Agency and Deontic Logic, by John F. Horty. Oxford: Oxford University Press, 2001. Pp. xi + 192. H/b £27.00/$35.00.

Horty develops a new precise utilitarian (really consequentualist) account of agential oughts, sentences saying what agents ought to do. He does this without relying on expected value theory, while building upon a growing indeterminist formal research program where future contingent propositions, including propositions about agents' actions, are neither true nor false at a moment. This unexplored combination of constraints provides a sound rationale for the book, in which a novel indeterminist value-theoretic framework is used to explore precise versions of act and rule utilitarianism. Since there are lots of links to ethical literature, and formalities are toned down (most proofs are in an appendix), this is a work in formal philosophy, rather than pure deontic logic, one reminiscent of F. Feldman's *Doing the Best We Can* (D, Reidel Publishing Co., 1986). I will concentrate on chapters 2-4 of Horty's book, which are foundational.

Ch. 2 introduces a formal indeterministic temporal framework stemming from A. Prior, championed by R. Thomason and N. Belnap, and now the basis of a robust research program, to which Horty has previously contributed. (See N. Belnap *et al*, *Facing the Future: Agents and Choices in Our Indeterminist World*, Oxford: Oxford University Press, 2001.) A set of 'moments', *Tree*, and a two-place ordering relation defined on it, *after*, are primitives. Roughly, moments may be thought of as possible world slices (not seconds), where m' is (possibly) after m iff m' is some still possible future moment of m. Moments can branch forward, toward the future, not backward. A history (world), h, is construed as a maximal path or branch on a tree. In models with two or more histories, some moment is not comparable to another, and some moment is common to distinct histories. Contingent future tensed statements, and thus all statements

for uniformity, are assigned truth values at a moment-history *pair, m-h.* (I will sometimes ignore histories in formulations where uniformity is the only reason to mention them.) The only case where we can say a statement will be true at *m simpliciter* is where its future truth is historically necessary: *p* is *historically necessary* at *m* iff *p* holds at all histories passing through *m*. *A history passes through m* when *m* is in that history.

Next, a set of *Agents*, and a *Choice* function are introduced. *Choice* partitions the histories passing through a moment for an agent. Each cell is taken to represent a possible complete *action* ('choice' is also used) for that agent at that moment. Actions are essentially instantaneous constraints placed on the future. No intentions are represented. Where there is more than one cell, no particular action is determined at that moment. For each history within an available action, that action is *the action performed by the agent at that moment in that history*. An agent then *c-sees to it that (c-stit) p* at *m-h* iff the action the agent performs in *h* at *m* guarantees *p* (contains only *p*-histories). An agent *d-sees to it that (d-stit) p* at *m-h* iff the agent c-stit *p* at *m-h* and it is not historically necessary that *p* at *m*. The latter is the 'deliberative stit' because the negative condition is to assure that there is a real choice—the agent might have chosen in a way that left *p*'s falsity open. An agent is *able to c-see (d-see) to it that p* at a *m-h* iff it is historically possible that the agent c-stit (d-stit) *p* at *m-h*. In a pattern nicely employed throughout the book, Horty generalizes the account to cover groups (as sets of individuals). Since the two notions of agency are interdefinable in the underlying framework, Horty concentrates on the simpler c-stit version. I henceforth intend c-stit by 'seeing to it that' unless otherwise indicated.

In Ch. 3, Horty introduces a general value-theoretic framework inspired by van Fraasen, but he quickly zeroes in on more specific models intended to represent

utilitarianism. Roughly, histories have a rank-reflecting numerical value that does not vary from moment to moment (so histories are all mutually comparable with respect to value). Here, I will assume we always have best members, since this gives the flavor of Horty's analyses without the complications. Horty now provides his analysis of impersonal oughts: *it ought to be that p* holds at *m*-*h* iff *p* holds at the best histories passing through *m*. All the principles of Standard Deontic Logic (SDL), including no ought-conflicts, follow. A non-agential version of Kant's Law (it ought to be that *p* only if *p* is historically possible) also follows.

Horty next turns to the central question of the book: How do we represent agential oughts? Horty examines the widely discussed 'Meinong-Chisholm reduction' (M-C reduction), that what an agent ought to do is what it ought to be that she does. Recast in the stit framework, *an agent ought to see to p* iff it ought to be the case that the agent sees to it that *p*. Given the previously proposed semantics, an agent ought to see to *p* holds at *m*-*h* iff that agent chooses a p-guaranteeing action at the best histories passing through *m*. This analysis again yields a normal modal operator satisfying the principles of SDL (but not if d-stit is used instead of c-stit). An agential version of Kant's Law follows (an agent ought to see to *p* only if she is able to). It also follows that what an agent ought to do, ought to be.

With this precise representation of the M-C reduction, Horty turns to objections by P. Geach and G. Harman to the MC-reduction. Essentially, Geach's objection is:

> Necessarily, Fred dances with Ginger iff Ginger dance with Fred. But then it ought to be that Fred dances with Ginger iff it ought to be that Ginger dances with Fred, so the M-C reduction wrongly implies that Fred ought to dance with Ginger iff Ginger ought to dance with Fred as well.

Horty argues that this objection (like the one by Harman) fails because it turns on an oversight that his explicit agential resources help reveal. Horty asks us to distinguish three claims: that Fred and Ginger dance together (*p*), that Fred sees to it that *p*, and that Ginger sees to it that *p*. Necessarily, Fred and Ginger dance together iff Ginger and Fred dance together, but it does not follow from that that necessarily, it ought to be the case that *Fred sees to p* iff to it ought to be the case that *Ginger sees to p*.

Horty next introduces his 'Gambling Problem' (cf. Feldman, ibid, pp.194-5). Suppose I have two actions available to me, gamble $5 or not. Now suppose that if I gamble and win, I get $10; and if I gamble and lose, I get $0. Suppose the only values at stake are the dollar values, and thus the value of not gambling is $5 saved. Ignore probabilities. Since I cannot determine whether or not I win, it is not true that what I ought to do is gamble. But in the best histories, I win, and my gambling is a necessary condition of my winning, so the M-C reduction implies that I ought to gamble. Horty takes this to decisively defeat the M-C reduction. Horty does not turn to an expected utility approach to agential oughts because he believes that probabilities about outcomes, especially those involving agent's choices, are often unavailable or meaningless. In decision-theoretic terms, Horty aims at an account for conditions of 'uncertainty', not 'risk'.

In Ch. 4, Horty develops dominance act utilitarianism. The main task is to use the values of histories to identify a dominance ordering over actions that can be used to generate a utilitarian account of agential oughts. This is done by first defining an ordering over (moment-relative) propositions (any subset of histories passing through a moment), then a distinct one over actions. Where two propositions, *p* and *q*, are relative to the same

moment, *p is as good as q* iff the value of each history making up *p* is ranked as high as any making up *q*. *p is better than q* iff *p* is as good as *q* but not vice versa, so that we are sure to do at least as well, and maybe better, with *p* (cf. {1,0} with {0,0}). Since my actions at a moment are themselves taken as subsets of histories through that moment, the simplest proposal for a utilitarian ranking of actions would be the one just arrived at for propositions generally. However, Horty shows that two actions incomparable on this simple proposal might still be comparable. He shows this using 'sure thing reasoning': if A is better than B in each agent-independent state, then A is better than B period. Since the only source of agent-independent states that are stipulated in the models are others' possible choices, Horty defines the *states* for an agent at *m* as the set of joint actions then available to the group consisting of all other agents. In such a context, *one action, A, available to me at m weakly dominates another, B,* iff the combination (intersection) of *A* with any collective pattern of action available to the other agents is as good as the combination of *B* with that pattern; *A strongly dominates B* iff *A* weakly dominates *B*, but not vice versa. The *optimal actions* for an agent at *m* are then defined as those actions in the agent's ability at *m* that are not dominated by any other such actions. *Dominance Act Utilitarianism (DAU)* is thus the view that your *right actions* at *m* are your optimal ones. In turn, *an agent ought to see to it that p* at *m* iff each of her right actions at *m* guarantees *p*. This account yields an agential version of Kant's Law and entails a logic containing SDL. Even conflicting oughts between agents are excluded. Horty also identifies relationships between the various oughts.

This novel orientation is then elaborated upon in the remaining chapters. In Ch. 5 Horty generalizes the account to cover conditionally right acts and conditional oughts,

showing that a widely accepted version of sure thing reasoning with conditional oughts isn't universally valid. Next, in contrast to DAU, he defines 'Orthodox Act Utilitarianism' (OAU). He then examines a challenging moral luck case discussed in the literature by H. Goldman and I. Humberstone, where no agent choice is unambiguously right, since that appears to depend in all cases on the independent choices of others. Horty's distinction between DAU and OAU seems to track, and perhaps explain, our vacillation about what's right in such cases. In Ch. 6, Horty generalizes DAU and OAU to cover groups, and explores links between individual and group satisfaction of each. He examines the question raised by R. Brandt, explored by A. Gibbard and J. H. Soble, and made famous by D. Regan: Is satisfaction of utilitarianism by each member of a group sufficient for satisfaction of utilitarianism by the group (or vice versa)? Horty also explores various issues of upward and downward inheritance discussed by D. Parfit, T. Tannsjo, and others (e.g. if a group ought to see to *p* and can do so only if some member sees to *q*, does it follow that that member ought to see to *q*). Next, Horty defines a version of *Rule Utilitarianism* (cf. D. Regan's 'coordinated optimization theory') and explores its relations to DAU and OAU, showing that certain associated arguments of D. Regan and F. Jackson are sound for OAU, but not for DAU. As cast in his framework, Horty settles a number of such issues in these two chapters. Finally, in Ch.7, Horty gives a thumbnail sketch of a generalization of DAU, one involving temporally extended sequences of actions and sensitive to the current normative impact of actions available to an agent in future possible moments. He closes with an associated brief discussion of utilitarian possibilism and actualism.

The widely employed framework for indeterminism and agency is not defended in Horty's book. The best source for that is *Facing the Future* (ibid.), which incorporates earlier work by Horty and others. However, since an assessment of the significance of this project derives in part from the underlying framework for indeterminism and agency, I comment on this first. Firstly, it is fundamental to this approach that future contingents are neither true nor false at a moment. This still controversial view is defended in *Facing the Future*. Similarly for the importance of the "sees to it that" idiom for modeling agency. Regarding actions, there is no analysis of instantaneous actions that are not maximal (e.g. calling *someone*). Some might take the notion of an instantaneous action as both the lower limit of action length, and a contradiction in terms. Perhaps 'choice' is safer in this regard. There is nothing here like the analysis of actions as transformations of sets of moments to other sets of moments, as in dynamic logic (K. Segerberg, 'Getting Started: Beginnings in the Logic of Action,' *Studia Logica* **51**, 1992, pp. 347-378). However, the branching time framework has hardly been exhausted in this regard (M. Xu, 'Causation in Branching Time (I): Transitions, Events, and Causes,' *Synthese* **112**, 1997, pp. 137-192).

There are challenges facing both versions of 'seeing to it that'. In the case of c-stit, an agent sees to all necessary truths. D-stit is intended to get around this, but first of all, it seems to involve a potential conflation in its conception as a 'deliberative stit'. Recalling Locke's prisoner, in order for an agent to deliberatively choose, perhaps she must believe there is an option; but it is quite contentious that there must actually be an option. Here, seeing to p requires that it still might be that ~p, thus making agency depend *logically* on the falsity of compatibilism. (See D. Elgesem, 'The Modal Logic of

Agency', *Nordic Journal of Philosophical Logic* **2**(2), 1997, pp. 1-46.) Furthermore, it seems too easy here to undermine genuine agency. (This holds for the 'achievement stit' in *Facing the Future*, as well as the c-stit and d-stit operators used here and there.) Suppose I pull the trigger of a gun aiming at you, and you are hit by the bullet. Now add that when I pulled the trigger a random gust of wind could have occurred and knocked the bullet off target, though it didn't occur. On the current analyses, it follows that I did not see to it that you were hit, because no choice I made *guaranteed* that. The mere fact that the wind *could have* interfered with the course of the bullet is enough to undermine the claim that I was the agent of your being hit. This smacks of getting away with murder.

This may be related to a problem with Horty's dancing case. It takes two to Tango. So Fred needs Ginger's cooperation for *p* (that they dance together), since Ginger has the ability to see to it that *p* fails. But then on the current analysis, how can Fred see to it that *p*, since that requires an action that *guarantees p*? The formal model Horty uses to argue that Geach's objection to the M-C reduction is flawed reflects this oddity as well, since in it, neither party has the ability to see to it that s/he *doesn't* dance (p.52). However, short of irresistible charm, can't I bring it about that we dance together by asking you?

Consider Horty's analysis of impersonal oughts. It is now settled that some children will die of starvation, but that ought to not be the case. It seems false that whatever ought to be still could be. Talk of what would be ideal is not so constrained. It is not even clear why a utilitarian analysis of what ought to be (as opposed to what an agent ought to do) must meet this constraint. Can't a utilitarian consistently use her value

theory to say at once that we ought to cause Jones pain (to avoid greater pain, where some is inevitable) even though it ought to be the case that no one suffers, since suffering is intrinsically bad? Also, Horty's analysis of impersonal oughts ratifies the principle that *what is necessitated by what ought to be the case itself ought to be the case*. It is unfortunate that Horty does not discuss this controversial principle. It has been widely questioned in response to various paradoxes of deontic logic. If this principle in fact fails for 'ought to be', then Horty's pivotal use of the gambling problem to overthrow the M-C reduction of agential oughts, and motivate his own non-reductive analysis, is problematic.

Turning to his dominance ranking of propositions, note that propositions will be incomparable if each contains a history whose value outranks a history in the other (cf. {1,0} with {1,0}). Horty notes that a lot of familiar properties for ordering relations follow, but he doesn't note that no moment-relative proposition will be as good as itself unless it is made up of histories all exactly equal in value. Similarly, the derivative ranking relation for actions (in terms of independent states) is non-reflexive. However, does the indeterminism of some features of the future entail that some propositions and actions are not as good as themselves?

Horty gives central place to the decision-theoretic strategy of dominance to handle conditions of uncertainty in practical decision making. But dominance utilitarianism is too weak to account for our actual decision making in ordinary situations of uncertainty. Suppose I am obligated to be somewhere and can get there only by driving. Now consider the fact that at numerous intersections that I drive through, someone *could* choose to run the lights or stop signs, and thereby kill me. These are choices other drivers could make, thus nothing I do can rule out this sort of possible

history other than to not drive to my appointment. (Also note that I do not have knowledge of numerical probabilities here.) So in ordinary cases of this sort, driving to my appointment will not dominate staying home, so I will not be obligated to see to it that I keep the appointment. But many of my obligations are subject to this sort of objection. Given the independence of other's choices, coupled with their ability to undermine my efforts, their will be far too few dominance oughts. This seems to raise questions about the power of this sort of approach to represent moral reasoning in conditions of uncertainty, other than the extreme where I have no idea what others will do. Presumably, in ordinary conditions of uncertainty, we are routinely relying on either more than dominance or something different, yet without falling back on expected value theory. This is not to say that Horty is not right about what would be reasonable if all we had to go on is what he has in his dominance models. I think his contribution there is secured. The trouble is that since we appear to be relying on either something more, or on something else altogether, until we uncover what that is, we will not know whether it builds on Horty's foundation, or sidesteps it, nor will we know if there is a utilitarian reconstruction of whatever we are in fact relying on. However, this really reflects a general challenge facing decision theory.

This fine book is resource-rich and thought-provoking. I recommend it to all analytic philosophers interested in exploring the implications of a robust indeterminism for consequentialism. It is must reading for those who work on formal approaches to normative reasoning.

Department of Philosophy
University of New Hampshire
Durham, NH 03824-3574 USA
paulm@cisunix.unh.edu

PAUL MCNAMARA