Journal of Philosophical Logic (2008) 37:1–21
DOI 10.1007/s10992-007-9049-z

BARTELD KOOI AND ALLARD TAMMINGA

MORAL CONFLICTS BETWEEN GROUPS OF AGENTS

Received 16 June 2006

ABSTRACT. Two groups of agents, $\mathcal{G}_1$ and $\mathcal{G}_2$, face a *moral conflict* if $\mathcal{G}_1$ has a moral obligation and $\mathcal{G}_2$ has a moral obligation, such that these obligations cannot both be fulfilled. We study moral conflicts using a multi-agent deontic logic devised to represent reasoning about sentences like 'In the interest of group $\mathcal{F}$ of agents, group $\mathcal{G}$ of agents ought to see to it that ϕ'. We provide a formal language and a consequentialist semantics. An illustration of our semantics with an analysis of the Prisoner's Dilemma follows. Next, necessary and sufficient conditions are given for (1) the possibility that a single group of agents faces a moral conflict, for (2) the possibility that two groups of agents face a moral conflict within a single moral code, and for (3) the possibility that two groups of agents face a moral conflict.

KEY WORDS: consequentialism, moral conflicts, multi-agent deontic logic, *stit* logic

for Erik Krabbe

1. INTRODUCTION

Moral conflicts, their existence, and the conditions for their existence are among the divisive elements in meta-ethics. In the philosophical literature, moral conflicts are usually studied from the standpoint of a single agent.[1] From such a single-agent point of view, an agent faces a moral conflict if the agent has two moral obligations that cannot both be fulfilled. By raising the study of moral conflicts from a single-agent to a multi-agent perspective, we generalize the concept of moral conflict: two groups of agents, $\mathcal{G}_1$ and $\mathcal{G}_2$, face a moral conflict if $\mathcal{G}_1$ has a moral obligation and $\mathcal{G}_2$ has a moral obligation, such that these obligations cannot both be fulfilled. (If the two groups are identical and consist of a single agent, the moral conflict boils down to a single-agent moral conflict in the usual sense.)

Given two moral obligations, one may ask whether these obligations stem from the same moral code or not. In Sophocles's *Antigone*, Creon's obligation not to bury the traitor Polynices stems from the civic values of the city he represents, whereas Antigone's obligation to bury her brother Polynices stems from religious and family values. For reasons of uniformity, we shall confine our present investigation to moral codes that can be formulated in a consequentialist fashion. More specifically, each group $\mathcal{F}$ of agents defines a moral code stipulating that $\mathcal{F}$'s collective

interest is to be maximized. Accordingly, the group of *all* agents defines the moral code of utilitarianism: an agent has a certain utilitarian obligation if this obligation stems from the moral code that the collective interest of the group of all agents is to be maximized. Moreover, an agent only accepting the moral code defined by himself is an ethical egoist, who is to maximize his own self-interest. Henceforth, a moral obligation is indexed by two groups of agents, $\mathcal{G}$ and $\mathcal{F}$. $\mathcal{G}$ indicates the group who has the obligation. $\mathcal{F}$ refers to the interest group who defines the consequentialist moral code from which the obligation stems.

Fulfilling a moral obligation involves doing what one ought to do. Hence, J.L. Austin hit the mark with his suggestion that "before we consider what actions are good or bad, right or wrong, it is proper to consider first what is meant by, and what not, (..) the expression 'doing an action' or 'doing something'" (Austin, 1957, p. 178).[2] In our logical analysis of moral conflicts between groups of agents, we adopt Austin's suggestion. On the basis of the well established *stit* logics of agency developed by Nuel Belnap and others (Belnap et al., 2001), we present a consequentialist system of multi-agent deontic logic, which is a generalization of John Horty's utilitarian deontic logic (Horty, 2001). By means of our consequentialist deontic logic, we investigate the logical interaction between (1) *alethic statements* having the form 'It is possible that ϕ' (abbreviated as $\Diamond\phi$), (2) *agentive statements* having the form 'Group $\mathcal{G}$ of agents sees to it that ϕ' (abbreviated as $[\mathcal{G}]\phi$), and (3) *deontic statements* having the form 'In the interest of group $\mathcal{F}$ of agents, group $\mathcal{G}$ of agents ought to see to it that ϕ' (abbreviated as $\odot^{\mathcal{F}}_{\mathcal{G}}\phi$).

This language enables us to give a formal definition of moral conflicts. Two groups of agents, $\mathcal{G}_1$ and $\mathcal{G}_2$, face a moral conflict if and only if there are formulas ϕ and ψ, such that both $\odot^{\mathcal{F}_1}_{\mathcal{G}_1}\phi$ and $\odot^{\mathcal{F}_2}_{\mathcal{G}_2}\psi$ are true, whereas $\Diamond(\phi \wedge \psi)$ is false. Note that if ϕ and ψ cannot both be true, the truth of ψ implies the falsity of ϕ and, hence, having a moral obligation to see to it that ψ implies having a moral obligation to see to it that $\neg\phi$. Therefore, any moral conflict between $\mathcal{G}_1$ and $\mathcal{G}_2$ implies a *basic moral conflict* between those groups: two groups of agents, $\mathcal{G}_1$ and $\mathcal{G}_2$, face a basic moral conflict if and only if there is a formula ϕ, such that both $\odot^{\mathcal{F}_1}_{\mathcal{G}_1}\phi$ and $\odot^{\mathcal{F}_2}_{\mathcal{G}_2}\neg\phi$ are true. (Obviously, non-existence of basic moral conflicts implies non-existence of moral conflicts *tout court*.[3]) Thus, formulas of the form

$$\odot^{\mathcal{F}_1}_{\mathcal{G}_1}\phi \wedge \odot^{\mathcal{F}_2}_{\mathcal{G}_2}\neg\phi$$

will be central to our current investigation.

As a useful leg up to our formal results concerning moral conflicts between groups of agents, we illustrate our deontic logic with a formal analysis of the Prisoner's Dilemma. We shall show that the game theoretic dilemma can be completely translated into our model theory and that our semantics rules that each prisoner ought to confess in his own interest, but also that each prisoner ought not to confess in the collective interest of both prisoners: hence, each prisoner faces a basic moral conflict. Roughly, we shall show that a single group of agents may face a basic moral conflict if and only if the pertinent obligations stem from different moral codes. Or equivalently, a single group of agents cannot face a basic moral conflict if and only if the pertinent obligations stem from a single moral code. How about two groups? Might two groups of agents face a moral conflict of which the pertinent obligations stem from a single moral code?

Standard deontic logic entirely rules out such moral conflicts. The non-existence of moral conflicts within a single moral code is ensured by the joint validity of the principles $\mathcal{O}\phi \rightarrow \Diamond\phi$ ('ought' implies 'can') and $(\mathcal{O}\phi \wedge \mathcal{O}\psi) \rightarrow \mathcal{O}(\phi \wedge \psi)$ (deontic agglomeration).[4] In our logic, the translates of these principles are the following:

$$\odot^{\mathcal{F}}_{\mathcal{G}}\phi \rightarrow \Diamond[\mathcal{G}]\phi \qquad (\text{'}\textit{ought}\text{' } \textit{implies } \text{'}\textit{can}\text{'})$$
$$(\odot^{\mathcal{F}}_{\mathcal{G}}\phi \wedge \odot^{\mathcal{F}}_{\mathcal{G}}\psi) \rightarrow \odot^{\mathcal{F}}_{\mathcal{G}}(\phi \wedge \psi) \quad (\textit{deontic agglomeration})$$

Both principles are valid according to our semantics. Nevertheless, their validity does not preclude the possibility of moral conflicts within a single moral code.[5] Roughly, we shall show that two groups of agents may face a basic moral conflict of which the pertinent obligations stem from a single moral code if and only if the two groups have at least one common member and neither of the two groups is a subgroup of the other.

The set-up of the paper is as follows. In the next section, we fix a formal language for our multi-agent deontic logic and provide a consequentialist semantics for it. We illustrate our semantics in Section 2.5.2 with an instance of the Prisoner's Dilemma. In Section 3.1 we give a formal characterization of the possibility that a single group of agents faces a basic moral conflict. In Section 3.2 we formally characterize the possibility that two groups of agents face a basic moral conflict stemming from a single moral code. These two formal characterizations are combined in Section 3.3 to obtain a formal characterization of the possibility that two groups of agents face a moral conflict. We conclude the paper with a short discussion of our results.

2. Language and Semantics

2.1. *Language*

Throughout the paper, we use a propositional modal language $\mathfrak{L}$ built from a countable set $\mathfrak{P} = \{p_1, p_2, \ldots\}$ of atomic propositions and a finite set $A = \{a_1, \ldots, a_n\}$ of individual agents. $\mathfrak{L}$ is the smallest set (in terms of set-theoretical inclusion) satisfying the conditions (i) through (v):[6]

(i) $\mathfrak{P} \subseteq \mathfrak{L}$
(ii) If $\phi \in \mathfrak{L}$ and $\psi \in \mathfrak{L}$, then $(\phi \wedge \psi) \in \mathfrak{L}$ and $(\phi \rightarrow \psi) \in \mathfrak{L}$
(iii) If $\phi \in \mathfrak{L}$, then $\neg\phi \in \mathfrak{L}$ and $\Diamond\phi \in \mathfrak{L}$
(iv) If $\phi \in \mathfrak{L}$ and $\mathcal{G} \subseteq A$, then $[\mathcal{G}]\phi \in \mathfrak{L}$
(v) If $\phi \in \mathfrak{L}$ and $\mathcal{F} \subseteq A$ and $\mathcal{G} \subseteq A$, then $\odot^{\mathcal{F}}_{\mathcal{G}}\phi \in \mathfrak{L}$.

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

This formal language enables us to formalize a rather broad class of moral obligations, because there is no necessary connection between the group $\mathcal{G}$ who has a certain obligation and the group $\mathcal{F}$ who defines the moral code from which the obligation stems. Two examples: a utilitarian obligation like 'In the interest of all agents, group $\mathcal{G}$ of agents ought to see to it that ϕ' can now be formalized as $\odot^{A}_{\mathcal{G}}\phi$, since A denotes the group of all agents. An egoistic obligation like 'In his own interest, the agent a ought to see to it that ϕ' can be rendered as $\odot^{a}_{a}\phi$.

We shall evaluate formulas of the language $\mathfrak{L}$ in consequentialist models.

2.2. *Consequentialist Models*

DEFINITION 1. A *consequentialist model* $\mathfrak{M}$ is an ordered pair $\langle \mathfrak{S}, \mathfrak{I} \rangle$, where $\mathfrak{S}$ is a choice structure and $\mathfrak{I}$ an interpretation.

Choice structures are defined over a non-empty set W of possible worlds and a finite set A of agents. Interpretations assign agent-relative utilities to possible worlds and world-relative truth-values to atomic propositions.

For clarity's sake we do not take up Horty's branching-time models to evaluate deontic formulas. Instead, we here adopt a rather standard possible worlds approach.[7] Hence, our models represent possibilities, group actions, and group obligations at a single moment in time.

2.3. *Choice Structures*

DEFINITION 2. A *choice structure* $\mathfrak{S}$ is a triple $\langle W, A, Choice \rangle$, where W is a non-empty set of possible worlds, A a finite set of agents, and $Choice$ a choice function.

2.3.1. *Choice Functions*

Given a non-empty set W of possible worlds and a finite set A of individual agents, we define choice sets of individual agents by a choice function from individual agents to sets of sets of possible worlds, *i.e.*, $Choice : A \mapsto \wp(\wp(W))$, meeting the conditions that (1) for each individual agent a in A it holds that $Choice(a)$ is a partition of W, and (2) for each selection function s assigning to each individual agent a in A a set of possible worlds $s(a)$ such that $s(a) \in Choice(a)$ it holds that $\bigcap_{a \in A} s(a)$ is non-empty.[8]

For example, let $W = \{w_1, w_2, w_3, w_4\}$ and $A = \{a, b\}$. Define $Choice(a) = \{\{w_1, w_2\}, \{w_3, w_4\}\}$ and $Choice(b) = \{\{w_1, w_3\}, \{w_2, w_4\}\}$. Then $Choice$ is a choice function, since it meets the two conditions: (1) both $Choice(a)$ and $Choice(b)$ are partitions of W, and (2) for each of the four selection functions s assigning to agent a an option $s(a)$ in $Choice(a)$ and to agent b an option $s(b)$ in $Choice(b)$ it holds that $s(a) \cap s(b) \neq \emptyset$.[9]

Choice sets for groups of agents are given by a choice function from sets of individual agents to sets of sets of possible worlds, *i.e.*, $Choice : \wp(A) \mapsto \wp(\wp(W))$. Just like Horty does, we define group choices in terms of individual choices, thereby giving an affirmative answer to von Wright's question "whether acts attributed to collective agents could not be regarded as 'logical constructions' of acts of some individual agents" (von Wright, 1963, pp. 38–39). To be precise, given a choice function $Choice$ from individual agents to sets of sets of possible worlds and given the corresponding set $Select$ of selection functions s assigning to each individual agent a in A an option $s(a)$ in $Choice(a)$, we define

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

if $\mathcal{G}$ is non-empty. Otherwise, $Choice(\mathcal{G}) = \{W\}$. Thus, in our present example, $Choice(\{a, b\}) = \{\{w_1\}, \{w_2\}, \{w_3\}, \{w_4\}\}$.[10]

The proof of our theorem on moral conflicts within a single moral code (Section 3.2) makes use of some facts concerning choice functions:

LEMMA 1. Let $\mathfrak{S}(= \langle W, A, Choice \rangle)$ be a choice structure. Let $\mathcal{G}_1, \mathcal{G}_2 \subseteq A$. Then

(i) If $Choice(\mathcal{G}_1) = Choice(\mathcal{G}_2)$, then $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$
(ii) If $Choice(\mathcal{G}_1) = \{W\}$, then $Choice(\mathcal{G}_2) = Choice(\mathcal{G}_1 \cup \mathcal{G}_2)$
(iii) If $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$, then $Choice(\mathcal{G}_1) = Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$
(iv) If $Choice(\mathcal{G}_1) = Choice(\mathcal{G}_2)$, then $Choice(A - \mathcal{G}_1) = Choice(A - \mathcal{G}_2)$.

Proof.

(i) Assume $Choice(\mathcal{G}_1) = Choice(\mathcal{G}_2)$. Suppose $Choice(\mathcal{G}_1 - \mathcal{G}_2) \neq \{W\}$. Then there are $K, K' \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ such that $K \cap K' = \emptyset$. Take an $M \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$. It holds that $K \cap M, K' \cap M \in Choice(\mathcal{G}_1)$ and $(K \cap M) \cap (K' \cap M) = \emptyset$. By our assumption, $K \cap M, K' \cap M \in Choice(\mathcal{G}_2)$. Then there are $L, L' \in Choice(\mathcal{G}_2 - \mathcal{G}_1)$ such that $K \cap M = L \cap M$ and $K' \cap M = L' \cap M$. Since $K \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ and $L' \in Choice(\mathcal{G}_2 - \mathcal{G}_1)$ and $M \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$, it must be that $K \cap L' \cap M \in Choice(\mathcal{G}_1 \cup \mathcal{G}_2)$ and $K \cap L' \cap M \neq \emptyset$. But $(K \cap M) \cap (L' \cap M) = \emptyset$. Contradiction. Therefore, for all $K, K' \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ it holds that $K \cap K' \neq \emptyset$. Given the conditions on *Choice*, it must be that $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$.

(ii) Assume $Choice(\mathcal{G}_1) = \{W\}$. Then it must be that $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$. ($\Rightarrow$) Suppose $K \in Choice(\mathcal{G}_2)$. Since $W \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ and $K \in Choice(\mathcal{G}_2)$, it must be that $W \cap K \in Choice(\mathcal{G}_1 \cup \mathcal{G}_2)$. Hence, $K \in Choice(\mathcal{G}_1 \cup \mathcal{G}_2)$. ($\Leftarrow$) Suppose $K \in Choice(\mathcal{G}_1 \cup \mathcal{G}_2)$. Then there are $L \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ and $M \in Choice(\mathcal{G}_2)$ such that $K = L \cap M$. Then $L = W$ and $K = M$. Hence, $K \in Choice(\mathcal{G}_2)$.

(iii) Assume $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$. By (ii), it must be that $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) = Choice((\mathcal{G}_1 - \mathcal{G}_2) \cup (\mathcal{G}_1 \cap \mathcal{G}_2))$. Therefore, $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) = Choice(\mathcal{G}_1)$.

(iv) Assume $Choice(\mathcal{G}_1) = Choice(\mathcal{G}_2)$. By (i), it must be that $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\} = Choice(\mathcal{G}_2 - \mathcal{G}_1)$. Then $Choice(A - \mathcal{G}_1)$

$$
\begin{aligned}
&= Choice((A - (\mathcal{G}_1 \cup \mathcal{G}_2)) \cup (\mathcal{G}_2 - \mathcal{G}_1)) && \text{by elementary set theory}\\
&= Choice(A - (\mathcal{G}_1 \cup \mathcal{G}_2)) && \text{by (ii)}\\
&= Choice((A - (\mathcal{G}_1 \cup \mathcal{G}_2)) \cup (\mathcal{G}_1 - \mathcal{G}_2)) && \text{by (ii)}\\
&= Choice(A - \mathcal{G}_2) && \text{by elementary set theory.}
\end{aligned}
$$

Therefore, $Choice(A - \mathcal{G}_1) = Choice(A - \mathcal{G}_2)$. □

2.3.2. *$\mathcal{G}$-choice Equivalence of Worlds*

In choosing an option K from $Choice(\mathcal{G})$, the group $\mathcal{G}$ of agents restricts the total set of possible worlds to the possible worlds in the set K. A formula of the form $[\mathcal{G}]\phi$, informally interpreted as 'Group $\mathcal{G}$ of agents sees to it that ϕ', is true in a world w if and only if ϕ is true in all possible worlds that are elements of the option of $\mathcal{G}$ that contains w. Or, equivalently, if and

only if for all possible worlds w' that are $\mathcal{G}$-*choice equivalent* to world w it holds that ϕ is true in world w'. $\mathcal{G}$-choice equivalence is defined as follows:

DEFINITION 3. ($\mathcal{G}$-choice Equivalence) Let $\mathfrak{S}(= \langle W, A, \mathit{Choice}\rangle)$ be a choice structure. Let $\mathcal{G} \subseteq A$. Let $w, w' \in W$. Then $w \sim_{\mathcal{G}} w'$ (w *and* w' *are* $\mathcal{G}$-*choice equivalent*) is defined to be:

$$w \sim_{\mathcal{G}} w' \text{ iff for all } K \in \mathit{Choice}(\mathcal{G}) \text{ with } w \in K \text{ it holds that } w' \in K.$$

Thus, in our previous example, it holds that $w_1 \sim_a w_2$ and that $w_1 \not\sim_b w_2$.

After this discussion of choice structures, we must define their interpretations.

2.4. *Interpretations*

DEFINITION 4. An *interpretation* $\mathfrak{I}$ is an ordered pair $\langle \mathit{Utility}, V\rangle$, where *Utility* is a utility function and V a valuation function.

2.4.1. *Utility Functions*

The relation between individual utilities and group utilities is a subject bristling with pitfalls. Without taking a definite stance on this issue, we adopt, for the sake of the argument, John Harsanyi's proposal and conceive of group utility as the arithmetical mean of the individual utilities of the agents involved.[11] Obviously, the arithmetical mean of individual utilities can only be given a clear meaning if we can make interagential comparisons of individual utilities. To make possible such comparisons, we here start from the assumption that all individual utilities are normalized and that they are given by a utility function from ordered pairs consisting of an individual agent and a possible world to the real numbers between, say, -5 and 5, *i.e.*, $\mathit{Utility} : A \times W \mapsto [-5, 5]$. Thus, if an individual agent a assigns to a possible world w a utility of 4, we write $\mathit{Utility}(a, w) = 4$.

Group utilities are given by a utility function from ordered pairs consisting of a set of individual agents and a possible world to real numbers between -5 and 5, *i.e.*, $\mathit{Utility} : \wp(A) \times W \mapsto [-5, 5]$. We define group utilities in terms of individual utilities, assuming that (1) in assessing the group utility of a given possible world, the individual utilities that are assigned to that world by the individual agents in the group are to be weighed equally, and (2) group utilities of groups of different sizes are to be comparable. Hence, the group utility a group $\mathcal{F}$ of agents assigns to

a possible world w is defined as the arithmetical mean of the individual utilities the individual agents in $\mathcal{F}$ assign to w:

$$Utility(\mathcal{F}, w) = \frac{1}{|\mathcal{F}|} \sum_{a \in \mathcal{F}} Utility(a, w),$$

if $\mathcal{F}$ is non-empty. Otherwise, $Utility(\mathcal{F}, w) = 0$. Thus, if $Utility(a, w) = 4$ and $Utility(b, w) = 0$, then $Utility(\{a, b\}, w) = 2$.

2.4.2. *Valuation Functions*

Valuations are given by a valuation function from ordered pairs consisting of an atomic proposition and a possible world to the truth-values TRUE and FALSE, *i.e.*, $V : \mathfrak{P} \times W \mapsto \{\text{TRUE}, \text{FALSE}\}$, where $\text{TRUE} \neq \text{FALSE}$. Thus, if an atomic proposition p is true in a possible world w, we write $V(p, w) = \text{TRUE}$.

2.4.3. *$\mathcal{F}$-dominance between $\mathcal{G}$'s Options*

Roughly, a formula of the form $\odot^{\mathcal{F}}_{\mathcal{G}}\phi$, informally interpreted as 'In the interest of group $\mathcal{F}$ of agents, group $\mathcal{G}$ of agents ought to see to it that ϕ', is true in a world w if and only if for all options K in $Choice(\mathcal{G})$ that do not ensure ϕ there is a strictly $\mathcal{F}$-better option K' in $Choice(\mathcal{G})$ such that (1) option K' ensures ϕ, and (2) all options K'' that are at least as $\mathcal{F}$-good as K' also ensure ϕ. Following (Horty, 1996) and (Horty, 2001), we interpret "$\mathcal{F}$-betterness" decision-theoretically and define it as $\mathcal{F}$-dominance. Our relation of $\mathcal{F}$-dominance is a generalization of Horty's dominance relation.[12]

When a group $\mathcal{G}$ performs a collective action by choosing an option K from $Choice(\mathcal{G})$, it constrains the set W of possible worlds to that set K of possible worlds. It may be, however, that the agents who are not members of $\mathcal{G}$ (and who therefore are members of the group $A - \mathcal{G}$) perform a collective action by choosing an option S from $Choice(A - \mathcal{G})$, thereby constraining the set K to the non-empty set of possible worlds $K \cap S$. (Note that the condition of agent independence ensures that $K \cap S$ is non-empty.) Hence, $\mathcal{G}$ usually will not be able to fully determine the outcome of its collective actions, since the final outcome also depends on the actions of agents in $A - \mathcal{G}$. Nevertheless, we can define an $\mathcal{F}$-dominance relation, denoted by $\succeq^{\mathcal{F}}_{\mathcal{G}}$, over $\mathcal{G}$'s options. If K and K' both are in $Choice(\mathcal{G})$, then, intuitively, $K \succeq^{\mathcal{F}}_{\mathcal{G}} K'$ is true if and only if K promotes the utility of group $\mathcal{F}$ at least as well as K', regardless of the collective action of the agents in $A - \mathcal{G}$.

Hence, we insert an interest group $\mathcal{F}$ in Horty's Definitions 4.1 and 4.5 (Horty, 2001, p. 60 and p. 68) to define $\mathcal{F}$-dominance:

DEFINITION 5. ($\mathcal{F}$-dominance) Let $\mathfrak{M}(= \langle \mathfrak{S}, \mathfrak{I} \rangle)$ be a consequentialist model. Let $\mathcal{F}, \mathcal{G} \subseteq A$ and let $K, K' \in Choice(\mathcal{G})$. Then $K \succeq^{\mathcal{F}}_{\mathcal{G}} K'$ (K *weakly $\mathcal{F}$-dominates K' for $\mathcal{G}$*) is defined to be:

$$K \succeq^{\mathcal{F}}_{\mathcal{G}} K' \text{ iff } \begin{array}{l} \text{for all } S \in Choice(A - \mathcal{G}) \text{ and for all } w, w' \in W \\ \text{it holds that if } w \in K \cap S \text{ and } w' \in K' \cap S, \text{ then} \\ Utility(\mathcal{F}, w) \geq Utility(\mathcal{F}, w'). \end{array}$$

As usual, $K \succ^{\mathcal{F}}_{\mathcal{G}} K'$ (K *strongly $\mathcal{F}$-dominates K' for $\mathcal{G}$*) if and only if $K \succeq^{\mathcal{F}}_{\mathcal{G}} K'$ and $K' \not\succeq^{\mathcal{F}}_{\mathcal{G}} K$.

The proof of our theorem on moral conflicts within a single moral code (Section 3.2) relies on the following lemmas about $\mathcal{F}$-dominance:

LEMMA 2. Let $\mathfrak{M}(= \langle \mathfrak{S}, \mathfrak{I} \rangle)$ be a consequentialist model. Let $\mathcal{F}, \mathcal{G}_1, \mathcal{G}_2 \subseteq A$. Then

$$\text{If } Choice(\mathcal{G}_1) = Choice(\mathcal{G}_2), \text{ then } K \succeq^{\mathcal{F}}_{\mathcal{G}_1} K' \text{ iff } K \succeq^{\mathcal{F}}_{\mathcal{G}_2} K'.$$

Proof. Immediate from Lemma 1(iv). □

LEMMA 3. Let $\mathfrak{M}(= \langle \mathfrak{S}, \mathfrak{I} \rangle)$ be a consequentialist model. Let $\mathcal{F}, \mathcal{G}_1, \mathcal{G}_2 \subseteq A$ such that $\mathcal{G}_1 \cap \mathcal{G}_2 = \emptyset$. Let $K, K' \in Choice(\mathcal{G}_1)$ and $L \in Choice(\mathcal{G}_2)$. Then

$$\text{If } K \succeq^{\mathcal{F}}_{\mathcal{G}_1} K', \text{ then } K \cap L \succeq^{\mathcal{F}}_{(\mathcal{G}_1 \cup \mathcal{G}_2)} K' \cap L.$$

Proof. Assume $K \succeq^{\mathcal{F}}_{\mathcal{G}_1} K'$. Take $L \in Choice(\mathcal{G}_2)$. Suppose that $S \in Choice(A - (\mathcal{G}_1 \cup \mathcal{G}_2))$ and $w \in K \cap L \cap S$ and $w' \in K' \cap L \cap S$. Since $\mathcal{G}_1 \cap \mathcal{G}_2 = \emptyset$, it holds that $L \cap S \in Choice(A - \mathcal{G}_1)$. Hence, by our assumption, it holds that $Utility(\mathcal{F}, w) \geq Utility(\mathcal{F}, w')$. □

2.5. *Semantics*

Having defined the notions of a consequentialist model, of $\mathcal{G}$-choice equivalence, and of $\mathcal{F}$-dominance, we now give the semantical rules stipulating the conditions under which a formula ϕ from $\mathfrak{L}$ is true in a world w in a consequentialist model $\mathfrak{M}$. Next, we list some standard deontic formulas that are true according to this semantics. As usual, $[\![\phi]\!]_{\mathfrak{M}}$ refers to the set of possible worlds in $\mathfrak{M}$ that validate ϕ.

2.5.1. *Semantical Rules and Tautologies*

DEFINITION 6. (Semantical Rules) Let $\mathfrak{M}(=\langle \mathfrak{S}, \mathfrak{I}\rangle)$ be a consequentialist model. Let $w \in W$ and let $\phi, \psi \in \mathfrak{L}$. Then

(i)	$\mathfrak{M}, w \models p$	iff	$V(p, w) = \text{TRUE}, \quad \textit{if } p \in \mathfrak{P}$
(ii)	$\mathfrak{M}, w \models \neg\phi$	iff	$\mathfrak{M}, w \not\models \phi$
(iii)	$\mathfrak{M}, w \models \phi \wedge \psi$	iff	$\mathfrak{M}, w \models \phi$ and $\mathfrak{M}, w \models \psi$
(iv)	$\mathfrak{M}, w \models \phi \rightarrow \psi$	iff	$\mathfrak{M}, w \not\models \phi$ and/or $\mathfrak{M}, w \models \psi$
(v)	$\mathfrak{M}, w \models \Diamond\phi$	iff	there is a w' in W such that $\mathfrak{M}, w' \models \phi$
(vi)	$\mathfrak{M}, w \models [\mathcal{G}]\phi$	iff	for all w' in W with $w \sim_{\mathcal{G}} w'$ it holds that $\mathfrak{M}, w' \models \phi$
(vii)	$\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}}\phi$	iff	for all K in $\textit{Choice}(\mathcal{G})$ with $K \not\subseteq [\![\phi]\!]_{\mathfrak{M}}$ there is a K' in $\textit{Choice}(\mathcal{G})$ with $K' \subseteq [\![\phi]\!]_{\mathfrak{M}}$ such that (1) $K' \succ^{\mathcal{F}}_{\mathcal{G}} K$, and (2) for all K'' in $\textit{Choice}(\mathcal{G})$ with $K'' \succeq^{\mathcal{F}}_{\mathcal{G}} K'$ it holds that $K'' \subseteq [\![\phi]\!]_{\mathfrak{M}}$.

For the purpose of practicality, we introduce the following notational conventions: Given a model $\mathfrak{M}$, we write $\mathfrak{M} \models \phi$, if for all worlds w in W it holds that $\mathfrak{M}, w \models \phi$. We write $\models \phi$, if for all models $\mathfrak{M}$ it holds that $\mathfrak{M} \models \phi$. Given a choice structure $\mathfrak{S}$, we write $\mathfrak{S} \models \phi$, if for all interpretations $\mathfrak{I}$ of $\mathfrak{S}$ it holds that $\langle \mathfrak{S}, \mathfrak{I}\rangle \models \phi$.

LEMMA 4. Let $\phi, \psi \in \mathfrak{L}$. Then

(i) $\models \odot^{\mathcal{F}}_{\mathcal{G}}\phi \rightarrow \Diamond[\mathcal{G}]\phi$ ('ought' implies 'can')

(ii) If $\models \phi \leftrightarrow \psi$, then $\models \odot^{\mathcal{F}}_{\mathcal{G}}\phi \leftrightarrow \odot^{\mathcal{F}}_{\mathcal{G}}\psi$

(iii) If $\models \phi$, then $\models \odot^{\mathcal{F}}_{\mathcal{G}}\phi$

(iv) $\models \odot^{\mathcal{F}}_{\mathcal{G}}(\phi \wedge \psi) \rightarrow (\odot^{\mathcal{F}}_{\mathcal{G}}\phi \wedge \odot^{\mathcal{F}}_{\mathcal{G}}\psi)$

(v) $\models (\odot^{\mathcal{F}}_{\mathcal{G}}\phi \wedge \odot^{\mathcal{F}}_{\mathcal{G}}\psi) \rightarrow \odot^{\mathcal{F}}_{\mathcal{G}}(\phi \wedge \psi)$ (deontic agglomeration)

Proof. The proofs of (i) through (iv) are straightforward. The proof of (v) is analogous to the one in (Horty, 2001, pp. 166–167). □

2.5.2. *An Example: the Prisoner's Dilemma*

We now fulfill the promise made in the introduction and illustrate our consequentialist semantics for multi-agent deontic logic with an analysis of

the Prisoner's Dilemma. The Prisoner's Dilemma is a two-player strategic game, represented by the following payoff matrix (Osborne and Rubinstein, 1994, p. 17):

	Don't confess	*Confess*
Don't confess	3, 3	0, 4
Confess	4, 0	1, 1

This payoff matrix can be translated into a consequentialist model $\mathfrak{M}(=\langle \mathfrak{S}, \mathfrak{I}\rangle)$, where we interpret game theoretic utilities as normalized individual utilities. The choice structure $\mathfrak{S}$ is given by $W=\{w_1, w_2, w_3, w_4\}$, $A=\{a, b\}$, $Choice(a)=\{\{w_1, w_2\}, \{w_3, w_4\}\}$, and $Choice(b) = \{\{w_1, w_3\}, \{w_2, w_4\}\}$. The interpretation $\mathfrak{I}$ is given by

$$\begin{array}{ll} Utility(a, w_1) = 3 & Utility(b, w_1) = 3 \\ Utility(a, w_2) = 0 & Utility(b, w_2) = 4 \\ Utility(a, w_3) = 4 & Utility(b, w_3) = 0 \\ Utility(a, w_4) = 1 & Utility(b, w_4) = 1, \end{array}$$

and $V(p, w) = \text{TRUE}$ if and only if $w \in \{w_3, w_4\}$, and $V(q, w) = \text{TRUE}$ if and only if $w \in \{w_2, w_4\}$. We read p as 'Agent a confesses' and q as 'Agent b confesses'.

In this model $\mathfrak{M}$, each individual agent faces a basic moral conflict. Both statements 'In the interest of agent a, agent a ought to see to it that p' and 'In the interest of the group of agents consisting of a and b, agent a ought to see to it that $\neg p$' are true in $\mathfrak{M}$. The situation for b is analogous. Hence, $\mathfrak{M}$ gives rise to two single-agent basic moral conflicts:

$$\mathfrak{M} \models \odot_a^a p \wedge \odot_a^{a,b} \neg p \text{ and } \mathfrak{M} \models \odot_b^b q \wedge \odot_b^{a,b} \neg q.$$

Let us see why this follows from our semantics. We show that for all $w \in W$ it holds that (i) $\mathfrak{M}, w \models \odot_a^a p$ and (ii) $\mathfrak{M}, w \models \odot_a^{a,b} \neg p$. Let $w \in W$.

Ad (i). $\mathfrak{M}, w \models \odot_a^a p$ if and only if for each $K \in Choice(a)$ with $K \not\subseteq [\![p]\!]_{\mathfrak{M}}$ there is a $K' \in Choice(a)$ with $K' \subseteq [\![p]\!]_{\mathfrak{M}}$, such that $K' \succ_a^a K$, and for each $K'' \in Choice(a)$ with $K'' \succeq_a^a K'$ it holds that $K'' \subseteq [\![p]\!]_{\mathfrak{M}}$. By definition of $\mathfrak{M}$, it holds that $Choice(a) = \{\{w_1, w_2\}, \{w_3, w_4\}\}$. Moreover, it holds that $\{w_1, w_2\} \not\subseteq [\![p]\!]_{\mathfrak{M}}$ and $\{w_3, w_4\} \subseteq [\![p]\!]_{\mathfrak{M}}$. Hence, we only need to check $\{w_3, w_4\} \succ_a^a \{w_1, w_2\}$.

Notice that $\{w_3, w_4\} \succ_a^a \{w_1, w_2\}$ if and only if $\{w_3, w_4\} \succeq_a^a \{w_1, w_2\}$ and $\{w_1, w_2\} \not\succeq_a^a \{w_3, w_4\}$. The first conjunct holds if and only if for all

$S \in \mathit{Choice}(b)$ and for all $w, w' \in W$ it holds that if $w \in \{w_3, w_4\} \cap S$ and $w' \in \{w_1, w_2\} \cap S$, then $\mathit{Utility}(a, w) \geq \mathit{Utility}(a, w')$. By definition of $\mathfrak{M}$, it holds that $\mathit{Choice}(b) = \{\{w_1, w_3\}, \{w_2, w_4\}\}$ and $\mathit{Utility}(a, w_3) \geq \mathit{Utility}(a, w_1)$ and $\mathit{Utility}(a, w_4) \geq \mathit{Utility}(a, w_2)$. Hence, the first conjunct holds. The second conjunct holds if and only if there is an $S \in \mathit{Choice}(b)$ and there are $w, w' \in W$ such that $w \in \{w_1, w_2\} \cap S$ and $w' \in \{w_3, w_4\} \cap S$ and $\mathit{Utility}(a, w) < \mathit{Utility}(a, w')$. Any $S \in \mathit{Choice}(b)$ suffices. Hence, the second conjunct holds.

Ad (ii). $\mathfrak{M}, w \models \odot_a^{a,b} \neg p$ can be shown analogously. Note that $\mathit{Utility}(\{a, b\}, w_1) \geq \mathit{Utility}(\{a, b\}, w_3)$ and $\mathit{Utility}(\{a, b\}, w_2) \geq \mathit{Utility}(\{a, b\}, w_4)$.[13]

Two additional remarks are in order. First, both statements 'In the interest of the group of agents consisting of a and b, the group of agents consisting of a and b ought to see to it that $\neg p$' and 'In the interest of the group of agents consisting of a and b, the group of agents consisting of a and b ought to see to it that $\neg q$' are true in $\mathfrak{M}$. Hence, $\mathfrak{M}$ gives rise to two multi-agent basic moral conflicts:

$$\mathfrak{M} \models \odot_a^a p \wedge \odot_{a,b}^{a,b} \neg p \text{ and } \mathfrak{M} \models \odot_b^b q \wedge \odot_{a,b}^{a,b} \neg q.$$

Second, notice that the agent a cannot see to it that agent b confesses and that the agent b cannot see to it that agent a confesses. Accordingly, both statements 'In the interest of the group of agents consisting of a and b, agent a ought to see to it that $\neg p \wedge \neg q$' and 'In the interest of the group of agents consisting of a and b, agent b ought to see to it that $\neg p \wedge \neg q$' are *false* in $\mathfrak{M}$. Hence, it holds that

$$\mathfrak{M} \not\models \odot_a^{a,b}(\neg p \wedge \neg q) \text{ and } \mathfrak{M} \not\models \odot_b^{a,b}(\neg p \wedge \neg q).$$

In sum, our consequentialist semantics for multi-agent deontic logic provides a formal and fairly accurate account of some important senses of "moral obligation" in the Prisoner's Dilemma. Our logic does not, of course, *solve* the Prisoner's Dilemma, since it does not prescribe individual agents a and b to further the interest of the group $\{a, b\}$ rather than to advance their individual interest, nor the other way round.

3. Three Characterizations of Moral Conflicts

Let us now address the problem of basic moral conflicts from a metalogical viewpoint. We take up Johan van Benthem's notion of a modal formula characterizing a frame property and adapt it to the present situation.[14] By proving that certain deontic formulas characterize certain properties of

choice structures, we give necessary and sufficient conditions for (1) the possibility that a single group of agents faces a basic moral conflict, for (2) the possibility that two groups of agents face a basic moral conflict within a single moral code, and for (3) the possibility that two groups of agents face a basic moral conflict.

DEFINITION 7. (Characterization) Let $\mathfrak{C}$ be a class of choice structures and let $\phi \in \mathfrak{L}$. Then ϕ *characterizes* $\mathfrak{C}$, if for all choice structures $\mathfrak{S}$ it holds that $\mathfrak{S} \in \mathfrak{C}$ if and only if $\mathfrak{S} \models \phi$.

3.1. *Moral Conflicts of Type* $\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$

We show that a moral conflict of type $\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$ might occur in a choice structure $\mathfrak{S}$ if and only if there are groups $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}$ of agents in $\mathfrak{S}$ such that $\mathcal{F}_1$ is non-empty, $\mathcal{F}_2$ is non-empty, $\mathcal{F}_1$ and $\mathcal{F}_2$ are not identical, and $\mathcal{G}$ has at least two non-identical options for acting:

THEOREM 1. Let $\mathfrak{C}$ be the class of choice structures $\mathfrak{S}$ such that for all $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A$ it holds that $\mathcal{F}_1 = \emptyset$ or $\mathcal{F}_2 = \emptyset$ or $\mathcal{F}_1 = \mathcal{F}_2$ or $\mathit{Choice}(\mathcal{G}) = \{W\}$. Let $p \in \mathfrak{P}$. Then

$$\bigwedge_{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A} \neg \left(\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p \right)$$

characterizes $\mathfrak{C}$.

Proof. We show that (i) for all $\mathfrak{S} \in \mathfrak{C}$ it holds that $\mathfrak{S} \models \bigwedge_{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A} \neg \left(\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p \right)$ and (ii) for all $\mathfrak{S} \notin \mathfrak{C}$ it holds that $\mathfrak{S} \not\models \bigwedge_{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A} \neg \left(\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p \right)$.

Ad (i). Suppose $\mathfrak{S} \in \mathfrak{C}$. Suppose $\mathfrak{S} \not\models \bigwedge_{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A} \neg \left(\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p \right)$. Then there is an interpretation $\mathfrak{I}$ of $\mathfrak{S}$ such that the model $\mathfrak{M} = \langle \mathfrak{S}, \mathfrak{I} \rangle$ falsifies $\bigwedge_{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A} \neg \left(\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p \right)$. Then there are $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A$ and a w in W, such that $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} p$ and $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$. Since $\mathfrak{S} \in \mathfrak{C}$, it must be that $\mathcal{F}_1 = \emptyset$ or $\mathcal{F}_2 = \emptyset$ or $\mathcal{F}_1 = \mathcal{F}_2$ or $\mathit{Choice}(\mathcal{G}) = \{W\}$.

[Case 1]. Suppose $\mathcal{F}_1 = \emptyset$. Then, since $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\emptyset} p$, it must be that $\mathfrak{M}, w \models \Box p$ and, hence, $[\![\neg p]\!]_{\mathfrak{M}} = \emptyset$. By $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$ and Lemma 4(i), it must be that $\mathfrak{M}, w \models \Diamond[\mathcal{G}] \neg p$. Then there must be a (non-empty) K in $\mathit{Choice}(\mathcal{G})$ such that $K \subseteq [\![\neg p]\!]_{\mathfrak{M}}$. Contradiction.

[Case 2]. Suppose $\mathcal{F}_2 = \emptyset$. Analogous to Case 1.

[Case 3]. Suppose $\mathcal{F}_1 = \mathcal{F}_2$. Then $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_1} \neg p$. By Lemma 4(v), it must be that $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} (p \wedge \neg p)$. By Lemma 4(i), it

must be that $\mathfrak{M}, w \models \Diamond[\mathcal{G}](p \wedge \neg p)$. Then there must be a (non-empty) K in $Choice(\mathcal{G})$ such that $K \subseteq [\![p \wedge \neg p]\!]_{\mathfrak{M}}$. Contradiction.

[Case 4]. Suppose $Choice(\mathcal{G}) = \{W\}$. Then it must be that $W \subseteq [\![p]\!]_{\mathfrak{M}}$, since $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} p$, and it must be that $W \subseteq [\![\neg p]\!]_{\mathfrak{M}}$, since $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$. Note that W is non-empty. Contradiction.

Ad (ii). Suppose $\mathfrak{S} \notin \mathfrak{C}$. Then there must be $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A$ such that $\mathcal{F}_1 \neq \emptyset$ and $\mathcal{F}_2 \neq \emptyset$ and $\mathcal{F}_1 \neq \mathcal{F}_2$ and $Choice(\mathcal{G}) \neq \{W\}$. To prove that $\mathfrak{S} \not\models \bigwedge_{\mathcal{F}_1, \mathcal{F}_2, \mathcal{G} \subseteq A} \neg\big(\odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p\big)$, it suffices to construct a model $\mathfrak{M} = \langle \mathfrak{S}, \mathfrak{I} \rangle$ in which there is a w such that $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$. Without loss of generality, we may conclude from the four properties that there is an agent a with $a \in \mathcal{F}_1$, there is an agent b with $b \notin \mathcal{F}_1$ and $b \in \mathcal{F}_2$, and there are at least two non-identical options K_1 and K_2 in $Choice(\mathcal{G})$. We now define a suitable interpretation $\mathfrak{I} = \langle Utility, V \rangle$. First, $Utility$ is defined as follows:

$$Utility(a, w) = \begin{cases} 1, & \text{if } w \in K_1 \\ 0, & \text{otherwise}, \end{cases}$$

$$Utility(b, w) = \begin{cases} 2, & \text{if } w \in K_2 \\ 0, & \text{otherwise}, \end{cases}$$

and for all agents c in $A - \{a, b\}$ and for all worlds w in W, we fix $Utility(c, w) = 0$. Second, we stipulate that $V(p, w) = \text{TRUE}$ if and only if $w \in K_1$.

Let $\mathfrak{M} = \langle \mathfrak{S}, \mathfrak{I} \rangle$ and let $w \in W$. Now it is easy to show that (1) for all $K \in Choice(\mathcal{G})$ with $K \neq K_1$ it holds that $K_1 \succ_{\mathcal{G}}^{\mathcal{F}_1} K$ and (2) for all $K \in Choice(\mathcal{G})$ with $K \neq K_2$ it holds that $K_2 \succ_{\mathcal{G}}^{\mathcal{F}_2} K$. Hence, $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} p$ and $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$. Therefore, $\mathfrak{M}, w \models \odot_{\mathcal{G}}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}}^{\mathcal{F}_2} \neg p$. □

3.2. *Moral Conflicts of Type* $\odot_{\mathcal{G}_1}^{\mathcal{F}} p \wedge \odot_{\mathcal{G}_2}^{\mathcal{F}} \neg p$

Some authors have claimed that axiological approaches to moral obligations leave no room for moral conflicts.[15] A common argument for this claim runs as follows: "For suppose that A and B are incompatible. Then if it ought to be the case that A, higher values attach to some outcomes satisfying A than to any satisfying *not* A. But, because of the assumed incompatibility, all outcomes that satisfy B satisfy *not* A. Hence it is better to opt for A than for B. So, whenever A and B are mutually incompatible, it cannot be that both ought to be the case" (van Fraassen, 1973, p. 8). Things are different in our consequentialist multi-agent deontic logic.

We show that a moral conflict of type $\odot_{\mathcal{G}_1}^{\mathcal{F}} p \wedge \odot_{\mathcal{G}_2}^{\mathcal{F}} \neg p$ might occur in a choice structure $\mathfrak{S}$ if and only if there are groups $\mathcal{F}, \mathcal{G}_1, \mathcal{G}_2$ of agents in $\mathfrak{S}$

such that $\mathcal{F}$ is non-empty, $\mathcal{G}_1 - \mathcal{G}_2$ has at least two non-identical options for acting, $\mathcal{G}_2 - \mathcal{G}_1$ has at least two non-identical options for acting, and $\mathcal{G}_1 \cap \mathcal{G}_2$ has at least two non-identical options for acting.[16] Since there is only one interest group involved, the theorem also applies to Horty's original system. Note also that the proof is independent from our definition of group utility as the mean of the individual utilities concerned.

THEOREM 2. Let $\mathfrak{C}'$ be the class of choice structures $\mathfrak{S}$ such that for all $\mathcal{F}, \mathcal{G}_1, \mathcal{G}_2 \subseteq A$ it holds that $\mathcal{F} = \emptyset$ or $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$ or $Choice(\mathcal{G}_2 - \mathcal{G}_1) = \{W\}$ or $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) = \{W\}$. Let $p \in \mathfrak{P}$. Then

$$\bigwedge_{\mathcal{F},\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg \left(\odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p\right)$$

characterizes $\mathfrak{C}'$.

Proof. We show that (i) for all $\mathfrak{S} \in \mathfrak{C}'$ it holds that $\mathfrak{S} \models \bigwedge_{\mathcal{F},\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg \left(\odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p\right)$ and (ii) for all $\mathfrak{S} \notin \mathfrak{C}'$ it holds that $\mathfrak{S} \not\models \bigwedge_{\mathcal{F},\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg \left(\odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p\right)$.

Ad (i). Suppose $\mathfrak{S} \in \mathfrak{C}'$. Suppose $\mathfrak{S} \not\models \bigwedge_{\mathcal{F},\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg \left(\odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p\right)$. Then there is an interpretation $\mathfrak{I}$ of $\mathfrak{S}$ such that the model $\mathfrak{M} = \langle \mathfrak{S}, \mathfrak{I} \rangle$ falsifies $\bigwedge_{\mathcal{F},\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg \left(\odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p\right)$. Then there are $\mathcal{F}, \mathcal{G}_1, \mathcal{G}_2 \subseteq A$ and a w in W, such that $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_1} p$ and $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p$. Since $\mathfrak{S} \in \mathfrak{C}'$, it must be that $\mathcal{F} = \emptyset$ or $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$ or $Choice(\mathcal{G}_2 - \mathcal{G}_1) = \{W\}$ or $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) = \{W\}$.

[Case 1]. Suppose $\mathcal{F} = \emptyset$. Then, since $\mathfrak{M}, w \models \odot^{\emptyset}_{\mathcal{G}_1} p$ and $\mathfrak{M}, w \models \odot^{\emptyset}_{\mathcal{G}_2} \neg p$, it must be that $\mathfrak{M}, w \models \Box p$ and $\mathfrak{M}, w \models \Box \neg p$. Contradiction.

[Case 2]. Suppose $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$. Then, by Lemma 1(iii), it must be that $Choice(\mathcal{G}_1) = Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$. Suppose for all M with $M \in Choice(\mathcal{G}_2)$ it holds that $M \subseteq [\![\neg p]\!]_{\mathfrak{M}}$. Then $W \subseteq [\![\neg p]\!]_{\mathfrak{M}}$. Then $\mathfrak{M}, w \not\models \odot^{\mathcal{F}}_{\mathcal{G}_1} p$. Contradiction. Hence, there is an $M \in Choice(\mathcal{G}_2)$ with $M \not\subseteq [\![\neg p]\!]_{\mathfrak{M}}$.

Hence, since $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p$, there is a (non-empty) $M' \in Choice(\mathcal{G}_2)$ with $M' \subseteq [\![\neg p]\!]_{\mathfrak{M}}$, such that $M' \succ^{\mathcal{F}}_{\mathcal{G}_2} M$ and for all $M'' \in Choice(\mathcal{G}_2)$ with $M'' \succeq^{\mathcal{F}}_{\mathcal{G}_2} M'$ it holds that $M'' \subseteq [\![\neg p]\!]_{\mathfrak{M}}$. Note that $M' = K' \cap L'$ with $K' \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$ and $L' \in Choice(\mathcal{G}_2 - \mathcal{G}_1)$. Hence, $K' \in Choice(\mathcal{G}_1)$. It must be that $K' \not\subseteq [\![p]\!]_{\mathfrak{M}}$. Hence, since $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_1} p$, there is a $K'' \in Choice(\mathcal{G}_1)$ with $K'' \subseteq [\![p]\!]_{\mathfrak{M}}$ and $K'' \succ^{\mathcal{F}}_{\mathcal{G}_1} K'$. Hence, $K'' \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$ and, by Lemma 2, it must be that $K'' \succ^{\mathcal{F}}_{\mathcal{G}_1 \cap \mathcal{G}_2} K'$. Obviously, $K'' \cap L' \in Choice(\mathcal{G}_2)$ and $K'' \cap L' \neq \emptyset$. By Lemma 3, it must be that $K'' \cap L' \succeq^{\mathcal{F}}_{\mathcal{G}_2}$

$K' \cap L'$. Note that $K'' \cap L' \subseteq [\![p]\!]_{\mathfrak{M}}$. Finally, by substituting $K'' \cap L'$ for M'', we find that $K'' \cap L' \subseteq [\![\neg p]\!]_{\mathfrak{M}}$. Contradiction.

[Case 3]. Suppose $Choice(\mathcal{G}_2 - \mathcal{G}_1) = \{W\}$. Analogous to Case 2.

[Case 4]. Suppose $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) = \{W\}$. By $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_1} p$, there must be a K in $Choice(\mathcal{G}_1)$ such that $K \subseteq [\![p]\!]_{\mathfrak{M}}$. By $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p$, there must be an L in $Choice(\mathcal{G}_2)$ such that $L \subseteq [\![\neg p]\!]_{\mathfrak{M}}$. Note that there must be an $M \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ and an $S \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$ such that $K = M \cap S$. Moreover, there must be an $N \in Choice(\mathcal{G}_2 - \mathcal{G}_1)$ and an $S' \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$ such that $L = N \cap S'$. By our supposition, it holds that $S = S' = W$ and, hence, $K \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ and $L \in Choice(\mathcal{G}_2 - \mathcal{G}_1)$. By the condition of agent independence, it must be that $K \cap L \neq \emptyset$, since $(\mathcal{G}_1 - \mathcal{G}_2) \cap (\mathcal{G}_2 - \mathcal{G}_1) = \emptyset$. Contradiction.

Ad (ii). Suppose $\mathfrak{S} \notin \mathfrak{C}'$. Then there must be $\mathcal{F}, \mathcal{G}_1, \mathcal{G}_2 \subseteq A$ such that $\mathcal{F} \neq \emptyset$ and $Choice(\mathcal{G}_1 - \mathcal{G}_2) \neq \{W\}$ and $Choice(\mathcal{G}_2 - \mathcal{G}_1) \neq \{W\}$ and $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) \neq \{W\}$. To prove that $\mathfrak{S} \not\models \bigwedge_{\mathcal{F},\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg\big(\odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p\big)$, it suffices to construct a model $\mathfrak{M} = \langle \mathfrak{S}, \mathfrak{I} \rangle$ in which there is a w such that $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p$. We conclude from the four properties that there are at least two non-identical options K_1 and K_2 in $Choice(\mathcal{G}_1 - \mathcal{G}_2)$, at least two non-identical options L_1 and L_2 in $Choice(\mathcal{G}_2 - \mathcal{G}_1)$, and at least two non-identical options M_1 and M_2 in $Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$, and that there is an agent a in $\mathcal{F}$. Note that if $K \in Choice(\mathcal{G}_1 - \mathcal{G}_2)$ and $M \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$, then $K \cap M \neq \emptyset$ and $K \cap M \in Choice(\mathcal{G}_1)$. Note that if $L \in Choice(\mathcal{G}_2 - \mathcal{G}_1)$ and $M \in Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$, then $L \cap M \neq \emptyset$ and $L \cap M \in Choice(\mathcal{G}_2)$. We now define a suitable interpretation $\mathfrak{I} = \langle Utility, V \rangle$. First, *Utility* is defined as follows:

$$Utility(a, w) = \begin{cases} 1, \textit{ if } w \in K_1 \cap M_2 \textit{ or } w \in L_2 \cap M_1 \\ 0, \textit{ otherwise}, \end{cases}$$

and for all agents b in $A - \{a\}$ and for all worlds w in W, we fix $Utility(b, w) = 0$. Second, we stipulate $V(p, w) = \text{TRUE}$ if and only if $w \in K_1 \cap M_2$.

Let $\mathfrak{M} = \langle \mathfrak{S}, \mathfrak{I} \rangle$ and let $w \in W$. Now it is easy to show that (1) for all $R \in Choice(\mathcal{G}_1)$ with $R \neq (K_1 \cap M_2)$ it holds that $(K_1 \cap M_2) \succ^{\mathcal{F}}_{\mathcal{G}_1} R$ and (2) for all $S \in Choice(\mathcal{G}_2)$ with $S \neq (L_2 \cap M_1)$ it holds that $(L_2 \cap M_1) \succ^{\mathcal{F}}_{\mathcal{G}_2} S$. Hence, $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_1} p$ and $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p$. Therefore, $\mathfrak{M}, w \models \odot^{\mathcal{F}}_{\mathcal{G}_1} p \wedge \odot^{\mathcal{F}}_{\mathcal{G}_2} \neg p$. □

Let us take a closer look at the countermodel of part (ii) to interpret it properly. The group $\mathcal{G}_1 \cap \mathcal{G}_2$ of agents cannot make a principled choice from $Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$ to maximize the interest of group $\mathcal{F}$. If $\mathcal{G}_1 \cap \mathcal{G}_2$ is taken to belong to group $\mathcal{G}_1$, it has to choose option M_2 to maximize $\mathcal{F}$'s

interest. On the other hand, if $\mathcal{G}_1 \cap \mathcal{G}_2$ is seen as a subgroup of group $\mathcal{G}_2$, it must rather choose option M_1 to maximize $\mathcal{F}$'s interest. Obviously, $\mathcal{G}_1 \cap \mathcal{G}_2$ cannot choose both options. The group $\mathcal{G}_1 \cap \mathcal{G}_2$ of agents is wearing two hats here.

Earl Conee contends that in cases "where competing moral considerations have exactly the same force (..) [w]e have the familiar option of holding that (..) each act is permitted and none is absolutely obligatory" (Conee, 1982, p. 92). Unfortunately, this advice does not free $\mathcal{G}_1 \cap \mathcal{G}_2$ from its precarious predicament. In our example, only two of $\mathcal{G}_1 \cap \mathcal{G}_2$'s possible courses of action maximize $\mathcal{F}$-utility. Clearly, in order to maximize $\mathcal{F}$-utility, $\mathcal{G}_1 \cap \mathcal{G}_2$ *must* perform one of these two $\mathcal{F}$-maximizing options. Conee's recommendation just to waive the obligatory character of both $\mathcal{F}$-maximizing options simply does not wash. The existence of a moral conflict does not relieve $\mathcal{G}_1 \cap \mathcal{G}_2$ of the obligation to do the best it can.[17]

Hence, an additional decision procedure must be invoked to enforce a unique $\mathcal{F}$-maximizing course of action. Ruth B. Marcus suggests that "[i]n the unlikely cases where in fact two conflicting courses of action have the same utility, it is open to the act utilitarian to adopt a procedure for deciding, such as tossing a coin" (Marcus, 1980, p. 126). If $\mathcal{G}_1 \cap \mathcal{G}_2$ consists of a single individual agent and if this single agent has identified the $\mathcal{F}$-maximizing options, Marcus's suggestion would indeed save Buridan's ass from starvation. Does it also solve the decision problem if $\mathcal{G}_1 \cap \mathcal{G}_2$ consists of two (or more) individual agents?

In his simile of the soul as a chariot driven by reason and pulled by two horses embodying the spirited and the appetitive element, Plato notes that "the task of our charioteer is difficult and troublesome".[18] If $\mathcal{G}_1 \cap \mathcal{G}_2$ consists of two (or more) individual agents, the situation is even worse: a team of Buridan's asses has to co-ordinate its actions even without the whip of reason. Such a co-ordinated collective action entails at least the following four steps. First, the agents in $\mathcal{G}_1 \cap \mathcal{G}_2$ must collectively identify the $\mathcal{F}$-maximizing options in $Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$. Second, they have to agree upon which $\mathcal{F}$-maximizing option M in $Choice(\mathcal{G}_1 \cap \mathcal{G}_2)$ is going to be realized. (Here, they might indeed agree to flip a coin.) Third, each agent a in $\mathcal{G}_1 \cap \mathcal{G}_2$ must identify the unique option in $Choice(a)$ required for realizing M. Fourth, each agent a in $\mathcal{G}_1 \cap \mathcal{G}_2$ has to perform this unique course of action. Obviously, at each step a slip is easily made.

3.3. *Moral Conflicts of Type* $\odot_{\mathcal{G}_1}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}_2}^{\mathcal{F}_2} \neg p$

To conclude, we show that a moral conflict of type $\odot_{\mathcal{G}_1}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}_2}^{\mathcal{F}_2} \neg p$ might occur in a choice structure $\mathfrak{S}$ if and only if there are groups $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}_1, \mathcal{G}_2$ of agents in $\mathfrak{S}$ such that $\mathcal{F}_1$ is non-empty, $\mathcal{F}_2$ is non-empty, $\mathcal{G}_1 \cap \mathcal{G}_2$ has

at least two non-identical options for acting, and, finally, if $\mathcal{F}_1$ and $\mathcal{F}_2$ are identical, then both $\mathcal{G}_1 - \mathcal{G}_2$ and $\mathcal{G}_2 - \mathcal{G}_1$ have at least two non-identical options for acting:

THEOREM 3. Let $\mathfrak{C}''$ be the class of choice structures $\mathfrak{S}$ such that for all $\mathcal{F}_1, \mathcal{F}_2, \mathcal{G}_1, \mathcal{G}_2 \subseteq A$ it holds that $\mathcal{F}_1 = \emptyset$ or $\mathcal{F}_2 = \emptyset$ or $Choice(\mathcal{G}_1 \cap \mathcal{G}_2) = \{W\}$ or ($\mathcal{F}_1 = \mathcal{F}_2$ and $Choice(\mathcal{G}_1 - \mathcal{G}_2) = \{W\}$) or ($\mathcal{F}_1 = \mathcal{F}_2$ and $Choice(\mathcal{G}_2 - \mathcal{G}_1) = \{W\}$). Let $p \in \mathfrak{P}$. Then

$$\bigwedge_{\mathcal{F}_1,\mathcal{F}_2,\mathcal{G}_1,\mathcal{G}_2 \subseteq A} \neg \left(\odot_{\mathcal{G}_1}^{\mathcal{F}_1} p \wedge \odot_{\mathcal{G}_2}^{\mathcal{F}_2} \neg p \right)$$

characterizes $\mathfrak{C}''$.

Proof. Use the proofs of Theorem 1 and Theorem 2. □

4. Conclusion

In the present paper, we studied three different types of moral conflicts. On such conflicts we proved three theorems, the third of which is a generalization of the other two. Theorem 1 implies that if a single group of agents faces a moral conflict, this group must have obligations with respect to different non-empty interest groups. Hence, such a single-group moral conflict includes at least two agents. From Theorem 2 it follows that if two groups of agents face a moral conflict within a single moral code, then there must be at least three agents involved. Apparently, some deontological properties which are central to meta-ethics can only emerge in multi-agent settings, thereby establishing, or so it seems, multi-agent deontic logic as a proper field of study.[19]

Notes

[1] Notable exceptions are (Hamblin, 1972; Marcus, 1980) and (McConnell, 1988).
[2] In the same vein G.H. von Wright deems a logic of action a necessary requirement for deontic logic. See (von Wright, 1963, p. vii) and (von Wright, 1966, p. 134).
[3] Compare (Goble, 2005, pp. 462–463).
[4] The proof is straightforward and familiar: suppose that $\mathcal{O}\phi$ and $\mathcal{O}\neg\phi$ are simultaneously true. Then, by deontic agglomeration and modus ponens, it must be that $\mathcal{O}(\phi \wedge \neg\phi)$. Hence, by 'ought' implies 'can' and modus ponens, it must be that $\Diamond(\phi \wedge \neg\phi)$, which is absurd.

To clear the way for moral dilemmas in deontic logic, E.J. Lemmon proposes to reject the principle that 'ought' implies 'can' (see (Lemmon, 1962, p. 150, n. 8) and (Lemmon, 1965, pp. 47–50)), whereas Bas van Fraassen and Bernard Williams suggest to dismiss the principle of deontic agglomeration (see (van Fraassen, 1973, p. 15) and (Williams, 1973, pp. 181–182)). For a recent discussion of moral conflicts in standard deontic logic, see (Goble, 2005).

[5] Terrance McConnell writes: "[T]he existence of single-agent dilemmas forces us to give up either the principle that 'ought' implies 'can' or the [agglomeration] principle of deontic logic; but the reality of interpersonal moral conflicts forces no such concessions" (McConnell, 1988, p. 32).

[6] Logical rigour would demand that we draw a sharp distinction between (1) the *names* of (sets of) agents and (2) the *objects* that are being named, *i.e.*, the (sets of) agents themselves. We waive this distinction and thereby avoid unnecessary complications, as our present aims can be reached without it.

[7] Definitions of branching-time models for *stit* logics can be found in (Belnap et al., 2001) and (Horty, 2001).

[8] The latter requirement is the condition of *agent independence*. It ensures that there is a possible world in which each individual agent performs the action of his choice, regardless of the courses of action adopted by all other individual agents. Hence, *Choice* is defined so that at a single moment in time no individual agent can prevent any other individual agent from performing an action. See, for instance, (Belnap et al., 2001, pp. 217–218 and p. 283), and (Horty, 2001, pp. 30–31).

[9] The four selection functions are s_1, s_2, s_3, and s_4, where $s_1(a) = \{w_1, w_2\}$, $s_1(b) = \{w_1, w_3\}$; $s_2(a) = \{w_1, w_2\}$, $s_2(b) = \{w_2, w_4\}$; $s_3(a) = \{w_3, w_4\}$, $s_3(b) = \{w_1, w_3\}$; and $s_4(a) = \{w_3, w_4\}$, $s_4(b) = \{w_2, w_4\}$.

[10] Given the four selection functions of footnote 9, it holds that $\{w_1\} = s_1(a) \cap s_1(b)$, $\{w_2\} = s_2(a) \cap s_2(b)$, $\{w_3\} = s_3(a) \cap s_3(b)$, and $\{w_4\} = s_4(a) \cap s_4(b)$.

[11] Harsanyi contends: "[T]he more complete our factual information and the more completely individualistic our ethics, the more the different individuals' social welfare functions will converge toward the same objective quantity, namely, the unweighted sum (or rather the unweighted arithmetic mean) of all individual utilities" (Harsanyi, 1955, p. 320).

[12] Horty's deontic logic aims to model utilitarian obligations only. Therefore, Horty assumes that all obligations stem from the single moral code of utilitarianism, which he defines in terms of agent-neutral utilities. Compare (Horty, 2001, pp. 36–37 and 41–42). Our consequentialist multi-agent deontic logic requires utility functions based on normalized agent-dependent utilities.

[13] Observe that our analysis is sensitive to a theory of group utility. If we had defined group utility in terms of, for example, Amartya Sen's leximin rule which maximizes the utility of the worst-off individual (Sen, 1970, p. 138), rather than via Harsanyi's arithmetical mean, we would obtain neither $\mathfrak{M} \models \odot_a^{a,b} \neg p$ nor $\mathfrak{M} \models \odot_b^{a,b} \neg q$.

[14] See (van Benthem, 1984) and (Blackburn et al., 2001, p. 126).

[15] Compare (Feldman, 1986, p. 209).

[16] As $Choice(\emptyset) = \{W\}$, these conditions imply that $\mathcal{G}_1 - \mathcal{G}_2$, $\mathcal{G}_2 - \mathcal{G}_1$, and $\mathcal{G}_1 \cap \mathcal{G}_2$ are all non-empty. Thus, any moral conflict within a single moral code involves at least *three* individual agents. Consequently, the present system of consequentialist logic throws doubts upon Terrance McConnell's claim that as regards multi-agent moral conflicts generated by

the same moral code "the two-person case may be taken as typical" (McConnell, 1988, p. 25).

[17] Compare (Lemmon, 1962, p. 151) and (McConnell, 1988, p. 31).

[18] Plato, *Phaedrus*, translated by R. Hackforth (in E. Hamilton & H. Cairn (eds.), *The Collected Dialogues of Plato*, Princeton: Princeton University Press, 1963), 246b4.

[19] We wish to thank John Horty, Erik Krabbe, Martin van Hees, the members of the Groningen research colloquium in theoretical philosophy, and an anonymous referee of this journal for their critical comments on earlier versions of this paper.

REFERENCES

Austin, J. L.: 1957, A plea for excuses, in J. L. Austin, *Philosophical Papers*, 2nd ed., Oxford University Press, Oxford, 1970, pp. 175–204.

Belnap, N., Perloff, M., and Xu, M.: 2001, *Facing the Future*, Oxford University Press, New York.

Blackburn, P., de Rijke, M., and Venema, Y.: 2001, *Modal Logic*, Cambridge University Press, Cambridge.

Conee, E.: 1982, Against moral dilemmas, *Philosophical Review* **91**, 87–97.

Feldman, F.: 1986, *Doing the Best We Can*, D. Reidel Publishing Company, Dordrecht.

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

Hamblin, C. L.: 1972, Quandaries and the logic of rules, *Journal of Philosophical Logic* **1**, 74–85.

Harsanyi, J. C.: 1955, Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *Journal of Political Economy* **63**, 309–321.

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

Horty, J. F.: 2001, *Agency and Deontic Logic*, Oxford University Press, New York.

Lemmon, E. J.: 1962, Moral dilemmas, *Philosophical Review* **71**, 139–158.

Lemmon, E. J.: 1965, Deontic logic and the logic of imperatives, *Logique & Analyse* **29**, 39–71.

Marcus, R. B.: 1980, Moral dilemmas and consistency, *Journal of Philosophy* **77**, 121–136.

McConnell, T.: 1988, Interpersonal moral conflicts, *American Philosophical Quarterly* **25**, 25–35.

Osborne, M. and Rubinstein, A.: 1994, *A Course in Game Theory*, The MIT Press, Cambridge, MA.

Sen, A. K.: (1970), *Collective Choice and Social Welfare*, Holden-Day, San Francisco.

van Benthem, J. F. A. K.: 1984, Correspondence theory, in D. Gabbay and F. Guenther (eds.), *Handbook of Philosophical Logic*, Vol. II, D. Reidel Publishing Company, Dordrecht, pp. 167–247.

van Fraassen, B. C.: 1973, Values and the heart's command, *Journal of Philosophy* **70**, 5–19.

von Wright, G. H.: 1963, *Norm and Action*, Routledge & Kegan Paul, London.

von Wright, G. H.: 1966, The logic of action: a sketch, in N. Rescher (ed.), *The Logic of Decision and Action*, University of Pittsburgh Press, Pittsburgh, pp. 121–136.

Williams, B.: 1973, Ethical consistency, in B. Williams (ed.), *Problems of the Self*, Cambridge University Press, Cambridge, pp. 166–186.

Department of Theoretical Philosophy
Faculty of Philosophy
University of Groningen
Oude Boteringestraat 52
9712 GL Groningen
The Netherlands
E-mail: B.P.Kooi@rug.nl
E-mail: A.M.Tamminga@rug.nl