Binmore states: "[morality] includes the rules for *sustaining* an equilibrium in a game, which specify those actions that are permitted and those actions that will be punished" (Voorhoeve 2009, p. 140).

[12] Absolute $\mathcal{F}$-dominance was first defined in Kooi and Tamminga (2008b, p. 9). We insert an interest group $\mathcal{F}$, an acting group $\mathcal{H}$, and an option $L \in Choice(\mathcal{H})$ in Horty's Definition 5.2 (Horty 2001, p. 97) to define conditional $\mathcal{F}$-dominance.

Fig. 1 A consequentialist frame with preference relations $\succeq_j$ and $\succeq_{i,j,k}$

	L_1	L_2
K_1	$1,1$	$1,1$
K_2	$0,0$	$1,1$

M_1

	L_1	L_2
K_1	$0,0$	$0,1$
K_2	$0,1$	$0,0$

M_2

As for absolute $\mathcal{F}$-dominance, note that agent i's action K_1 furthers the interests of agent j at least as well as agent i's action K_2, regardless of the four possible collective actions of the group $\{j, k\}$. Hence, it holds that $K_1 \geq_i^{j} K_2$. Note, however, that $K_1 \not\geq_i^{i,j,k} K_2$.

As for conditional $\mathcal{F}$-dominance, note that given that agent k performs action M_1, agent i's action K_1 furthers the interests of the group $\{i, j, k\}$ at least as well as agent i's action K_2, regardless of the two possible actions of agent j. Hence, it holds that $K_1 \geq_{(i/k,M_1)}^{i,j,k} K_2$. Note, however, that $K_1 \not\geq_{(i/k,M_2)}^{i,j,k} K_2$.

In fact, absolute $\mathcal{F}$-dominance is a special case of conditional $\mathcal{F}$-dominance. Since $Choice(\emptyset) = \{\mathcal{W}\}$, it holds that $K \geq_{\mathcal{G}}^{\mathcal{F}} K'$ iff $K \geq_{(\mathcal{G}/\emptyset,\mathcal{W})}^{\mathcal{F}} K'$.

2.4 Semantics

Now that we have defined the notions of a consequentialist model, of absolute and conditional $\mathcal{F}$-dominance, we are in a position to give the semantical rules that state the conditions under which a formula φ from the language $\mathfrak{L}$ is true at a world w in a consequentialist model $\mathfrak{M}$. We write $\mathfrak{M}, w \models \varphi$ if this is the case.

Definition 4 (*Semantical Rules*) Let $\mathfrak{M} = \langle \mathfrak{F}, v \rangle$ be a consequentialist model. Let $\mathcal{F}, \mathcal{G} \subseteq \mathcal{N}$ and let $\mathcal{H} \subseteq \mathcal{N} - \mathcal{G}$. Let $w \in \mathcal{W}$ and let $p \in \mathfrak{P}$ and $\alpha_{\mathcal{G}}, \alpha_{\mathcal{H}} \in \mathfrak{A}$ and $\varphi, \psi \in \mathfrak{L}$. Then

$\mathfrak{M}, w \models p$	iff	$w \in v(p)$
$\mathfrak{M}, w \models \alpha_{\mathcal{G}}$	iff	$w \in v(\alpha_{\mathcal{G}})$
$\mathfrak{M}, w \models \neg\varphi$	iff	$\mathfrak{M}, w \not\models \varphi$
$\mathfrak{M}, w \models \varphi \wedge \psi$	iff	$\mathfrak{M}, w \models \varphi$ and $\mathfrak{M}, w \models \psi$
$\mathfrak{M}, w \models \Diamond\varphi$	iff	there is a w' such that $\mathfrak{M}, w' \models \varphi$
$\mathfrak{M}, w \models [\alpha_{\mathcal{G}}]\varphi$	iff	$w \in v(\alpha_{\mathcal{G}})$ and for all $w' \in v(\alpha_{\mathcal{G}})$ it holds that $\mathfrak{M}, w' \models \varphi$
$\mathfrak{M}, w \models \mathsf{O}_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}}$	iff	for all K in $Choice(\mathcal{G})$ with $K \neq v(\alpha_{\mathcal{G}})$ it holds that $v(\alpha_{\mathcal{G}}) >_{\mathcal{G}}^{\mathcal{F}} K$
$\mathfrak{M}, w \models \mathsf{P}_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}}$	iff	for all K in $Choice(\mathcal{G})$ with $K \neq v(\alpha_{\mathcal{G}})$ it holds that $v(\alpha_{\mathcal{G}}) \geq_{\mathcal{G}}^{\mathcal{F}} K$
$\mathfrak{M}, w \models \mathsf{O}_{\mathcal{G}}^{\mathcal{F}} (\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})$	iff	for all K in $Choice(\mathcal{G})$ with $K \neq v(\alpha_{\mathcal{G}})$ it holds that $v(\alpha_{\mathcal{G}}) >_{(\mathcal{G}/\mathcal{H},v(\alpha_{\mathcal{H}}))}^{\mathcal{F}} K$
$\mathfrak{M}, w \models \mathsf{P}_{\mathcal{G}}^{\mathcal{F}} (\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})$	iff	for all K in $Choice(\mathcal{G})$ with $K \neq v(\alpha_{\mathcal{G}})$ it holds that $v(\alpha_{\mathcal{G}}) \geq_{(\mathcal{G}/\mathcal{H},v(\alpha_{\mathcal{H}}))}^{\mathcal{F}} K$.

We adopt a standard notational convention: given a model $\mathfrak{M}$, we write $\mathfrak{M} \models \varphi$, if for all worlds w in $\mathcal{W}$ it holds that $\mathfrak{M}, w \models \varphi$. We write $\models \varphi$, if for all models $\mathfrak{M}$ it holds that $\mathfrak{M} \models \varphi$.

The following formulas are true in all models:

Lemma 1   *Let $\alpha_{\mathcal{G}}, \alpha_{\mathcal{H}}, \alpha_{\emptyset} \in \mathfrak{A}$. Then*

(i) $\models \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}}) \rightarrow \Diamond\alpha_{\mathcal{G}}$

(ii) $\models \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}}) \rightarrow \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})$

(iii) $\models \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}}) \rightarrow (\mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha'_{\mathcal{G}}/\alpha_{\mathcal{H}}) \leftrightarrow \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha'_{\mathcal{G}}/\alpha_{\mathcal{H}}))$

(iv) $\models \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}}) \rightarrow (\mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha'_{\mathcal{G}}/\alpha_{\mathcal{H}}) \leftrightarrow \Box(\alpha_{\mathcal{G}} \leftrightarrow \alpha'_{\mathcal{G}}))$

(v) $\models \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}\alpha_{\mathcal{G}} \leftrightarrow \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\emptyset})$

(vi) $\models \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}\alpha_{\mathcal{G}} \leftrightarrow \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\emptyset})$.

The first formula states that a conditional permission to perform an action implies the ability to perform that action. The second formula states that a conditional obligation to perform an action implies the corresponding conditional permission to perform that action.[13] The third and the fourth formulas are true because of the uniqueness requirement in the semantic rule for conditional obligations. The last two formulas show that an absolute obligation is a conditional obligation with a vacuous condition, and that an absolute permission is a conditional permission with a vacuous condition.

The following formulas are not true in all models:

Lemma 2   *Let $\alpha_{\mathcal{G}}, \alpha_{\mathcal{H}} \in \mathfrak{A}$. Then*

(i) $\not\models (\alpha_{\mathcal{H}} \wedge \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})) \rightarrow \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}\alpha_{\mathcal{G}}$ (factual detachment)

(ii) $\not\models (\alpha_{\mathcal{H}} \wedge \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})) \rightarrow \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}\alpha_{\mathcal{G}}$

(iii) $\not\models (\mathsf{O}^{\mathcal{F}}_{\mathcal{H}}\alpha_{\mathcal{H}} \wedge \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})) \rightarrow \mathsf{O}^{\mathcal{F}}_{\mathcal{G}}\alpha_{\mathcal{G}}$ (deontic detachment)

(iv) $\not\models (\mathsf{P}^{\mathcal{F}}_{\mathcal{H}}\alpha_{\mathcal{H}} \wedge \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})) \rightarrow \mathsf{P}^{\mathcal{F}}_{\mathcal{G}}\alpha_{\mathcal{G}}$.

The first invalidity shows that if both an action and an obligation conditional on that action are true, then the corresponding absolute obligation might still be false. The third invalidity shows that if both an absolute obligation to perform an action and an obligation conditional on that action are true, then the corresponding absolute obligation might still be false. For absolute and conditional permissions the situation is the same.

There is, however, a different way to characterize the relation between absolute and conditional permissions: group $\mathcal{G}$ may perform action $\alpha_{\mathcal{G}}$ in the interest of group $\mathcal{F}$ if and only if group $\mathcal{G}$ may perform action $\alpha_{\mathcal{G}}$ in the interest of group $\mathcal{F}$ regardless of what action any other group of agents were to perform. (This equivalence does *not* hold for absolute and conditional obligations.[14])

---

[13] In Standard Deontic Logic, $\mathsf{O}\varphi \rightarrow \mathsf{P}\varphi$ or, equivalently, $\neg(\mathsf{O}\varphi \wedge \mathsf{O}\neg\varphi)$ is known as the principle (D). Rejecting its validity is tantamount to admitting incompatible obligations. Goble (2005) presents a fairly standard deontic logic that does not validate (D).

[14] Let $\mathfrak{M}$ be a consequentialist model consisting of four worlds and two agents i and j, where $Choice(i) = \{K_1, K_2\}$ and $Choice(j) = \{L_1, L_2\}$. The preference relation $\succeq_i$ is given by the utilities $u_i(K_1 \cap L_1) = 1$ and $u_i(K_1 \cap L_2) = u_i(K_2 \cap L_1) = u_i(K_2 \cap L_2) = 0$. Let $v(\alpha_i) = K_1$ and $v(\alpha_j) = L_2$. Then it holds that $\mathfrak{M} \models \mathsf{O}^i_i\alpha_i$ and $\mathfrak{M} \not\models \mathsf{O}^i_i(\alpha_i/\alpha_j)$.

Lemma 3 (Absolute and Conditional Permissions) *Let* $\mathfrak{M} = \langle \mathfrak{F}, v \rangle$ *be a conse-quentialist model. Let* $\mathcal{F}, \mathcal{G} \subseteq \mathcal{N}$ *and let* $\mathcal{H} \subseteq \mathcal{N} - \mathcal{G}$. *Let* $\alpha_{\mathcal{G}} \in \mathfrak{A}$. *Then the following statements are equivalent:*

(i) $\mathfrak{M} \models \mathsf{P}_{\mathcal{G}}^{\mathcal{F}} \alpha_{\mathcal{G}}$
(ii) $\mathfrak{M} \models \mathsf{P}_{\mathcal{G}}^{\mathcal{F}} (\alpha_{\mathcal{G}} / \alpha_{\mathcal{H}})$ *for all* $\alpha_{\mathcal{H}} \in \mathfrak{A}$.

Proof   See Appendix.   □

It should be noted that our formal semantics provides truth conditions for a wide variety of conditional obligations and permissions, not all of which have their counterpart in game theory. Given the general form $\mathsf{P}_{\mathcal{G}}^{\mathcal{F}} (\alpha_{\mathcal{G}} / \alpha_{\mathcal{H}})$ of a conditional permission, we can distinguish at least four different types according to the specifics of $\mathcal{F}, \mathcal{G},$ and $\mathcal{H}$: (1) conditional permissions where $\mathcal{G}$ is a non-singleton group of agents, (2) conditional permissions where $\mathcal{G}$ and $\mathcal{H}$ do not partition the grand coalition, that is, where $\mathcal{G} \cup \mathcal{H} \neq \mathcal{N}$, (3) conditional permissions where the acting group $\mathcal{G}$ does not coincide with the interest group $\mathcal{F}$, that is, where $\mathcal{F} \neq \mathcal{G}$, and (4) conditional permissions of the form $\mathsf{P}_i^i (\alpha_i / \alpha_{\mathcal{N}-i})$, which differs from any of the other categories. As we have seen, these four types of conditional permissions are all evaluated by the same semantical rule. We only need the last type of conditional permissions, however, to show that our multi-agent deontic action logic is expressive enough to capture a central game theoretical notion: Nash equilibria of strategic games.

3 Nash Equilibria of Strategic Games

Strategic games are built from a finite set $N = \{1, \ldots, n\}$ of players. Each player i in N is assigned a non-empty set A_i of actions. An outcome $(a_1, \ldots, a_n)$ is reached if each player i in N chooses an action a_i from his set A_i of actions. The full set A of possible outcomes is given by the Cartesian product of all the players' sets of actions. Each player i has its own preference relation $\succsim_i$ over the full set of possible outcomes. The following definition, provided by Osborne and Rubinstein (1994, p. 14), makes these ideas precise. We also adopt their notational conventions.[15]

Definition 5   A *strategic game* G is a triple $\langle N, (A_i), (\succsim_i) \rangle$, where N is a finite set of players, for each player $i \in N$ it holds that A_i is a non-empty set of actions available to player i, and for each player $i \in N$ it holds that $\succsim_i$ is a preference relation on the set of outcomes $A = \times_{i \in N} A_i$.

Unlike Osborne and Rubinstein, we assume each A_i to be finite or countably infinite, because if we want to use our multi-agent deontic action logic to study strategic games, we have to make sure that there is a name for each action. Just like in Osborne and Rubinstein (1994, p. 7), preference relations $\succsim_i$ are assumed to be reflexive, transitive, and complete. We use a_i and a_i^* as variables for actions in A_i. Likewise, a and a^* are variables for outcomes in A. Given an outcome

---

[15] See Osborne and Rubinstein (1994, Section 1.7).

$a = (a_1, \ldots, a_n)$, we use a_{-i} to refer to the combination $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ of all the actions in a except for player i's action a_i. Given such a combination a_{-i} of actions and an action a_i, we use (a_{-i}, a_i) to refer to $(a_1, \ldots, a_n)$.

Given a strategic game $\langle N, (A_i), (\succsim_i) \rangle$, for each non-empty coalition $\mathcal{G} \subseteq N$ we define the set $A_{\mathcal{G}}$ of actions available to coalition $\mathcal{G}$ as $A_{\mathcal{G}} = \times_{i \in \mathcal{G}} A_i$. We use $a_{\mathcal{G}}$ and $a_{\mathcal{G}}^*$ as variables for actions in $A_{\mathcal{G}}$. Obviously, $A_{\mathcal{G}}$ is non-empty for each non-empty coalition $\mathcal{G}$. There are two borderline cases: (1) the set of actions $A_{\{i\}}$ available to the singleton coalition $\{i\}$ equals the set of actions A_i available to player i, and (2) the set of actions A_N available to the grand coalition N equals the set of outcomes A.

An outcome $a^* = (a_1^*, \ldots, a_n^*)$ is a Nash equilibrium of a strategic game G if no player can unilaterally enforce an outcome that he strictly prefers to a^* by choosing a different course of action. We use the following definition (Osborne and Rubinstein 1994, p. 14):

Definition 6 An outcome $a^* \in A$ is a *Nash equilibrium* of a strategic game $G = \langle N, (A_i), (\succsim_i) \rangle$ if and only if for each player $i \in N$ it holds that

$$(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i) \quad \text{for all } a_i \in A_i.$$

3.1 From Strategic Games to Consequentialist Models

Our multi-agent deontic action logic can be used to study properties of strategic games. To do so, we first have to transform strategic games into consequentialist models. The models that result from this transformation make certain formulas from the language $\mathfrak{L}$ true and others false. Some properties of strategic games can hence be expressed in terms of the formulas that are true in the consequentialist models that result from the transformation of those games. As an illustration of this claim, we show that an outcome a^* is a Nash equilibrium of a strategic game if and only if a conjunction of certain conditional permissions is true in the consequentialist model that results from the transformation of that strategic game. To make things precise, we first define a transformation $\mathfrak{T}$ that converts any strategic game G into a consequentialist frame $\mathfrak{T}(G)$. To obtain an appropriate consequentialist model $\langle \mathfrak{T}(G), v \rangle$ from this frame, we then define a suitable valuation function v.

The transformation $\mathfrak{T}$ of a strategic game G into a consequentialist frame $\mathfrak{T}(G)$ takes four steps. First, we define the set of possible worlds in the consequentialist frame as the set of outcomes in the strategic game. Second, we define the set of agents in the consequentialist frame as the set of players in the strategic game. Third, for each non-empty group of agents we define its choice set in the consequentialist frame as the set of smallest sets of outcomes in the strategic game for which that group is α-effective (the choice set of the empty group of agents is simply defined as the set of all possible worlds).[16] Fourth, each singleton group's

---

[16] A coalition $\mathcal{G}$ of players is α-effective for a set B of outcomes in a strategic game G if and only if there is an action $a_{\mathcal{G}}$ in $A_{\mathcal{G}}$ such that $\{(a_{\mathcal{G}}, a_{-\mathcal{G}}) \in A : a_{-\mathcal{G}} \in A_{-\mathcal{G}}\} \subseteq B$. Informally: regardless of the actions performed by the other players, the coalition $\mathcal{G}$ can ensure that the outcome is in the set B. The notion of α-effectivity was first defined in Moulin and Peleg (1982, p. 118). The notion is also central to Pauly's (2001) characterization of strategic games in terms of playable effectivity functions.

preference relation in the consequentialist frame is defined as its preference relation in the strategic game (preference relations for non-singleton groups of agents are simply defined as the universal relation).

Let us make these ideas clear with the following definition:

Definition 7 Let $G = \langle N, (A_i), (\succsim_i) \rangle$ be a strategic game. The quadruple $\mathfrak{T}(G) = \langle \mathcal{W}, \mathcal{N}, \text{Choice}, (\succeq_{\mathcal{F}}) \rangle$ is defined as follows:

(i) $\mathcal{W} = A$

(ii) $\mathcal{N} = N$

(iii) $\text{Choice}(\mathcal{G}) = \begin{cases} \{\{(a_{\mathcal{G}}, a_{-\mathcal{G}}) \in A : a_{-\mathcal{G}} \in A_{-\mathcal{G}}\} : a_{\mathcal{G}} \in A_{\mathcal{G}}\}, & \text{if } \mathcal{G} \neq \emptyset \\ \{\mathcal{W}\}, & \text{otherwise} \end{cases}$

(iv) $\succeq_{\mathcal{F}} = \begin{cases} \succsim_i, & \text{if } \mathcal{F} = \{i\} \\ \mathcal{W} \times \mathcal{W} & \text{otherwise.} \end{cases}$

The operator $\mathfrak{T}$ transforms strategic games into a consequentialist frames:

Theorem 1 *Let G be a strategic game. Then $\mathfrak{T}(G)$ is a consequentialist frame.*

Proof   See Appendix.   □

Note that there is no straightforward transformation T that converts consequentialist frames $\mathfrak{F}$ into strategic games $T(\mathfrak{F})$, since in a consequentialist frame $\mathfrak{F}$ it does not necessarily hold that $\text{Choice}(\mathcal{N}) = \{\{w\} : w \in \mathcal{W}\}$, that is, it might be that there are two different worlds, w_1 and w_2, such that $\{w_1, w_2\} \in \text{Choice}(\mathcal{N})$. In cases as these, the grand coalition does not have the ability to enforce a choice between w_1 and w_2. Consequentialist models, however, can simply distinguish these worlds by making an atomic proposition p true in w_1 and false in w_2. This additional structure can, of course, be studied with action statements of the type $[\alpha_G]\varphi$ and $\Diamond[\alpha_G]\varphi$. In this respect, consequentialist models offer a finer-grained analysis of strategic interaction than strategic games. Furthermore, the preference relations $\succeq_{\mathcal{F}}$ in a consequentialist frame also present an obstacle for its transformation into a strategic game. Given two options $K_1, K_2 \in \text{Choice}(\mathcal{N})$ such that $K_1 = \{w_1, w_2\}$ and $K_2 = \{w_3, w_4\}$ and $w_1 \succ_i w_3$ and $w_4 \succ_i w_2$, it is unclear how to convert the preference relation $\succeq_i$ on $\mathcal{W}$ into a preference relation $\succsim_i$ on $\text{Choice}(\mathcal{N})$.

Given a consequentialist frame $\mathfrak{T}(G)$, we now must define a valuation function v to obtain a consequentialist model $\langle \mathfrak{T}(G), v \rangle$. Since there are no atomic propositions p in $\mathfrak{P}$ to take care of, we simply put $v(p) = \mathcal{W}$ for all $p \in \mathfrak{P}$. As for the atomic action propositions, we have to be more subtle, since any valuation function v that assigns to each α_G in $\mathfrak{A}$ an action $K \in \text{Choice}(\mathcal{G})$ such that each possible action of each group of agents has a name would in principle suffice. Nevertheless, to establish a formal connection between best responses and conditional permissions, we need to keep track of which atomic action proposition α_G in $\mathfrak{A}_G$ is validated by the performance of which action a_G in A_G.

To ensure this, we use an injective map f that for each $\mathcal{G} \subseteq \mathcal{N}$ assigns to each action a_G in each A_G an atomic action proposition α_G in $\mathfrak{A}_G$. If there is an action a_G in A_G such that $f(a_G) = \alpha_G$, then we define $v_f(\alpha_G) = \{(a_G, a_{-G}) \in A : a_{-G} \in A_{-G}\}$ (note that a_G is unique, since f is injective). If there is no action a_G in A_G such that

$f(a_{\mathcal{G}}) = \alpha_{\mathcal{G}}$, then we simply put $v_f(\alpha_{\mathcal{G}}) = K$ for some unique designated $K \in Choice(\mathcal{G})$. Although $f(a_{\mathcal{G}})$ itself is strictly speaking not an atomic action proposition in $\mathfrak{A}_{\mathcal{G}}$, for convenience we write $\mathfrak{M}, w \models f(a_{\mathcal{G}})$ if and only if $\mathfrak{M}, w \models \alpha_{\mathcal{G}}$ and $f(a_{\mathcal{G}}) = \alpha_{\mathcal{G}}$.[17] Given these conventions, we ensure that for each $a_{\mathcal{G}}$ in each $A_{\mathcal{G}}$ it holds that

$$\langle \mathfrak{T}(G), v_f \rangle, w \models f(a_{\mathcal{G}}) \quad \text{iff} \quad w \in \{(a_{\mathcal{G}}, a_{-\mathcal{G}}) \in A : a_{-\mathcal{G}} \in A_{-\mathcal{G}}\}$$

Any valuation function v_f for $\mathfrak{T}(G)$ that is based on such an injection f will henceforth be called a *suitable valuation function*.

   In sum, any strategic game is embeddable in a consequentialist model, provided that all players have at most countably many actions available to them. Given these embeddings, we can now characterize Nash equilibria of strategic games in terms of conditional permissions from multi-agent deontic action logic.

3.2 Nash Equilibria and Conditional Permissions

Conditional permissions enable us to give a formal characterization of Nash equilibria of strategic games. To do so, we first characterize the notion of a best response in terms of a conditional permission, that is, we show that an action $a_i^* \in A_i$ in a strategic game G is a best response to a collective action $a_{-i}^* \in A_{-i}$ if and only if the conditional permission $\mathsf{P}_i^i(f(a_i^*)/f(a_{-i}^*))$ is true in the consequentialist model $\langle \mathfrak{T}(G), v_f \rangle$:

Theorem 2   *Let G be a strategic game and let v_f be a suitable valuation function for $\mathfrak{T}(G)$. Then*

$$(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i) \quad \text{for all } a_i \in A_i \quad \text{iff} \quad \langle \mathfrak{T}(G), v_f \rangle \models \mathsf{P}_i^i(f(a_i^*)/f(a_{-i}^*)).$$

Proof   See Appendix.   □

   Since the set N of players is finite, we immediately obtain from the previous theorem a characterization of Nash equilibria in terms of conditional permissions:

Theorem 3   *Let G be a strategic game and let v_f be a suitable valuation function for $\mathfrak{T}(G)$. Then*

$$a^* \text{ is a Nash equilibrium of } G \quad \text{iff} \quad \langle \mathfrak{T}(G), v_f \rangle \models \bigwedge_{i \in \mathcal{N}} \mathsf{P}_i^i(f(a_i^*)/f(a_{-i}^*)).$$

4 Conclusion

In this paper we have further explored the opportunities that the framework of consequentialist models offers for the study of strategic interaction. As Kooi and Tamminga (2008b) showed, the framework can be used to interpret *stit*-like obligations of the type "In the interest of group $\mathcal{F}$ of agents, group $\mathcal{G}$ of agents

---

[17] Likewise, we write $\mathfrak{M} \models \mathsf{P}_{\mathcal{G}}^{\mathcal{G}}(f(a_{\mathcal{G}})/f(a_{-\mathcal{G}}))$ if and only if $\mathfrak{M} \models \mathsf{P}_{\mathcal{G}}^{\mathcal{G}}(\alpha_{\mathcal{G}}/\alpha_{\mathcal{H}})$ and $f(a_{\mathcal{G}}) = \alpha_{\mathcal{G}}$ and $f(a_{-\mathcal{G}}) = \alpha_{\mathcal{H}}$.

ought to see to it that φ" and to study the exact structural conditions under which moral conflicts between groups of agents are possible. In the present paper the consequentialist models were used to formally interpret multi-agent deontic action logic. We gave truth conditions for an important class of multi-agent conditional obligations and permissions. It was showed that strategic games where all players have at most countably many actions available to them can be systematically transformed into consequentialist models, and that an outcome is a Nash equilibrium of a strategic game if and only if a conjunction of certain conditional permissions is true in the consequentialist model that is obtained from the transformation of that strategic game. For this characterization only a special type of conditional permissions was needed, although, as was noted in the last paragraph of Sect. 2, our semantics gives truth conditions for more general conditional permissions as well. The level of generality of consequentialist models might very well be used to further investigate systematic relations between deontic logic and game theory. We conclude with a brief illustration of this claim.

Let us first extend the relations $\succsim_i$ in a strategic game $\langle N, (A_i), (\succsim_i) \rangle$ to relations $\succsim_{\mathcal{F}}$ in an extended strategic game $\langle N, (A_i), (\succsim_{\mathcal{F}}) \rangle$ such that for each group of players $\mathcal{F} \subseteq N$ it holds that $\succsim_{\mathcal{F}}$ is a preference relation on the set of outcomes (we simplify clause (iv) of Definition 7 accordingly). Then it is easy to see that the present formalism allows for the introduction of a solution concept for cooperative games.

Definition 8   An outcome $a^* \in A$ is a *cooperative Nash equilibrium* of an extended strategic game $G = \langle N, (A_i), (\succsim_{\mathcal{F}}) \rangle$ for a partition $\{\mathcal{G}_1, \ldots, \mathcal{G}_k\}$ of N if and only if it holds that

$$\langle \mathfrak{T}(G), v_f \rangle \models \bigwedge_{i=1}^{k} \mathsf{P}_{\mathcal{G}_i}^{\mathcal{G}_i}(f(a_{\mathcal{G}_i}^*)/f(a_{-\mathcal{G}_i}^*)),$$

where v_f is a suitable valuation function for $\mathfrak{T}(G)$.

Note that this solution concept is a straightforward generalization of the notion of a Nash equilibrium. We leave the study of its properties to a future occasion.

Appendix

Proof of Lemma 3   Assume (i). Suppose $\alpha_{\mathcal{H}} \in \mathfrak{A}$. Suppose $K' \in Choice(\mathcal{G})$ with $K' \neq v(\alpha_{\mathcal{G}})$.   Suppose $S \in Choice((\mathcal{N} - \mathcal{G}) - \mathcal{H})$ and $w, w' \in \mathcal{W}$.   Suppose

$w \in v(\alpha_\mathcal{G}) \cap v(\alpha_\mathcal{H}) \cap S$ and $w' \in K' \cap v(\alpha_\mathcal{H}) \cap S$. Note that $v(\alpha_\mathcal{H}) \in Choice(\mathcal{H})$. Hence, $v(\alpha_\mathcal{H}) \cap S \in Choice(\mathcal{N} - \mathcal{G})$. By assumption, it holds that $v(\alpha_\mathcal{G}) \geq_\mathcal{G}^\mathcal{F} K'$. Hence $w \succeq_\mathcal{F} w'$. Therefore $v(\alpha_\mathcal{G}) \geq_{(\mathcal{G}/\mathcal{H}, v(\alpha_\mathcal{H}))}^\mathcal{F} K'$. Therefore $\mathfrak{M} \models \mathsf{P}_\mathcal{G}^\mathcal{F}(\alpha_\mathcal{G}/\alpha_\mathcal{H})$.

Assume (ii). Suppose $K' \in Choice(\mathcal{G})$ with $K' \neq v(\alpha_\mathcal{G})$. Suppose $S \in Choice(\mathcal{N} - \mathcal{G})$ and $w, w' \in \mathcal{W}$. Suppose $w \in v(\alpha_\mathcal{G}) \cap S$ and $w' \in K' \cap S$. Note that $S = L \cap M$ for an $L \in Choice(\mathcal{H})$ and an $M \in Choice((\mathcal{N} - \mathcal{G}) - \mathcal{H})$. Since every possible action has a name, there is an $\alpha_\mathcal{H} \in \mathfrak{A}$ such that $v(\alpha_\mathcal{H}) = L$. By assumption, it holds that $\mathfrak{M} \models \mathsf{P}_\mathcal{G}^\mathcal{F}(\alpha_\mathcal{G}/\alpha_\mathcal{H})$. Hence $v(\alpha_\mathcal{G}) \geq_{(\mathcal{G}/\mathcal{H}, v(\alpha_\mathcal{H}))}^\mathcal{F} K'$. Hence $w \succeq_\mathcal{F} w'$. Therefore $v(\alpha_\mathcal{G}) \geq_\mathcal{G}^\mathcal{F} K'$. Therefore $\mathfrak{M} \models \mathsf{P}_\mathcal{G}^\mathcal{F} \alpha_\mathcal{G}$. $\qquad \square$

Proof of Theorem 1 Consider $\mathfrak{T}(G) = \langle \mathcal{W}, \mathcal{N}, Choice, (\succeq_\mathcal{F}) \rangle$. We have to show that (1) $\mathcal{W}$ is non-empty, (2) $\mathcal{N}$ is finite, (3) $Choice(i)$ is a partition of $\mathcal{W}$ for each $i \in \mathcal{N}$, (4) $\bigcap_{i \in \mathcal{N}} s(i)$ is non-empty for each selection function s that assigns to each $i \in \mathcal{N}$ a set of possible worlds $s(i) \in Choice(i)$, (5) $Choice(\mathcal{G}) = \{\bigcap_{i \in \mathcal{G}} s(i) : s \in Select\}$ for each non-empty $\mathcal{G} \subseteq \mathcal{N}$, and (6) $\succeq_\mathcal{F}$ is a preference relation on $\mathcal{W}$ for each $\mathcal{F} \subseteq \mathcal{N}$.

Ad (1). By definition, each A_i in G is non-empty. Let $a_i \in A_i$ for each $i \in N$. Then $a = \times_{i \in N} a_i \in \times_{i \in N} A_i = A$. Therefore, $\mathcal{W}$ is non-empty.

Ad (2). By definition, N is finite. Therefore, $\mathcal{N}$ is finite.

Ad (3). Suppose that $K \in Choice(i)$ for some $i \in \mathcal{N}$. Then $K = \{(a_i, a_{-i}) \in A : a_{-i} \in A_{-i}\}$ for some $a_i \in A_i$. Since each A_j in G is non-empty, it must be that A_{-i} is non-empty as well. Let $a_{-i} \in A_{-i}$. Then $a = (a_i, a_{-i}) \in K$. Therefore, K is non-empty for each $K \in Choice(i)$ for each $i \in \mathcal{N}$.

Suppose that $a^* \in \bigcup Choice(i)$. Then there is a $K \in Choice(i)$ such that $a^* \in K$. It holds that $K = \{(a_i, a_{-i}) \in A : a_{-i} \in A_{-i}\}$ for some $a_i \in A_i$. Hence, $a^* \in A$. Therefore, $a^* \in \mathcal{W}$.

Suppose that $a^* \in \mathcal{W}$. It holds that $a^* \in A$. Hence, $a^* = (a_i^*, a_{-i}^*)$ with $a_i^* \in A_i$ and $a_{-i}^* \in A_{-i}$. Then $a^* \in \{(a_i^*, a_{-i}) \in A : a_{-i} \in A_{-i}\} \in Choice(i)$. Therefore, $a^* \in \bigcup Choice(i)$.

Suppose that there are $K, K' \in Choice(i)$ and $a^* \in \mathcal{W}$ such that $a^* \in K \cap K'$. Then $a^* \in A$ and $a_i^* \in A_i$. Hence, $K = \{(a_i^*, a_{-i}) \in A : a_{-i} \in A_{-i}\} = K'$.

Therefore, $Choice(i)$ is a partition of $\mathcal{W}$ for each $i \in \mathcal{N}$.

Ad (4). Suppose that s is a selection function that assigns to each $i \in \mathcal{N}$ a set $s(i) \in Choice(i)$. Since all $s(i)$'s are non-empty, we may take a possible world $a^i \in s(i)$ for each $i \in \mathcal{N}$. Note that $a^i \in A$ and $a_i^i \in A_i$ and $s(i) = \{(a_i^i, a_{-i}) \in A : a_{-i} \in A_{-i}\}$ for each $i \in \mathcal{N}$. Since $\mathcal{N} = \{i_1, \dots, i_n\}$ is finite, it must be that $(a_1^1, \dots, a_n^n) \in s(i)$ for each $i \in \mathcal{N}$. Therefore, $\bigcap_{i \in \mathcal{N}} s(i)$ is non-empty.

Ad (5). Suppose that $K \in Choice(\mathcal{G})$ for a non-empty $\mathcal{G} \subseteq \mathcal{N}$. Then $K = \{(a_\mathcal{G}^*, a_{-\mathcal{G}}) \in A : a_{-\mathcal{G}} \in A_{-\mathcal{G}}\}$ for some $a_\mathcal{G}^* \in A_\mathcal{G}$. Let $s(i) = \{(a_i^*, a_{-i}) \in A : a_{-i} \in A_{-i}\}$ for each $i \in \mathcal{G}$. It is easy to show that $K = \bigcap_{i \in \mathcal{G}} s(i)$. Therefore $Choice(\mathcal{G}) \subseteq \{\bigcap_{i \in \mathcal{G}} s(i) : s \in Select\}$.

Suppose that $K \in \{\bigcap_{i \in \mathcal{G}} s(i) : s \in Select\}$ for a non-empty $\mathcal{G} \subseteq \mathcal{N}$. Then $K = \bigcap_{i \in \mathcal{G}} s(i)$ for some $s \in Select$. Since $s(i) \in Choice(i)$ for each $i \in \mathcal{G}$, it must be that

for each $i \in \mathcal{G}$ there is an $a_i^* \in A_i$ such that $s(i) = \{(a_i^*, a_{-i}) \in A : a_{-i} \in A_{-i}\}$. Consider $a_{\mathcal{G}}^* = \times_{i \in \mathcal{G}} a_i^*$. It is easy to show that $K = \{(a_{\mathcal{G}}^*, a_{-\mathcal{G}}) \in A : a_{-\mathcal{G}} \in A_{-\mathcal{G}}\}$. Therefore $\{\bigcap_{i \in \mathcal{G}} s(i) : s \in Select\} \subseteq Choice(\mathcal{G})$.

Ad (6). By definition, $\succsim_i$ is a preference relation on A for each $i \in N$. Therefore, for each $\mathcal{F} \subseteq \mathcal{N}$ it holds that either $\succeq_{\mathcal{F}}$ is the preference relation $\succsim_i$ or $\succeq_{\mathcal{F}}$ is the preference relation $\mathcal{W} \times \mathcal{W}$. $\qquad\square$

Proof of Theorem 2 Assume that $(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i)$ for all $a_i \in A_i$. Consider $\langle \mathfrak{T}(G), v_f \rangle = \langle \mathcal{W}, \mathcal{N}, Choice, \succeq_{\mathcal{F}}, v_f \rangle$. Let $f(a_i^*) = \alpha_i^*$ and $f(a_{-i}^*) = \alpha_{-i}^*$. Note that $v_f(\alpha_i^*) \in Choice(i)$ and $v_f(\alpha_{-i}^*) \in Choice(\mathcal{N} - i)$. Suppose that $K \in Choice(i)$ such that $K \neq v_f(a_i^*)$. By the construction of $\mathfrak{T}(G)$ and v_f, there must be an $a_i \in A_i$ such that $f(a_i) = \alpha_i$ and $K = v_f(\alpha_i)$. It suffices to show that $v_f(\alpha_i^*) \geq_{(i/\mathcal{N}-i, v_f(\alpha_{-i}^*))}^{i} v_f(\alpha_i)$. Suppose that $S \in Choice((\mathcal{N} - i) - (\mathcal{N} - i))$ and that $w \in v_f(\alpha_i^*) \cap v_f(\alpha_{-i}^*) \cap S$ and $w' \in v_f(\alpha_i) \cap v_f(\alpha_{-i}^*) \cap S$. Note that $Choice((\mathcal{N} - i) - (\mathcal{N} - i)) = Choice(\emptyset) = \{\mathcal{W}\}$. Note that $v_f(\alpha_i^*) \cap v_f(\alpha_{-i}^*) = \{(a_{-i}^*, a_i^*)\}$ and $v_f(\alpha_i) \cap v_f(\alpha_{-i}^*) = \{(a_{-i}^*, a_i)\}$. Hence, $w = (a_{-i}^*, a_i^*)$ and $w' = (a_{-i}^*, a_i)$. By assumption, we have $(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i)$. Hence, by the construction of $\mathfrak{T}(G)$, it must be that $w \succeq_i w'$. Therefore, $\langle \mathfrak{T}(G), v_f \rangle \models \mathsf{P}_i^i(f(a_i^*)/f(a_{-i}^*))$.

Assume that $\langle \mathfrak{T}(G), v_f \rangle \models \mathsf{P}_i^i(f(a_i^*)/f(a_{-i}^*))$. Let $f(a_i^*) = \alpha_i^*$ and $f(a_{-i}^*) = \alpha_{-i}^*$. Suppose that $a_i \in A_i$. Suppose that $a_i = a_i^*$. By reflexivity of $\succsim_i$, we have $(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i)$. Suppose that $a_i \neq a_i^*$. Let $f(a_i) = \alpha_i$. By the construction of $\mathfrak{T}(G)$ and v_f, it must be that $v_f(\alpha_i^*), v_f(\alpha_i) \in Choice(i)$ and $v_f(\alpha_{-i}^*) \in Choice(\mathcal{N} - i)$ and $v_f(\alpha_i^*) \neq v_f(\alpha_i)$. By assumption, we have $v_f(\alpha_i^*) \geq_{(i/\mathcal{N}-i, v_f(\alpha_{-i}^*))}^{i} v_f(\alpha_i)$. Hence, it must be that $(a_{-i}^*, a_i^*) \succeq_i (a_{-i}^*, a_i)$. By the construction of $\mathfrak{T}(G)$, it must be that $(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i)$. Therefore, $(a_{-i}^*, a_i^*) \succsim_i (a_{-i}^*, a_i)$ for all $a_i \in A_i$. $\qquad\square$

References

Apostel, L. (1960). Game theory and the interpretation of deontic logic. *Logique & Analyse, 3*, 70–90.

Åqvist, L. (1974). A new approach to the logical theory of actions and causality. In S. Stenlund (Ed.), *Logical theory and semantic analysis* (pp. 73–91). Dordrecht: D. Reidel Publishing Company.

Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future*. New York: Oxford University Press.

Bonzon, E., Lagasquie-Schiex, M.-C., Lang, J., & Zanuttini, B. (2006). Boolean games revisited. In G. Brewka, S. Coradeschi, A. Perini, & P. Traverso (Eds.), *ECAI 2006. 17th European conference on artificial intelligence* (pp. 265–269). Amsterdam: IOS Press.

Darwall, S. L. (Ed.) (2003). *Consequentialism*. Malden, MA: Blackwell Publishers.

Goble, L. (2005). A logic for deontic dilemmas. *Journal of Applied Logic, 3*, 461–483.

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

Harrenstein, B. P., van der Hoek, W., Meyer, J.-J., & Witteveen, C. (2003). A modal characterization of Nash equilibrium. *Fundamenta Informaticae, 57*, 281–321.

Horty, J. F. (1996). Agency and obligation. *Synthese, 108*, 269–307.

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

Horty, J. F., & Belnap, N. (1995). The deliberative stit: A study of action, omission, ability, and obligation. *Journal of Philosophical Logic, 24*, 583–644.

Kanger, S. (1957). New foundations for ethical theory. Stockholm: Privately distributed pamphlet. First published in 1971, R. Hilpinen (Ed.), *Deontic logic: Introductory and systematic readings* (pp.

36–58). Dordrecht: D. Reidel Publishing Company. Reprinted in 2001, G. Holmström-Hintikka, S. Lindström, & R. Sliwinski (Eds.), *Collected papers of Stig Kanger with essays on his life and work, vol. 1* (pp. 99–119). Dordrecht: Kluwer.

Kanger, S. (1972). Law and logic. *Theoria, 38*, 105–132. Reprinted in 2001, G. Holmström-Hintikka, S. Lindström, & R. Sliwinski (Eds.), *Collected papers of Stig Kanger with essays on his life and work, vol. 1* (pp. 146–169). Dordrecht: Kluwer.

Kooi, B., & Tamminga, A. (2008a). Conditional obligations in strategic situations. In G. Boella, G. Pigozzi, M. Singh, & H. Verhagen (Eds.), *NorMAS 2008. Third international workshop on normative multiagent systems* (pp. 188–200). Luxembourg: Université du Luxembourg.

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

List, C., & Pettit, P. (2011). *Group agency*. Oxford: Oxford University Press.

Lorini, E. (2010). A dynamic logic of agency II: Deterministic $\mathcal{DLA}$, coalition logic, and game theory. *Journal of Logic, Language and Information, 19*, 327–351.

Moulin, H., & Peleg, B. (1982). Cores of effectivity functions and implementation theory. *Journal of Mathematical Economics, 10*, 115–145.

Osborne, M. J., & Rubinstein, A. (1994). *A course in game theory*. Cambridge, MA: The MIT Press.

Pauly, M. (2001). *Logic for social software*. Amsterdam: ILLC Dissertation Series.

Pörn, I. (1970). *The logic of power*. Oxford: Basil Blackwell.

Roy, O. (2008). *Thinking before acting*. Amsterdam: ILLC Dissertation Series.

van Benthem, J. (2007). Rational dynamics and epistemic logic in games. *International Game Theory Review, 9*, 13–45.

van Benthem, J., van Otterloo, S., & Roy, O. (2005). *Preference logic, conditionals and solution concepts in games*. ILLC Research Report PP-2005-28. Amsterdam: Universiteit van Amsterdam.

van der Hoek, W., Jamroga, W., & Wooldridge, M. (2005) A logic for strategic reasoning. In F. Dignum, V. Dignum, S. Koenig, S. Kraus, M.P. Singh, M. Wooldridge (Eds.), *AAMAS 2005. Proceedings of the fourth international joint conference on autonomous agents and multiagent systems* (pp. 157–164). New York: ACM Press.

van Hees, M. (1995). *Rights and decisions*. Dordrecht: Kluwer.

van Hees, M. (1996). A game-theoretic logic of norms and actions. *Logique & Analyse, 155–156*, 229–241.

von Kutschera, F. (1986). Bewirken. *Erkenntnis, 24*, 253–281.

von Wright, G. H. (1951). Deontic logic. *Mind, 60*, 1–15.

von Wright, G. H. (1956). A note on deontic logic and derived obligation. *Mind, 65*, 507–509.

von Wright, G. H. (1963). *The logic of preference*. Edinburgh: Edinburgh University Press.

Voorhoeve, A. (2009). *Conversations on ethics*. Oxford: Oxford University Press.