

# Chapter 11

## Perspectival Act Utilitarianism

John F. Horty

### 11.1 Introduction

This chapter works within a particular framework for reasoning about actions – sometimes known as the framework of “stit semantics” – originally due to Belnap and Perloff, based ultimately on the theory of indeterminism set out in Prior’s indeterministic tense logic, and developed in full detail by Belnap et al. (2001). The issues I want to consider arise when certain normative, or decision theoretic, notions are introduced into this framework: here I will focus on the notion of a *right action*, and so on the formulation of act utilitarianism within this indeterministic setting. The problem is simply that there are two different, and conflicting, ways of defining this notion, both well-motivated, and both carrying intuitive weight.

This problem was first pointed out in my Horty (2001), but here I address what I now think of as a mistake in that treatment. In that earlier book, in order to explain our conflicting judgments about right actions, I set out two substantially different accounts of the notion, which I labeled as the “dominance” and “orthodox” accounts. But here, there is only one account, only one theory of right actions, and our conflicting intuitions are instead explained by showing how this theory yields different results when actions are evaluated from different perspectives. In effect, a semantic explanation, which postulates an ambiguity in the notion of a right action, is replaced by a pragmatic explanation.

The chapter is structured as follows. In the next section, I review Prior’s indeterministic framework as well as the structures underlying stit semantics. Although these structures were originally introduced for the purpose of interpreting formal languages containing special modal operators – tense operators, agency operators – there is none of that here. The concepts I am concerned with in this chapter are

---

J.F. Horty (✉)

Philosophy Department and Institute for Advanced Computer Studies, University of Maryland,  
College Park, MD 20742, USA  
e-mail: horty@umiacs.umd.edu

defined entirely in terms of the underlying structures themselves; there is no need to introduce or interpret any formal language. In the third and fourth sections, I motivate the two ways of understanding the notion of a right action, and define the corresponding dominance and orthodox act utilitarian theories. Finally, in the fifth section, I show how these two theories can be unified, and how our conflicting intuitions about right actions can then be explained as resulting from the different perspectives from which actions might be evaluated. An appendix shows how the account can be generalized to group as well as individual actions, and how the relation between the right actions available to a group and to the individuals belonging to that group can then be seen to depend on the perspective from which these actions are evaluated.

## 11.2 Background

### 11.2.1 Individual Actions

Prior's theory of indeterminism, set out in his Prior (1967) and developed in more detail by Thomason (1970), is based on a picture of moments as ordered into a tree-like structure, with forward branching representing the openness or indeterminacy of the future and the absence of backward branching representing the determinacy of the past.

This picture can be represented as a nonempty set *Tree* of moments together with an ordering  $<$  on *Tree* that is transitive and irreflexive, and that satisfies the treelike property according to which, for any  $m_1, m_2$ , and  $m_3$  in *Tree*, if  $m_1 < m_3$  and  $m_2 < m_3$ , then either  $m_1 = m_2$  or  $m_1 < m_2$  or  $m_2 < m_1$ . A maximal set of linearly ordered moments from *Tree* is a *history*, representing some complete temporal evolution of the world. If  $m$  is a moment and  $h$  is a history, then the statement that  $m \in h$  can be taken to mean that  $m$  occurs at some point in the course of the history  $h$ , or that  $h$  passes through  $m$ . Of course, because of indeterminism, a single moment might be contained in several distinct histories. We let  $H_m = \{h : m \in h\}$  represent the set of histories passing through  $m$ , those histories in which  $m$  occurs; and when  $h$  belongs to  $H_m$ , we speak of a moment/history pair of the form  $m/h$  as an *index*.

In this framework, it is the histories themselves that represent possibilities, or "possible worlds." The set of possible worlds accessible at a moment  $m$  can thus be identified with the set  $H_m$  of histories passing through that moment; those histories lying outside of  $H_m$  are taken to represent worlds that are no longer accessible. We can therefore identify the *propositions* at  $m$  with the subsets of  $H_m$ , where of course,  $H_m$  itself is the least informative of these propositions.

These various ideas can be illustrated as in Fig. 11.1, where the upward direction represents the forward direction of time. This diagram depicts a branching time structure containing five histories,  $h_1$  through  $h_5$ . The moments  $m_1$  through  $m_4$  are highlighted; and we have, for example,  $m_2 \in h_3$  and  $H_{m_4} = \{h_4, h_5\}$ .

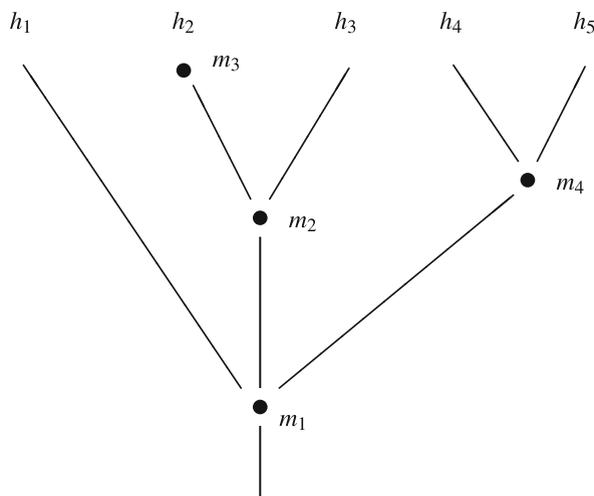


Fig. 11.1 Branching time

We now turn to the treatment of agency. The goal is to represent the notion that an agent, through its action, guarantees the truth of some proposition.<sup>1</sup> We must therefore be able to speak of individual agents, and also of their actions or choices; and so the basic framework of branching time is supplemented with two additional primitives.

The first is simply a set *Agent* of agents, individuals thought of as making choices, or acting, in time.

Now what is it for one of these agents to act, or choose, in this way? We idealize by ignoring any intentional components involved in the concept of action, by ignoring vagueness and probability, and also by treating actions as instantaneous. In this rarefied environment, acting can be thought of simply as constraining the course of events to lie within some definite subset of the possible histories still available. When an agent  $\alpha$  butters the toast, for example, the nature of its action is to constrain the history to be realized so that it must lie among those in which the toast is buttered. Of course, such an action still leaves room for a good deal of variation in the future course of events, and so cannot determine a unique history; but it does rule out all those histories in which the toast is not buttered.

Our second additional primitive, then, is a device for representing the possible constraints that an agent is able to exercise upon the course of events at a given moment, the actions or choices open to the agent at that moment. These constraints are encoded formally through a function *Choice*, mapping each agent  $\alpha$  and moment

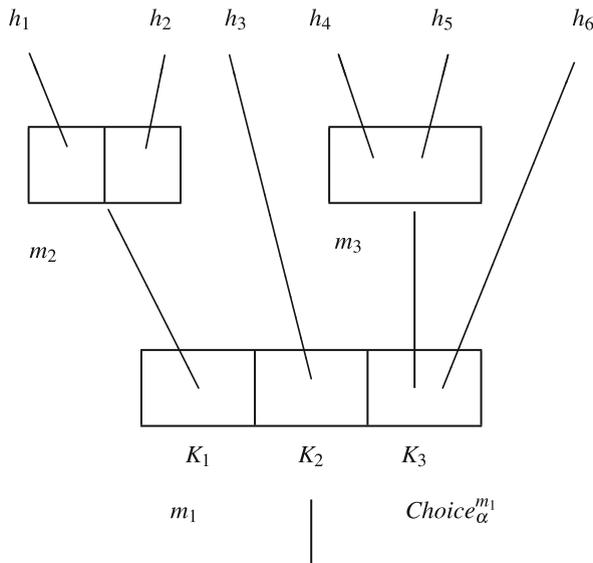
<sup>1</sup>In an effort to find language that is both gender neutral and unobtrusive, I assume here that the agents are impersonal acting devices, such as robots, which it is appropriate to refer to using the pronoun “it”.

$m$  into a partition  $Choice_\alpha^m$  of the set of histories  $H_m$  through  $m$ .<sup>2</sup> The idea behind this formalism is that, by acting at  $m$ , the agent  $\alpha$  selects a particular one of the equivalence classes, or choice cells, from  $Choice_\alpha^m$  within which the history to be realized must then lie, but that this is the extent of the agent’s influence. If  $K$  is such a choice cell, an equivalence class from  $Choice_\alpha^m$ , we speak of  $K$  as an *action* available to the agent  $\alpha$  at the moment  $m$ , and we speak of the histories belonging to  $K$  as the possible *outcomes* that might result from this action.

These various concepts relating to choice functions are illustrated in Fig. 11.2, which depicts a structure containing six histories, and in which the actions available to the agent  $\alpha$  at three moments are highlighted. The cells at the highlighted moments represent the actions available to  $\alpha$  at those moments. For example, there are three actions available to  $\alpha$  at  $m_1$  –  $Choice_\alpha^{m_1} = \{K_1, K_2, K_3\}$ , with  $K_1 = \{h_1, h_2\}$ ,  $K_2 = \{h_3\}$ , and  $K_3 = \{h_4, h_5, h_6\}$ . If the agent selects  $K_3$ , then the histories  $h_4, h_5,$  and  $h_6$  are the possible outcomes of its action.

### 11.2.2 Group Actions

To see how this account can be extended to group actions, it is best to begin with an example; so consider the multiple agent situation depicted in Fig. 11.3. Here, the actions open to the agent  $\alpha$  at the moment  $m$  are depicted by the vertical partitions of



**Fig. 11.2** An agent’s choices

<sup>2</sup> The *Choice* function is subject to two technical constraints of “no choice between undivided histories” and “independence of actions,” which I will not go into here. The constraints can be found in my Horty (2001), and are described in authoritative detail in Belnap et al. (2001).

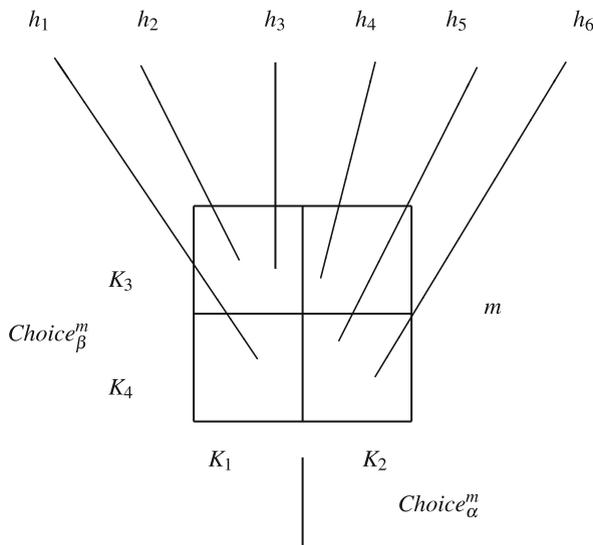


Fig. 11.3 Group actions

$H_m$ ; that is,  $Choice_\alpha^m = \{K_1, K_2\}$ , with  $K_1 = \{h_1, h_2, h_3\}$  and  $K_2 = \{h_4, h_5, h_6\}$ . The actions open to the agent  $\beta$  are depicted by the horizontal partitions;  $Choice_\beta^m = \{K_3, K_4\}$ , with  $K_3 = \{h_2, h_3, h_4\}$  and  $K_4 = \{h_1, h_5, h_6\}$ .

Now consider the proposition  $X = \{h_2, h_3, h_6\}$ . It should be clear that, in this situation, neither the agent  $\alpha$  nor the agent  $\beta$  acting alone has the ability to guarantee the truth of  $X$ . Each action available to each of these agents allows for a possible outcome in which  $X$  fails. Still, it seems that the group of agents  $\{\alpha, \beta\}$  acting together does have the ability to guarantee the truth of  $X$ . If  $\alpha$  performs the action  $K_1$  and  $\beta$  performs the action  $K_3$ , the group  $\{\alpha, \beta\}$  can be said to perform the action  $K_1 \cap K_3$ , and  $X$  holds at each possible outcome of this group action.

As this example suggests, group actions can usefully be defined as patterns of individual actions: an action available to a group of agents can be defined as an intersection of the actions available to the individual agents belonging to that group, one action for each agent.

In order to develop this suggestion, it is convenient to reify patterns of action by defining an *action selection* function at a moment  $m$  as a function assigning to each agent some action available to that agent at  $m$  – that is, a function  $s$  mapping each agent  $\alpha$  into some member of  $Choice_\alpha^m$ . Each of these action selection functions represents a possible pattern of action at the moment  $m$ , a selection of an available action for each agent. These patterns of action can be collected together into the set  $Select_m$ , containing the various action selection functions at  $m$ . And where  $\Gamma$  is a group of agents, the set  $Choice_\Gamma^m$  of action available to the group at the moment  $m$  – the patterns of action available to the members of the group – can then be defined as follows:

$$Choice_\Gamma^m = \left\{ \bigcap_{\alpha \in \Gamma} s(\alpha) : s \in Select_m \right\}.$$

It should be clear that this definition says what it should: the set of actions available to the group  $\Gamma$  is identified with the set of intersections of actions available to the agents belonging to that group, one action for each agent.

### 11.3 The Dominance Account

#### 11.3.1 Our Question

With this much of the framework in place, we now add one final primitive: a function Value mapping each history into a real number representing the overall value of that history, however that is conceived. This new primitive is illustrated in Fig. 11.4, where the numbers written beside histories indicate the values assigned to those histories, so that, for example,  $\text{Value}(h_1) = 10$ .

Now that values have been assigned to the various histories consistent with an agent’s actions – the various possible outcomes of those actions – we can turn to the central question of this chapter: How, in this indeterministic setting, can we characterize the act utilitarian notion of a *right action* for the agent to perform?

According to the standard formulation of act utilitarianism, an action is defined as right if there is no action among the available alternatives with better consequences, and wrong otherwise.<sup>3</sup> In the present framework, it is easy enough to define the

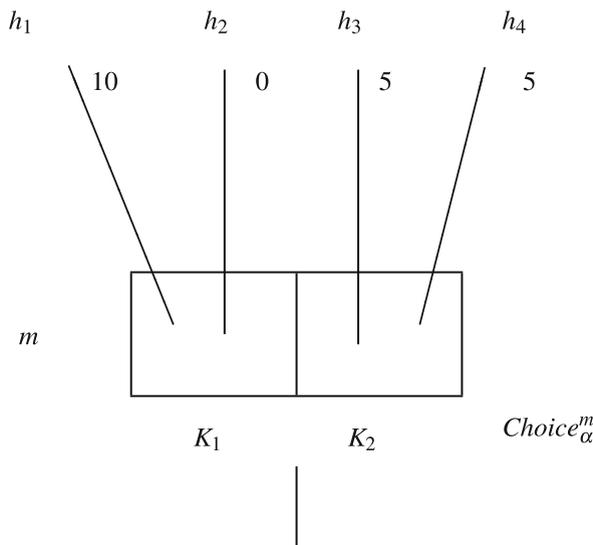


Fig. 11.4 Histories with values

<sup>3</sup> Perhaps the most careful formulation of act utilitarianism can be found in Bergström (1966); for work along similar lines, see Carlson (1995).

alternatives available to an agent  $\alpha$  at a moment  $m$ ; these are simply the actions from  $Choice_\alpha^m$ . And our Value function, of course, provides a straightforward ranking of possible outcomes. But in a setting that is genuinely indeterministic, how can we define the notion of an action's consequences?

The problem that a robust indeterminism presents for the characterization of an action's consequences – and so for a definition of act utilitarianism – was noted some time ago by Prior, in his contribution to a symposium on the topic:

Suppose that determinism is *not* true. Then there may indeed be a number of alternative actions which we could perform on a given occasion, but none of these actions can be said to have any “total consequences,” or to bring about a definite state of the world which is better than any other that might be brought about by other choices ... it's not merely that one cannot calculate the totality of what will happen if one decides in a certain way; the point is rather that there *is* no such totality. (Prior 1956, pp. 91–92)

And the general point is clear enough. In the case of Fig. 11.4, for example, the agent must choose between two available actions. The choice of  $K_2$  leads invariably to an outcome whose value is 5, while the choice of  $K_1$  leads to an outcome whose value is either 10 or 0, depending on whether things evolve along the lines of  $h_1$  or  $h_2$ . But since, if  $K_1$  is selected, it is then indeterminate whether  $h_1$  or  $h_2$  will be realized, how can we possibly say which of the two actions,  $K_1$  or  $K_2$ , has the better consequences?

In response to this problem, Prior himself offers the standard suggestion of appealing to probabilistic information, such as a probability distribution on the histories that might result from an action. Using this information, we could assign an expected value to each of the actions available to an agent, and the ordering of actions based on their expected values would then allow us to define a form of act utilitarianism that did not, in fact, rely on some definite notion of an action's consequences: an action could be defined as right whenever there is no alternative with greater expected value.

This approach – leading to a theory that might be described as *expected value act utilitarianism* – is, of course, very natural when the required probability distribution can be found. But there are many situations in which this information is either unavailable or meaningless; this is true, particularly, when the outcome resulting from an agent's action depends, not simply on a roll of the dice, but on the independent choice of another free agent. In the literature on decision theory, a situation in which the actions available to an agent might lead to their various possible outcomes with known probability is described as a case of *risk*, while a situation in which the probability with which the available actions might lead to their various possible outcomes is either unknown or meaningless is described as a case of *uncertainty*.<sup>4</sup>

---

<sup>4</sup> A discussion of this terminology can be found, for example, in Sections 2.1 and 13.1 of Luce and Raiffa (1957). Of course, the legitimacy of the distinction between uncertainty and risk is itself an issue: following Ramsey (1931) and Savage (1954), many writers in the Bayesian tradition assume that an agent's assessment of the possible outcomes in a given situation can always be represented through a probability measure, so that uncertainty always reduces to risk. However, there is an

Our concern here is with situations involving uncertainty, rather than risk, and we proceed by adapting a standard treatment of these situations from decision theory: since an ordering based on expected value is not possible, we instead define a notion of dominance that can be used to order the actions available to an agent.

### 11.3.2 Dominance Act Utilitarianism

We begin with a preference ordering on propositions, arbitrary sets of histories through a moment.

PREFERENCES ORDERING ON PROPOSITIONS: Let  $X$  and  $Y$  be propositions at a moment. Then  $X \leq Y$  ( $Y$  is *weakly preferred* to  $X$ ) if and only if  $\text{Value}(h) \leq \text{Value}(h')$  for each  $h \in X$  and each  $h' \in Y$ ; and  $X < Y$  ( $Y$  is *strongly preferred* to  $X$ ) if and only if  $X \leq Y$  and it is not the case that  $Y \leq X$ .

The idea is that, if  $Y$  is weakly preferred to  $X$ , each history from  $Y$  is at least as valuable as any history from  $X$ , so that we are sure to do at least as well in a history at which  $Y$  holds as we would in a history at which  $X$  holds. If  $Y$  is strongly preferred to  $X$ , then not only is each history from  $Y$  at least as valuable as any history from  $X$ , but some history from  $Y$  is actually more valuable than some history from  $X$ , so that we are not only sure to do at least as well with  $Y$  as with  $X$ , we might do better.

In the current framework, the actions available to an agent at a moment are reified as sets of histories through that moment. Each action is therefore a proposition, and so it is tempting to imagine that the dominance relations among actions might be identified with the preference orderings defined for propositions more generally. This idea is plausible, and there are a number of examples in which it seems to yield the correct results, including the earlier Fig. 11.4, where it tells us that neither of the two actions,  $K_1$  or  $K_2$ , dominates the other. However, the suggestion of simply identifying the dominance orderings over an agent's actions with the preference orderings on propositions fails in more complicated cases.

To see this, consider Fig. 11.5, depicting a situation of simultaneous choice by two agents, and interpreted as follows. We suppose that the agent  $\alpha$  is holding a nickel in its hand, and that at the moment  $m$ , the agent is faced with a choice between two actions: placing this nickel on a certain table either heads up, performing the action  $K_1$ , or tails up, performing the action  $K_2$ . At the same moment, the agent  $\beta$  must likewise choose between placing a dime on the table either heads up or tails up, performing either the action  $K_3$  or the action  $K_4$ . If  $\alpha$  places the nickel on the table heads up, then the resulting utility is 9 if  $\beta$  places the dime heads up and 4 if  $\beta$  places the dime tails up; but if  $\alpha$  places the dime on the table tails up, the resulting utility is 10 if  $\beta$  places the dime heads up and 5 if  $\beta$  places the dime tails up.

---

important tradition of resistance to the assimilation of uncertainty and risk in a single numerical measure. A classic paper in this tradition is Ellsberg (1961); for more recent work on decision theory in situations that mix elements of risk and uncertainty, see the papers contained in Parts II and IV of Gärdenfors and Sahlin (1988).

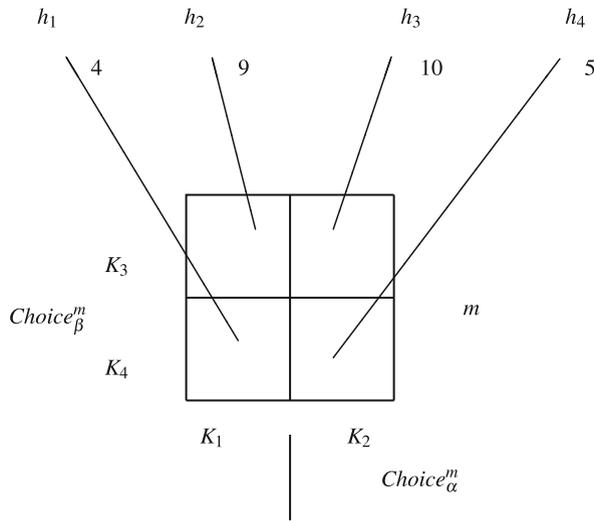


Fig. 11.5 The coin example

In this situation, neither of the two actions open to  $\alpha$  is even weakly preferred to the other in the sense of the propositional ordering, since each contains an outcome more valuable than some outcome belonging to the other. Nevertheless, there is a persuasive argument in favor of the conclusion that  $K_2$  is a better action than  $K_1$  for  $\alpha$  to perform: The agent  $\beta$  must place the dime on the table either heads up or tails up, performing either  $K_3$  or  $K_4$ . So suppose, first, that  $\beta$  places the dime heads up, performing  $K_3$ . In that case, it is clearly better for  $\alpha$  to place the nickel on the table tails up, performing  $K_2$  rather than  $K_1$ , since the unique history  $h_3$  belonging to  $K_2 \cap K_3$  is more valuable than the unique history  $h_2$  belonging to  $K_1 \cap K_3$ . Next, suppose that  $\beta$  places the dime tails up, performing  $K_4$ . Then it is again better for  $\alpha$  to place the nickel on the table tails up, again performing  $K_2$  rather than  $K_1$ , since the unique history  $h_4$  belonging to  $K_2 \cap K_4$  is more valuable than the unique history  $h_1$  belonging to  $K_1 \cap K_4$ . In each of these two cases, then, it is better for  $\alpha$  to perform  $K_2$  rather than  $K_1$ , and since these exhaust the possibilities, a pattern of reasoning sometimes described as the *sure-thing principle* suggests that  $K_2$  is simply a better action than  $K_1$  for  $\alpha$  to perform.<sup>5</sup>

The key to applying sure-thing reasoning in a given situation lies in identifying an appropriate partition of the possible outcomes into a set of *states* (sometimes called “states of nature” or “conditioning events”), against the background of which

<sup>5</sup> This pattern of reasoning is first explicitly characterized as the “sure-thing principle” in Savage (1954), but the principle appears already in some of Savage’s earlier work, such as (Savage 1951, p. 58), where he writes concerning situations of uncertainty that “there is one unquestionably appropriate criterion for preferring some act to some others: If for every possible state, the expected income of one act is never less and is in some cases greater than the corresponding income of another, then the former act is preferable to the latter.”

the actions available to an agent can then be evaluated through a state-by-state comparison of their results. This is often a difficult task, but we simplify in the current setting, not only by supposing that probabilistic information is unavailable, but also by imagining that the only sources of causality present are the actions of the various agents.

Given these assumptions, it is natural to identify the set of states confronting an agent  $\alpha$  at the moment  $m$  – here abbreviated as  $State_\alpha^m$  – with the possible patterns of action that might be performed at that moment by all other agents. In the case of Fig. 11.5, for example, if we assume that  $\alpha$  and  $\beta$  are the only two agents – that is,  $Agent = \{\alpha, \beta\}$  – then  $State_\alpha^m$  can be identified with  $Choice_\beta^m$ , the set  $\{K_3, K_4\}$  of actions available to  $\beta$ . Although we concentrate in this chapter on simple cases like this, with two agents at most, the definition of a state is more general. Where  $Agent$  contains an arbitrary group of agents, the set of agents other than  $\alpha$  is  $Agent - \{\alpha\}$ , of course, and we can then define the set of states confronting  $\alpha$  at  $m$  by stipulating that:

$$State_\alpha^m = Choice_{Agent - \{\alpha\}}^m.$$

Given this treatment of the states facing an agent, we can now define a dominance ordering on the actions available to the agent through a state-by-state comparison of their results. As an initial step, we must first specify a standard for comparing the possible results of two actions against the background of a particular state. The example depicted in Fig. 11.5 is deceptively simple in this regard, for in this situation, once a particular state from  $State_\alpha^m$  is fixed, each action available to  $\alpha$  then determines a unique outcome, so that these actions can simply be ranked along with their outcomes.

In the more general case, of course, even against the background of a fixed state, the actions available to an agent may determine only sets of outcomes, or propositions, rather than unique outcomes – but here, we can compare the results of different actions in a state by appealing to the preference ordering defined earlier on propositions. Where  $S$  is a state belonging to  $State_\alpha^m$ , and where  $K$  and  $K'$  are actions available to  $\alpha$  at  $m$ , we can say that the results of  $K'$  are at least as good as those of  $K$  in the state  $S$  whenever  $K \cap S \leq K' \cap S$  – whenever, that is, the proposition  $K' \cap S$ , determined by performing the action  $K'$  in the state  $S$ , is weakly preferred to the proposition  $K \cap S$ , determined by performing  $K$  in  $S$ .

With these various concepts in place, we are now in a position to define a dominance ordering on the actions available to an agent at a moment.

**DOMINANCE ORDERING ON ACTIONS:** Let  $\alpha$  be an agent and  $m$  a moment, and let  $K$  and  $K'$  be members of  $Choice_\alpha^m$ . Then  $K \leq K'$  ( $K'$  weakly dominates  $K$ ) if and only if  $K \cap S \leq K' \cap S$  for each state  $S \in State_\alpha^m$ ; and  $K < K'$  ( $K'$  strongly dominates  $K$ ) if and only if  $K \leq K'$  and it is not the case that  $K' \leq K$ .

The idea is that,  $K'$  weakly dominates  $K$ , then the results of performing  $K'$  are at least as good as those of performing  $K$  in every state, so that, no matter which state is realized, the agent is sure to do at least as well with  $K'$  as with  $K$ . If  $K'$  strongly

dominates  $K$ , then not only are the results of performing  $K'$  at least as good as those of performing  $K$  in every state, but there is some state in which  $K'$  yields better results, so that the agent is sure to do at least as well with  $K'$  as with  $K$ , and might do better.

Let us now return to our central question: how, in this indeterminist setting, can we define the utilitarian notion of a right action? The dominance account provides an answer that is both precise and intuitively plausible.

We begin by defining the set  $Optimal^m_\alpha$  containing the *optimal actions* available to an agent  $\alpha$  at a moment  $m$ , those actions available to the agent that are not strongly dominated by any others:

$$Optimal^m_\alpha = \{K \in Choice^m_\alpha : \neg \exists K' \in Choice^m_\alpha (K < K')\}.$$

It is then natural to formulate a theory that might be characterized as *dominance act utilitarianism* simply by identifying the right actions available to an agent at a moment with the optimal actions.

**DOMINANCE ACT UTILITARIANISM:** Let  $\alpha$  be an agent and  $m$  a moment, and suppose  $K \in Choice^m_\alpha$ . Then the action  $K$  is *right* at the moment  $m$  if and only if  $K \in Optimal^m_\alpha$ , and *wrong* otherwise.

The theory can be illustrated with our earlier examples. In the case of Fig. 11.4, we have  $Optimal^m_\alpha = \{K_1, K_2\}$ , so that both actions available to the agent at the moment  $m$  are right. In the case of Fig. 11.5, we have  $Optimal^m_\alpha = \{K_2\}$ , so that  $K_2$  is right and  $K_1$  is wrong.

## 11.4 The Orthodox Account

### 11.4.1 An Example

This theory of dominance act utilitarianism is, I suspect, not too surprising. It is perhaps even obvious. The underlying ideas of dominance and optimality are familiar from decision theory, generalized only slightly to allow for the fact that an action in a state yields a proposition, rather than a unique outcome.

What may be surprising, however – and particularly if the dominance theory does seem to be obvious – is the fact that the treatment of utilitarianism within the ethical literature does not follow this dominance account at all, but is based on an entirely different approach, which I will refer to, in deference to the literature, as the *orthodox* account.

In order to illustrate this orthodox account, let us consider an example that has figured prominently in the discussion of different forms of utilitarianism. Although the example was first introduced by Gibbard (1965), and was elaborated on shortly thereafter by Sobel (1968), I take the later but more extensive discussion by Regan (1980) as my primary source:

Suppose that there are only two agents in the moral universe, called Whiff and Poof. Each has a button in front of him which he can push or not. If both Whiff and Poof push their buttons, the consequences will be such that the overall state of the world has a value of ten units. If neither Whiff nor Poof pushes his button, the consequences will be such that the overall state of the world has a value of 6 units. Finally, if one and only one of the pair pushes his button (and it does not matter who pushes and who does not), the consequences will be such that the overall state of the world has a value of 0 (zero) units. Neither agent, we assume, is in a position to influence the other's choice. (Regan 1980, p. 19)

In the present framework, this example can be depicted as in Fig. 11.6, where  $\alpha$  represents Whiff,  $\beta$  represents Poof, and  $m$  is the moment at which each of these two agents must choose whether or not to push his button.<sup>6</sup> The action  $K_1$  represents Whiff's option of pushing his button, and  $K_2$  his option of refraining; likewise,  $K_3$  and  $K_4$  represent Poof's options of pushing or refraining; and the possible outcomes resulting from the choices by these agents are represented by the histories  $h_1$  through  $h_4$ , which are assigned the values indicated in Regan's description.

Now, when the example is set out in this way, it is easy to see that both agents will satisfy our previous theory of dominance act utilitarianism no matter what they do. Neither action available to either agent is dominated, and so we have both  $Optimal^m_\alpha = \{K_1, K_2\}$  and  $Optimal^m_\beta = \{K_3, K_4\}$ . Since both of the actions  $K_1$  and  $K_2$  available to Whiff are optimal, both are right according to the dominance theory; and both of the actions  $K_3$  and  $K_4$  available to Poof are right as well.

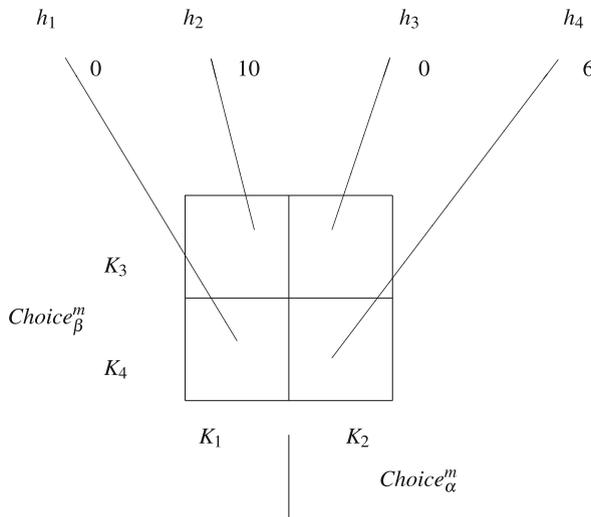


Fig. 11.6 Whiff and Poof

<sup>6</sup> Regan does not actually require that these choices must be simultaneous (though simultaneity is part of Gibbard's earlier description), but he does require the choices to be independent, and we guarantee independence through simultaneity.

The theory of dominance act utilitarianism, then, yields results that are at least definite in this case, even if not particularly constraining: each of the two agents can satisfy the theory by selecting either of the available actions. However, Regan's own conclusions – based on his own theory of act utilitarianism or, as he calls it, AU – are strikingly different:

Now, if we ask what AU directs Whiff to do, we find that we cannot say. If Poof pushes, then AU directs Whiff to push. If Poof does not push, then AU directs Whiff not to push. Until we specify how Poof behaves, AU gives Whiff no clear direction. The same is true, *mutatis mutandis*, of Poof. (Regan 1980, p. 18)

In saying that act utilitarianism gives Whiff no clear direction, Regan does not mean only that this theory, like the dominance theory, classifies multiple actions as right, allowing the agents to choose among them. Instead, he means that, on the basis only of the information provided so far, the theory is simply unable to generate any results at all: no actions can be classified either as right or as wrong. In order to arrive at a situation in which act utilitarianism is able to yield definite results, Regan feels that it is necessary to supplement the description of the example provided so far, and depicted in Fig. 11.6, with additional information concerning the actions actually performed by the individuals involved:

If we shift our attention to patterns of behavior for the pair, we can decide whether each agent satisfies AU in any specified pattern. (Regan 1980, p. 18)

And he illustrates the kind of reasoning allowed by this additional information as follows:

Suppose, for example, Whiff and Poof both push their buttons. The total value thereby achieved is ten units. Does Whiff satisfy AU? Yes. The only other thing he might do is not push his button. But under the circumstances, which include the fact that Poof pushes his button, Whiff's not pushing would result in a total utility of zero. Therefore Whiff's pushing his button has at least as good consequences as any other action available to him under the circumstances. Therefore, it is right according to AU (Regan 1980, pp. 18–19)

### 11.4.2 *Orthodox Act Utilitarianism*

Evidently, Regan is unwilling to classify actions as right or wrong absolutely, but only as right or wrong in particular circumstances. That is fair enough. But Regan, following Gibbard and Sobel, also takes the further, and more contentious step of supposing that an agent's circumstances must include whatever actions are simultaneously performed by other agents – so that he is unwilling to classify the actions available to Whiff and Poof as either right or wrong absolutely, but only as right or wrong under the circumstances determined by the actions of the other.<sup>7</sup>

---

<sup>7</sup> Gibbard adopts a similar viewpoint in his original discussion of this example, evaluating each agent's selection only under an assumption about the action selected by the other (Gibbard 1965, p. 215). And Sobel defends Gibbard's strategy as follows: "It is perhaps natural to feel that Gibbard's first case is objectionable just because it includes assumptions concerning what agents will

How can we represent the theory of act utilitarianism that guides Regan's judgments? In my Horty (2001), I adopted a strategy, which still seems reasonable to me, and which I review here, of first introducing a concept of conditional optimality, and then conditionalizing on a proposition that represents the agent's circumstances.

The concept of conditional optimality is introduced in three steps. First, taking  $X$  as a proposition, the set of actions available to an agent  $\alpha$  at  $m$  under the condition that  $X$  holds – expressed here as  $Choice_\alpha^m / X$  – is simply the set containing those actions open to  $\alpha$  at  $m$  that are consistent with  $X$ :

$$Choice_\alpha^m / X = \{K \in Choice_\alpha^m : K \cap X \neq \emptyset\}.$$

The next step is to generalize our earlier treatment of dominance to include conditional dominance.

CONDITIONAL DOMINANCE ORDERING ON ACTIONS: Let  $\alpha$  be an agent and  $m$  a moment, and let  $K$  and  $K'$  be members of  $Choice_\alpha^m$ , and  $X$  a proposition. Then  $K \preceq_X K'$  ( $K'$  weakly dominates  $K$  under the condition  $X$ ) if and only if  $K \cap X \cap S \leq K' \cap X \cap S$  for each state  $S \in State_\alpha^m$ ; and  $K \prec_X K'$  ( $K'$  strongly dominates  $K$  under the condition  $X$ ) if and only if  $K \preceq_X K'$  and it is not the case that  $K' \preceq_X K$ .

This conditional analysis follows the pattern of the absolute treatment set out earlier, except that, in comparing the results of two actions  $K$  and  $K'$  in a given state  $S$ , our attention is now restricted only to those outcomes that are consistent with the background proposition  $X$ .

Finally, having generalized both choice and dominance to the conditional setting, we can now combine these ideas to arrive at a concept of conditional optimality. Again taking  $X$  as a proposition, we define the set of optimal actions available to  $\alpha$  at  $m$  under the condition  $X$  – expressed as  $Optimal_\alpha^m / X$  – to be the set of those actions available to  $\alpha$  at  $m$  under the condition  $X$  that are not strongly dominated under this condition by any other such action:

$$Optimal_\alpha^m / X = \{K \in Choice_\alpha^m / X : \neg \exists K' \in Choice_\alpha^m / X (K \prec_X K')\}.$$

It is easy to verify, but worth noting explicitly that the conditional notions of choice, dominance, and optimality introduced here are, in fact, generalizations of our earlier concepts. When the background condition  $X$  is identified with the trivial proposition  $H_m$  – that is, when  $X = H_m$  – each of these three conditional notions coincides with its absolute counterpart. In particular, we have

$$Optimal_\alpha^m / H_m = Optimal_\alpha^m;$$

the actions available to  $\alpha$  at  $m$  that are optimal under the condition that the trivial proposition holds are simply the optimal actions.

---

and would do. But this can be no objection since it is obvious that such assumptions are essential to the application of AU; without such assumptions the dictates of AU could not be determined ...” (Sobel 1968, p. 152).

Now that the notion of conditional optimality has been introduced, it remains only to define the propositions on which we conditionalize.<sup>8</sup> Just as  $Choice_\alpha^m/X$  represents the set of actions available to  $\alpha$  at  $m$  that are consistent with  $X$ , we can likewise define

$$State_\alpha^m/X = \{K \in State_\alpha^m : K \cap X \neq \emptyset\}$$

as the set of states confronting  $\alpha$  at  $m$  that are consistent with  $X$ . And in this case, it is also convenient to represent the proposition formed by taking the union of these states – the proposition, that is, according to which one of these states holds – written  $State_\alpha^m(X)$  and defined as follows:

$$State_\alpha^m(X) = \bigcup State_\alpha^m/X.$$

To illustrate this notation, suppose in the case of Fig. 11.6, the Whiff and Poof example, that the proposition  $X = \{h_2, h_4\}$ . Then  $State_\alpha^m/X = \{K_3, K_4\}$  is the set of states confronting  $\alpha$  at  $m$  that are consistent with this proposition, and  $State_\alpha^m(X) = K_3 \cup K_4$  therefore represents the proposition that one of these states obtains.

In the special case in which  $X = \{h\}$  is a maximally specific proposition, containing only a single history, we write  $State_\alpha^m/h$  and  $State_\alpha^m(h)$  for convenience; and here,  $State_\alpha^m/h$  is a unit set containing the unique state consistent with that history, and  $State_\alpha^m(h)$  is simply this unique state itself. Thus, for example, again in the case of Fig. 11.6, we have  $State_\alpha^m/h_2 = \{K_3\}$  and so  $State_\alpha^m(h_2) = K_3$ .

With these concepts before us, we can now, as in Horty (2001), define a form of act utilitarianism designed to model the orthodox notion found in the work of Gibbard, Sobel, Regan, and others.

ORTHODOX ACT UTILITARIANISM: Let  $\alpha$  be an agent and  $m$  a moment, and suppose  $K \in Choice_\alpha^m$ . Then the action  $K$  is *right* at the index  $m/h$  if and only if  $K \in Optimal_\alpha^m/State_\alpha^m(h)$ , and *wrong* otherwise.

What the definition tells us, then, is simply that the action  $K$  is right at the index  $m/h$  whenever  $K$  is optimal under the condition specified by the state containing the history  $h$ .

Returning to the Whiff and Poof example, let us consider, for example, the index  $m/h_2$ , where both Whiff and Poof push their buttons. At this index, the situation confronting Whiff, determined by Poof's action, is  $K_3$ ; that is,  $State_\alpha^m(h_2) = K_3$ . We therefore have  $Optimal_\alpha^m/State_\alpha^m(h_2) = Optimal_\alpha^m/K_3$ . And it is easy to verify also that  $Optimal_\alpha^m/K_3 = \{K_1\}$ , so that the action  $K_1$  is classified as right at  $m/h_2$ . In the same way, however, we can see that  $Optimal_\alpha^m(h_1) = \{K_2\}$ , so that the action  $K_1$  is classified as wrong at the index  $m/h_1$ .

As this example shows, the orthodox classification of actions as right or wrong – in contrast to the dominance account – depends on a full index, not just a moment.

---

<sup>8</sup> These definitions may seem to be needlessly general, but please bear with me; the generality will help us later on.

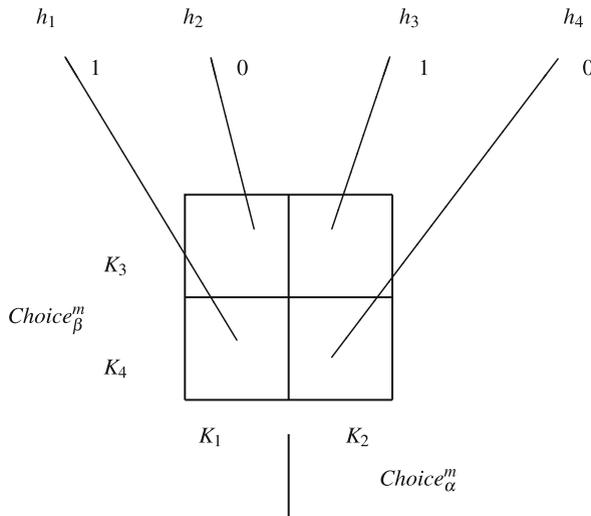
Here, the same action,  $K_1$ , is classified as right at the index  $m/h_2$  but wrong at the index  $m/h_1$ ; although Whiff performs the same action at each of these two indices, this agent satisfies orthodox act utilitarianism at the first, performing an action that is classified as right, but not at the second. It is as Regan says: we cannot define which of an agent’s actions are right or wrong until we know the circumstances under which the action is performed – that state confronting that agent, here defined as the actions simultaneously performed by the other agents involved.

## 11.5 The Perspectival Account

### 11.5.1 A Problem

At this point, we have before us two accounts of right action, dominance and orthodox. In order to compare these accounts, I now want to introduce yet another example, which I have found to be especially helpful in highlighting their differences.<sup>9</sup>

Imagine that two drivers are traveling toward each other on a one-lane road, with no time to stop or communicate, and with a single moment at which each must choose, independently, either to swerve or to continue along the road. There is only one direction in which the drivers might swerve, and so a collision can be avoided only if one of the drivers swerves and the other does not; if neither swerves, or both do, a collision occurs. This example is depicted in Fig. 11.7, where  $\alpha$  and  $\beta$  repre-



**Fig. 11.7** The driving example

<sup>9</sup> The example is due to Goldman (1976), but also discussed by Humberstone (1983), a paper that sets out in a different context some of the fundamental ideas underlying the orthodox account.

sent the two drivers,  $K_1$  and  $K_2$  represent the actions available to  $\alpha$  of swerving or continuing along the road,  $K_3$  and  $K_4$  likewise represent the swerving or continuing actions available to  $\beta$ , and  $m$  represents the moment at which  $\alpha$  and  $\beta$  must make their choice. The histories  $h_1$  and  $h_3$  are the ideal outcomes, resulting when one driver swerves and the other does not; collision is avoided. The histories  $h_2$  and  $h_4$ , resulting either when both drivers swerve or both continue along the road, are nonideal outcomes in which a collision occurs.

Now imagine that what actually happens is that both agents continue along the road, so that the resulting outcome is the history  $h_4$ , in which there is a collision. Suppose that, looking back at the situation from some later moment belonging to  $h_4$  – perhaps while recovering in the hospital – the agent  $\alpha$  says to itself: I performed the wrong action; it would have been right to swerve. And let us ask: is what the agent says correct, or not? The answer, I think, is that we can legitimately understand this statement either as correct or as incorrect, and that the contrast between these two different readings can be captured by appeal to our distinction between the orthodox and dominance accounts of right action.

On the one hand, it is clear from the standpoint of the later moment that, if the agent had swerved, there would have been no collision. Things would have gone much better for everyone had the agent swerved, and therefore, from a utilitarian point of view, the agent was wrong not to. This way of understanding the agent's statement is captured by the orthodox account, according to which  $Optimal_\alpha^m / State_\alpha^m(h_4) = \{K_1\}$ , so that the action  $K_1$  is classified as right and the action  $K_2$ , which the agent actually performed, as wrong at the index  $m/h_4$ . On the other hand, if we consider the situation from the standpoint of the earlier moment  $m$ , when the agent's action was actually performed, it is hard to see how we could have said at this moment that it would be right for the agent to swerve and wrong not to. Surely there is nothing in the situation as it appears at this moment – with the four histories each lying ahead as future possibilities – that could justify such a judgment. This way of evaluating the agent's statement is captured by the dominance account, according to which  $Optimal_\alpha^m = \{K_1, K_2\}$ , so that both actions are classified as right at the moment  $m$ .

The situation pictured in Fig. 11.7, then, seems to support two different evaluations of the agent's decision not to swerve – that it was wrong, or right – which can then be captured by our two theories of right and wrong actions, orthodox and dominance. This idea, originally set out in Horty (2001), of analyzing examples of this kind by appeal to two separate utilitarian theories carries some distinct advantages. It does not force us into the artificial position of classifying the agent's action either as unequivocally right or as unequivocally wrong, ignoring the pull of the opposite intuition. However, in allowing us this freedom, it also does not lead us into the muddled position of describing the action as somehow both right and wrong. What we can say, instead, is that the agent's action is right in one definite sense and wrong in another – that it is right in the dominance sense, but wrong in the orthodox sense.

Although I do not have space (or time) to justify this claim here, I believe that the contrast apparent in this example between the two different ways of evaluating the agent's action can be seen as underlying many of the debates in utilitarian theory

that commanded so much attention during the 1970 s and 1980 s. One of these, about which I will say nothing here at all, is the debate over the “actualist” and “possibilist” positions regarding the relations between an agent’s present obligations and future choices.<sup>10</sup> Another, about which I will say only a bit more in the appendix to this chapter, is the problem that occupied Gibbard, Sobel, Regan, and others concerning the relation between individual and group utilitarian theories.<sup>11</sup> The current proposal therefore has the real benefit of providing a rigorous explication of two different ways of understanding our normative evaluation that can be felt both in the example presented here and also, I believe, in other cases from the literature on utilitarian theory.

This benefit, however, comes with a cost. The cost is that the way in which the current proposal allows us to treat the same action as both right and wrong, respecting our conflicting intuitions, is by offering two different theories of right and wrong action. In effect, the proposal treats the words “right” and “wrong” as carrying two different senses, two different meanings. Of course, philosophy often proceeds like this, by discovering hidden ambiguities in items of ordinary language, which are then teased apart and provided with different formal explications. But in this case the idea simply seems wrongheaded. It is hard to think of these words as *semantically* ambiguous.

I now want to show that there is a better way. We can preserve the benefits of the account presented here, allowing appeal to both the orthodox and dominance perspectives in evaluating an agent’s actions, without postulating semantic ambiguity, by relying instead on a pragmatic difference.

### ***11.5.2 Perspectival Act Utilitarianism***

The basic idea is that an action performed by an agent at one moment is to be evaluated as right or wrong from the perspective of another moment, which may or may not be identical to the first. The key component of this idea – the appeal to “double time reference” – was first set out systematically by Belnap (2001), with an emphasis on the assessment of speech acts, particularly the speech act of assertion.<sup>12</sup> It was later developed in a somewhat different way by MacFarlane (2003) and elsewhere, who is concerned with the role of perspective in the assessment of a statement’s content: what is said, rather than the act of saying it.

---

<sup>10</sup> This problem was originally presented in a trio of papers: Goldman (1976), Sobel (1976), and Thomason (1981). Further discussion can be found, for example, in Bergström (1977), Carlson (1995), Feldman (1986), Goldman (1978), Greenspan (1978), Humberstone (1983), Jackson (1985), Jackson (1988), Jackson and Pargetter (1986), McKinsey (1979), and Zimmerman (1990).

<sup>11</sup> In addition to the work by Gibbard, Sobel, and Regan cited earlier, further discussion of this issue can be found in Carlson (1995), Feldman (1986), Jackson (1987), Jackson (1988), and of course Parfit (1984).

<sup>12</sup> Further discussion can be found at various points throughout Belnap et al. (2001) (see index entries under “double time reference”), and an informal presentation appears in Belnap (2004).

Let us take  $m$  as the *moment of action* and  $m'$  as the moment from which the action selected at  $m$  is evaluated – the *moment of evaluation*, which we can sensibly assume to be comparable to  $m$  in the treelike ordering of moments: either later than, earlier than, or identical with  $m$ . In that case,  $State_\alpha^m/H_{m'}$  – the set of states consistent with  $H_{m'}$ , the trivial proposition at  $m'$  – can be taken to represent the states confronting the agent at  $m$ , as judged from the standpoint of  $m'$ . As we have seen,  $State_\alpha^m(H_{m'})$  is simply the proposition that one of these states holds. And so the set  $Optimal_\alpha^m/State_\alpha^m(H_{m'})$  contains those actions available to the agent at  $m$  that are optimal under the conditions in which the agent finds itself, where these conditions are themselves judged from the standpoint of  $m'$ .

Using these ideas, we can therefore define *perspectival act utilitarianism* as the theory according to which an action available to the agent  $\alpha$  at the moment  $m$  is right from the standpoint of the moment  $m'$  just in case that action is optimal given the states that the agent is confronting at  $m$ , as judged from the standpoint of  $m'$ .

PERSPECTIVAL ACT UTILITARIANISM: Let  $\alpha$  be an agent and  $m$  and  $m'$  moments such that either  $m < m'$  or  $m' < m$  or  $m = m'$ , and suppose  $K \in Choice_\alpha^m$ . Then the action  $K$  is *right* at  $m$  from the standpoint of  $m'$  if and only if  $K \in Optimal_\alpha^m/State_\alpha^m(H_{m'})$ , and *wrong* otherwise.

This perspectival account allows us to capture the intuitions underlying the orthodox approach, as we can see by returning to the driving example. Suppose, again, that neither driver swerves, the crash occurs, and we are considering the incident from the standpoint of some later moment – call it  $m_1$  – lying on the history  $h_4$ . Since  $m_1$  lies on the history  $h_4$  at some time later than  $m$ , and  $h_4$  itself belongs to the state  $K_4$ , it follows that each history from  $H_{m_1}$ , the set of histories passing through  $m_1$ , must likewise belong to  $K_4$ .<sup>13</sup> From this it follows that  $K_4$  is the only state confronting  $\alpha$  at  $m$  that is consistent with  $H_{m_1}$  – that is,  $State_\alpha^m/H_{m_1} = \{K_4\}$ ; the set of states confronting the agent at  $m$ , as judged from the standpoint of  $m_1$ , contains  $K_4$  alone. From this it follows that  $State_\alpha^m(H_{m_1}) = K_4$ . We therefore have

$$\begin{aligned} Optimal_\alpha^m/State_\alpha^m(H_{m_1}) &= Optimal_\alpha^m/K_4 \\ &= \{K_1\}, \end{aligned}$$

so that, from the standpoint of  $m_1$ , we reach the orthodox judgment that the action  $K_2$  chosen by the agent was wrong and  $K_1$  would have been right, optimal under the circumstances in which the agent found itself.

On the other hand, suppose that, at the crucial moment, both drivers swerve, another crash occurs, things proceed along the history  $h_2$ , and that we are now reflecting on the incident from some later moment – say  $m_2$  – lying on that history. Parallel reasoning thus gives us  $State_\alpha^m(H_{m_2}) = K_3$ , from which we can conclude, just as before, that  $Optimal_\alpha^m/State_\alpha^m(H_{m_2}) = \{K_2\}$ . From the standpoint of  $m_2$ ,

---

<sup>13</sup> Although this point is “visually obvious,” it actually relies on the technical constraint of “no choice between undivided histories,” not discussed in this chapter, according to which histories that are still undivided at a given moment cannot be separated at that moment by the *Choice* partition.

then, we conclude that the action  $K_1$  was wrong and  $K_2$  would have been right. A different point of evaluation leads to a different result.

The perspectival approach, then, allows us to recover the intuitions underlying orthodox act utilitarianism, but interestingly, it subsumes the dominance account as well. This can be seen to hold quite generally. Suppose that the actions available to an agent at the moment  $m$  are evaluated from the standpoint of a moment  $m'$  that is either identical with or earlier than the moment  $m$  itself:  $m' = m$  or  $m' < m$ . Then it is easy to see that each member of  $State_\alpha^m$  contains some history from  $H_{m'}$ , so that the set of states confronting the agent at  $m$ , judged from the standpoint of  $m'$ , is simply  $State_\alpha^m$  itself:  $State_\alpha^m / H_{m'} = State_\alpha^m$ . From this it follows, since  $State_\alpha^m$  partitions the set  $H_m$ , that  $State_\alpha^m(H_{m'}) = H_m$ . As noted earlier, the set  $Optimal_\alpha^m / H_m$ , containing those actions available to  $\alpha$  at  $m$  that are optimal under the conditions specified by the trivial proposition, coincides with the set  $Optimal_\alpha^m$  itself. It therefore follows that

$$\begin{aligned} Optimal_\alpha^m / State_\alpha^m(H_{m'}) &= Optimal_\alpha^m / H_m \\ &= Optimal_\alpha^m, \end{aligned}$$

so that the set of actions available at  $m$  that are right from the standpoint of  $m'$  coincides with the set of actions available at  $m$  that are right according to the dominance account.

This general point can be illustrated with our driving example, Fig. 11.7, if we suppose that  $m'$  is some moment of evaluation identical with or earlier than the moment  $m$  of action. In that case, we have

$$\begin{aligned} Optimal_\alpha^m / State_\alpha^m(H_{m'}) &= Optimal_\alpha^m / H_m \\ &= Optimal_\alpha^m \\ &= \{K_1, K_2\}. \end{aligned}$$

Taking such a moment  $m'$  as our moment of evaluation, then, we arrive at the dominance intuition that either action available to the agent at  $m$  is right.

## 11.6 Conclusion

The theory of perspectival act utilitarianism set out here allows us see how we can say, in the driving example, for instance, that the agent's actions at the crucial moment might legitimately be viewed as both right and wrong. The theory thus preserves the advantages of my earlier account, from Horty (2001), by allowing us to respect our conflicting intuitions in cases like this. But it does not do so by postulating two separate senses of the words "right" and "wrong" – an orthodox and a dominance sense – captured by two separate utilitarian theories. These words can now be taken as semantically unambiguous.

When we say that an agent's action is right, from the standpoint of some moment of evaluation, we always mean exactly the same thing: the action is optimal under the conditions in which the agent finds itself at the moment of action, where these conditions are themselves judged from the standpoint of the moment of evaluation. Our conflicting intuitions about right and wrong can now be provided with a pragmatic, rather than a semantic, explanation, shifting with the relation between moment of action and moment of evaluation, and reflecting different evaluative judgments about the conditions confronting the agent at the moment of action. If the moment of evaluation is strictly later than the moment of action, then the perspectival theory agrees with orthodox act utilitarianism. But if the evaluation takes place at the very moment of action, or earlier, the perspectival theory agrees with dominance act utilitarianism. The difference between our orthodox and dominance intuitions is not, therefore, a substantial difference that needs to be explained by postulating two separate utilitarian theories, but only a matter of perspective.

## Appendix: Act Utilitarianism for Groups

This appendix shows how perspectival act utilitarianism can be extended from individual actions to group actions, and how the relation between the right actions available to groups and individuals can then be seen to depend on the standpoint from which these actions are evaluated.

The extension of perspectival act utilitarianism to group actions is straightforward, involving nothing more than a generalization of several of our previous notions. We have already seen, in the text, how the set  $Choice_{\Gamma}^m$  of actions available to the group  $\Gamma$  at the moment  $m$  can be defined, with each group action identified as a pattern of actions available to the individuals from that group. The states confronting the group  $\Gamma$  at  $m$  can then be defined as the patterns of actions available at  $m$  to all agents except those from that particular group:

$$State_{\Gamma}^m = Choice_{Agent-\Gamma}^m$$

And where  $X$  is some proposition, weak and strong dominance relations under the condition  $X$  can be defined among the actions available to a group in a way exactly parallel to the definition for individual actions.

**CONDITIONAL DOMINANCE ORDERING ON GROUP ACTIONS:** Let  $\Gamma$  be a group of agents and  $m$  a moment, and let  $K$  and  $K'$  be members of  $Choice_{\Gamma}^m$ , and  $X$  a proposition. Then  $K \preceq_X K'$  ( $K'$  weakly dominates  $K$  under the condition  $X$ ) if and only if  $K \cap X \cap S \leq K' \cap X \cap S$  for each state  $S \in State_{\Gamma}^m$ ; and  $K \prec_X K'$  ( $K'$  strongly dominates  $K$  under the condition  $X$ ) if and only if  $K \preceq_X K'$  and it is not the case that  $K' \preceq_X K$ .

The set of actions available to the group  $\Gamma$  under the condition  $X$  can be defined as those among the available actions that are consistent with this condition:

$$Choice_{\Gamma}^m / X = \{K \in Choice_{\Gamma}^m : K \cap X \neq \emptyset\}.$$

And the optimal actions available to the group under this condition can then be defined as the actions available under this condition that are not dominated under this condition by any other such actions:

$$\text{Optimal}_{\Gamma}^m/X = \{K \in \text{Choice}_{\Gamma}^m/X : \neg \exists K' \in \text{Choice}_{\Gamma}^m/X (K \prec_X K')\}.$$

Finally, the set of states confronting  $\Gamma$  at  $m$  that are consistent with the proposition  $X$  can be represented just as before:

$$\text{State}_{\Gamma}^m/X = \{K \in \text{State}_{\Gamma}^m : K \cap X \neq \emptyset\}.$$
<sup>14</sup>

And likewise the proposition that one of these states holds:

$$\text{State}_{\Gamma}^m(X) = \bigcup \text{State}_{\Gamma}^m/X.$$

Given these materials, we can now introduce a form a perspectival act utilitarianism for groups, according to which an action available to the group  $\Gamma$  at a moment  $m$  is right from the standpoint of the moment  $m'$  just in case that action is optimal under the conditions in which the group  $\Gamma$  finds itself at  $m$ , where these conditions are judged from the standpoint of  $m'$ :

PERSPECTIVAL ACT UTILITARIANISM FOR GROUPS: Let  $\Gamma$  be a group of agents and  $m$  and  $m'$  moments such that either  $m < m'$  or  $m' < m$  or  $m = m'$ , and suppose  $K \in \text{Choice}_{\Gamma}^m$ . Then the action  $K$  is *right* at  $m$  from the standpoint of  $m'$  if and only if  $K \in \text{Optimal}_{\Gamma}^m/\text{State}_{\Gamma}^m(H_{m'})$ , and *wrong* otherwise.

As with individual actions, this perspectival account supports the orthodox intuitions concerning group actions when the moment  $m'$  of evaluation is later than the moment  $m$  of action, while the dominance intuitions are supported when  $m'$  is earlier than or identical with  $m$ .

Now that the perspectival account has been extended from individuals to groups, let us turn briefly to two of the most central questions concerning the relation between individual and group act utilitarianism. First, if each individual belonging to a group performs a right action, does that entail that the group itself performs a right action? And second, if a group performs a right action, does that entail that the individuals belonging to the group do so?

The answer to the first question is No. This fact is well-known and can be illustrated with the Whiff and Poof example from Fig. 11.6, which was originally formulated to make exactly this point. Still, it is useful to consider the question separately from the dominance and orthodox perspectives, since the contours of this negative answer differ.

<sup>14</sup> In the group case, this fact actually follows from the previous definitions of  $\text{State}_{\Gamma}^m$  as the set of states confronting  $\Gamma$  at  $m$  and  $\text{Choice}_{\Gamma}^m/X$  as the actions available to  $\Gamma$  that are consistent with  $X$ . However, it is set out separately here in order to conform to our treatment of the individual case, where the corresponding notion must be introduced through a definition.

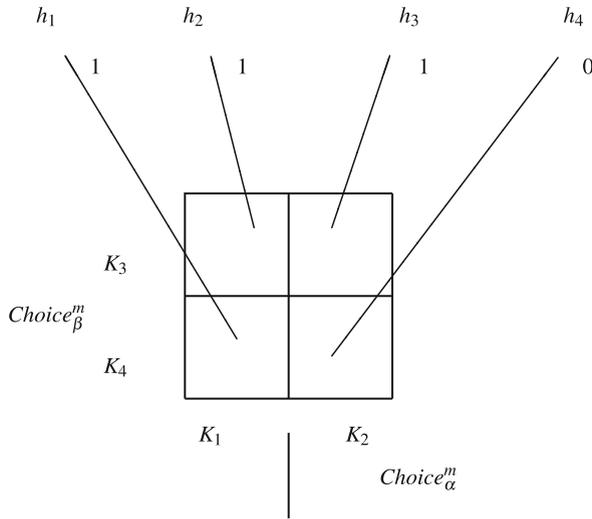
Suppose, first, that we evaluate the actions available at the moment  $m$  in Fig. 11.6 from the standpoint of  $m$  itself, adopting the dominance perspective. Then it is easy to verify that each action available to either agent is classified as right from the standpoint of  $m$ . So suppose that Whiff pushes his button, performing the action  $K_1$ , while Poof refrains, performing  $K_4$  – each agent therefore performing an action that is classified as right. Then the group  $\Gamma = \{\alpha, \beta\}$  containing both Whiff and Poof performs the action  $K_1 \cap K_4$ , which is clearly non-optimal, leading to a utility of 0 while 10 is possible, and so classified as wrong from the standpoint of  $m$ . Indeed the group action  $K_1 \cap K_4$  is not even in equilibrium: each agent would be better off with a different choice, given the action chosen by the other. Individual satisfaction of dominance act utilitarianism, then, not only fails to guarantee group satisfaction, but has the even more depressing consequence that the pattern of actions chosen, each right from an individual perspective, may not be an equilibrium pattern.

Next, suppose Whiff and Poof both refrain from pushing their buttons, performing the individual actions  $K_2$  and  $K_4$ . The outcome of this pair of actions is the history  $h_4$ . So let us evaluate these actions from the standpoint of some later moment along this history, thus adopting the orthodox perspective. It is easy to see that both of these actions are then classified as right from the standpoint of this later moment, and also that the pair of actions is in equilibrium: each agent is performing a best available action, given the actions performed by the other.

This example illustrates the general rule: whenever each individual member of a group of agents performs an action that is right from the standpoint of a later moment – and so right from the orthodox perspective – the pattern of actions performed by the entire group is in equilibrium. However, this does not mean that the group action is itself right. In this case, the group action  $K_2 \cap K_4$  is non-optimal, and so wrong, since it yields a utility of 6 while the available group action  $K_1 \cap K_3$  yields a utility of 10. If each member of a group performs an action that is right from the standpoint of a later moment, then, the overall pattern of actions will be in equilibrium, but it still may not be a right action for that group to perform, since there may be better equilibrium patterns.

Now to the second question: if a group action is right, does it follow from this that the actions of the individuals belonging to that group are also right? The standard answer to this question is Yes. Regan, for example, writes that “for any group of agents in any situation, any pattern of behaviour by that group of agents in that situation which produces the best consequences possible is a pattern in which the members of the group all satisfy AU” (Regan 1980, p. 54). And Jackson, that “if the right group action is actually performed, then that group action’s constituent individual actions must be right” (Jackson 1988, p. 264). In the case of this question, however, the dominance and orthodox perspectives yield different answers.

Both Regan and Jackson adopt the orthodox perspective, evaluating actions from the standpoint of a later moment, and from that perspective what they say is right. In our current language, it can be put like this: if a group action performed at  $m$  is right from the standpoint of a later moment  $m'$ , then the actions performed by the individual members of that group are also right from the standpoint of  $m'$ .



**Fig. 11.8** Group action right, individual action wrong

However, the implication fails if we consider the matter from the dominance perspective, evaluating actions from the standpoint of a moment at or before the moment of their performance: where  $m'$  is identical with or earlier than  $m$ , it might well be possible that a group action performed at  $m$  is right from the standpoint of  $m'$ , while the individual action of some member of that group is wrong from the standpoint of  $m'$ . This possibility is illustrated in Fig. 11.8. Here, it is easy to see that the action  $K_2 \cap K_3$  performed at the moment  $m$  by the group  $\Gamma = \{\alpha, \beta\}$  is right from the standpoint of the moment  $m$  itself, since this group action leads to an outcome of utility 1, the highest available, and is therefore optimal. But the component action  $K_2$  by the agent  $\alpha$  is wrong from the standpoint of  $m$ , since it is dominated by  $K_1$ . Of course, from the standpoint of some future moment along the history  $h_3$ , we can see the action  $K_2$  by  $\alpha$  was performed under circumstances in which  $\beta$  performed the action  $K_3$ , so that an outcome of utility 1 was achieved; from this later standpoint, the action  $K_2$  is therefore right. But at the moment  $m$  itself, while it is still unclear which action  $\beta$  will perform, the choice of  $K_2$  allows for an outcome of utility 0, and is therefore dominated by  $K_1$ , which guarantees an outcome of utility of 1.

## References

Belnap N (2001) Double time references: speech act reports as modalities in an indeterministic setting. In: Wolter F, Wansing H, de Rijke M, Zakharyashev M (eds) *Advances in modal logic*, vol 3, CSLI Publications, Stanford pp 1–21

Belnap N (2004) Future contingents and the sea battle tomorrow, manuscript, Philosophy Department, University of Pittsburgh

- Belnap N, Perloff M, Xu M (2001) *Facing the future: Agents and choices in our indeterminist world*. Oxford University Press, Oxford
- Bergström L (1966) The alternatives and consequences of actions, *Stockholm Studies in Philosophy*, vol 4. Almqvist and Wiksell, Stockholm
- Bergström L (1977) Utilitarianism and future mistakes. *Theoria* 43:84–102
- Carlson E (1995) *Consequentialism reconsidered, theory and decision library, series A: Philosophy and methodology of the social sciences*, vol 20. Kluwer Academic Publishers, Dordrecht
- Ellsberg D (1961) Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* 75:643–669
- Feldman F (1986) *Doing the best we can: An essay in informal deontic logic*. D. Reidel Publishing Company, Dordrecht
- Gärdenfors P, Sahlin NE (eds) (1988) *Decision, probability, and utility: Selected readings*. Cambridge University Press, Cambridge
- Gibbard A (1965) Rule-utilitarianism: merely an illusory alternative? *Australasian Journal of Philosophy* 43:211–220
- Goldman H (1976) Dated rightness and moral imperfection. *The Philosophical Review* 85:449–487
- Goldman H (1978) Doing the best one can. In: Goldman AI, Kim J (eds) *Values and morals*, D. Reidel Publishing Company, Dordrecht, pp 185–214
- Greenspan P (1978) Oughts and determinism: a response to Goldman. *Philosophical Review* pp 77–83
- Horty J (2001) *Agency and deontic logic*. Oxford University Press, Oxford
- Humberstone IL (1983) The background of circumstances. *Pacific Philosophical Quarterly* 64:19–34
- Jackson F (1985) On the semantics and logic of obligation. *Mind* 94:177–195
- Jackson F (1987) Group morality. In: Pettit P, Sylvan R, Norman J (eds) *Metaphysics and morality: Essays in honour of J. J.C. Smart*, Basil Blackwell Inc., Oxford pp 91–110
- Jackson F (1988) Understanding the logic of obligation. In: *Proceedings of the Aristotelian Society, Supplementary Volume 62*, Harrison and Sons
- Jackson F, Pargetter R (1986) Oughts, options, and actualism. *Philosophical Review* 99:233–255
- Luce RD, Raiffa H (1957) *Games and decisions*. John Wiley and Sons, New York
- MacFarlane J (2003) Future contingents and relative truth. *Philosophical Quarterly* 53:321–336
- McKinsey M (1979) Levels of obligation. *Philosophical Studies* 35:385–395
- Parfit D (1984) *Reasons and persons*. Oxford University Press, Oxford
- Prior A (1956) The consequences of actions. In: *Proceedings of the Aristotelian Society, Supplementary Volume 30*, Harrison and Sons
- Prior A (1967) *Past, present, and future*. Clarendon Press, Oxford
- Ramsey F (1931) Truth and probability. In: Braithwaite RB (ed) *The foundations of mathematics and other logical essays*, Routledge and Kegan Paul, Oxford pp 156–191, originally published in 1926
- Regan D (1980) *Utilitarianism and co-operation*. Clarendon Press, Oxford
- Savage L (1951) The theory of statistical decision. *Journal of the American Statistics Association* 46:55–67
- Savage L (1954) *The foundations of statistics*. John Wiley and Sons, New York second revised edition published by Dover Publications, 1972
- Sobel JH (1968) Rule-utilitarianism. *Australasian Journal of Philosophy* 46:146–165
- Sobel JH (1976) Utilitarianism and past and future mistakes. *Nous* 10:195–219
- Thomason R (1970) Indeterminist time and truth-value gaps. *Theoria* 36:264–281
- Thomason R (1981) Deontic logic and the role of freedom in moral deliberation. In: Hilpinen R (ed) *New Studies in Deontic Logic*, D. Reidel Publishing Company, Dordrecht pp 177–186
- Zimmerman M (1990) Where did I go wrong? *Philosophical Studies* 59:55–77