

## 6

## Deciding How to Decide: Is There a Regress Problem?

Holly Smith

Rational deliberation is itself an activity like any other, and the extent to which one should engage in it is subject to rational decision.

John Rawls, *A Theory of Justice*

A decision guide is a principle – such as “Maximize expected utility” – that evaluates actions and can be used by agents in selecting which action to perform. Most decision guides are designed to be usable in choosing actions even though the agent has incomplete knowledge about her options. Even so, an agent often wants or needs to acquire fuller information before making her choice. Thus choices are often preceded by activities in which the agent seeks, processes, and assesses relevant information. These activities, too, may be subject to choice. In deciding whether to invest her money in stocks or a money market account, should an agent consult government economic forecasters or the newsletter of her brokerage firm? It may be that the very same principle that guides an agent’s choice of her target action can also be used to guide her choice of the kind of information-gathering that she should employ beforehand: consulting forecasts or newsletters, like investing in stocks or money markets, can be evaluated in terms of expected utilities.

Information-gathering includes a wide variety of processes, which I shall divide into two rough categories, *deliberation* and *research*. By “deliberation” I shall mean all those processes aimed at revealing to the agent, for purposes of making a decision, information that she already implicitly possesses. Deliberation includes recalling facts stored in memory, deducing implications of one’s beliefs, working out one’s subjective probability assignments, assigning values to the outcomes of one’s choices, intellectually structuring the decision problem in order to apply one’s decision-guiding principle to it, and so forth. Although much

deliberation involves purely psychological operations, it need not: one often utilizes paper and pencil, or a hand calculator, to elicit the information desired. Research, on the other hand, involves acquiring information not already implicitly contained in the agent’s “mental store.” Straightforward examples of research include consulting authorities, conducting experiments, retrieving data from newspaper files, and similar activities.

Since deliberation and research are activities which one can choose, the selection of information-gathering activities may be guided by practical decision guides. Agents often have limited information, not only about their prospective acts, but also about their potential information-gathering activities. Hence they frequently have the need to gather information about information-gathering itself. In deciding whether consulting the government forecast or the brokerage newsletter has greater expected utility, the agent might ask advice from her economist friend. But she has other options as well: perhaps she can consult forecast track records in the financial journals. Which of these information-gathering activities should she employ? These options, too, may be evaluated, and a decision guide may be used in doing so.

It begins to appear that the use of decision guides in decision-making is threatened with some form of infinite regress. To decide how to act, we must first decide how to decide how to act. But to decide this, we must first decide how to decide how to decide how to act. But to decide this, we must first decide . . . *ad infinitum*. A variety of theorists have commented on this problem. A characteristic statement of it is given by Raiffa:

People often ask: “How do you know whether or not it is worth the effort to make a formal analysis of a decision problem? Is this a decision problem itself? Can you do a decision analysis of whether it is worth doing a decision analysis?” I don’t know anyone who can give definitive answers to these questions, and I suspect one runs into a messy and explosive infinite regression if he tries to incorporate considerations of these questions into the formal structure of a decision-theoretic model. (Raiffa, 1968, p. 266)

Another statement is found in Elster, quoting Sidney Winter:

collecting information involves costs for the firm and therefore should be undertaken only to the extent that it is (expected to be) profitable. To acquire no information about the environment is irrational; to go on collecting information for a very long time equally so; and hence there must be some optimal amount of information that the firm should acquire. But once again this begs the question, since the “choice of a profit maximizing information structure itself requires

information, and it is not apparent how the aspiring profit maximizer acquires this information or what guarantees that he does not pay an excessive price for it" [Winter, 1964-5, p. 262]. The demand for an optimal amount of evidence immediately leads to an infinite regress. (Elster, 1983, pp. 17-18)

Another expression of it can be found in Michael Resnik's recent introduction to decision theory:

The difficulty here can be put succinctly by observing that *whenever we apply decision theory we must make some choices*: At the least, we must pick the acts, states, and outcomes to be used in our problem specification. But if we use decision theory to make those choices, we must make yet another set of choices. This does not show that it is impossible to apply decision theory. But it does show that to avoid an infinite regress of decision analyses any application of the theory must be based ultimately on choices that are made without its benefit. Let us call such decisions *immediate decisions*. Now someone might object that insofar as decision theory defines rational decision making, only those decisions made with its benefit should count as rational. Thus immediate decisions are not rational, and because all decisions depend ultimately on these, no decisions are rational. (Resnik, 1987, p. 11)<sup>1</sup>

The problem has long been thought by a variety of moral philosophers to undermine utilitarianism as an ethical theory.<sup>2</sup> Russell Hardin, however, dismisses the problem. "It should embarrass philosophers that they have even taken this objection seriously," he states, and advises them to follow James March and Herbert Simon in escaping the quandary by noting that often we *satisfice*, we do not maximize: we stop calculating and considering when we find a merely adequate choice or action (Hardin, 1988, p. 4).

In this paper I shall investigate issues raised by the fact that decision guides can be used to guide choices over information-gathering activities as well as choices over the target acts themselves. Since most theorists have focused on the problem as it concerns deliberation rather than research, I shall confine my attention to the former. In a useful shorthand I shall call our topic the "regress problem." As we have seen, however, one of the issues is whether or not a regress need arise at all. I shall argue that it need not.

#### 1 ALLEGED DIFFICULTIES RAISED BY THE REGRESS PROBLEM

It is commonly feared that the regress problem threatens the possibility of rational decision-making. But what exactly is the nature of the threat?

Many commentators fear that the regress problem will lead to "an infinite regress." This suggests that what is feared is an infinite regress of deliberations, each one a deliberation regarding the rationality of the subsequent deliberation. But very little reflection is required to see that no threat of this simple sort looms.

One possible way that the threat might arise is through the agent's being required to keep deferring – infinitely many times – the target acts while she deliberates further. (By "target acts" or "terminal acts" I shall mean those nondeliberative acts, the necessity for choosing among which originally creates the opportunity to deliberate.) But of course no such infinite postponement can take place. Even if the target acts can be deferred, there is only a finite length of time available for postponement. After some finite length of time, the agent (if human) will die, and by that point she has either performed a target act or failed to do so. She cannot deliberate indefinitely. Nor can she deliberate indefinitely in cases where the target acts cannot be deferred, but must be performed at some fixed time, say,  $t_n$ . Suppose that it is now any time  $t_1$  in the agent's lifetime prior to  $t_n$ . The agent cannot move farther back in time in order to secure more time for deliberation about what to do at  $t_n$ . Hence there is no possibility of the agent's somehow being required to deliberate for an infinite amount of time before she performs some target act. If there is a genuine threat from the regress problem, it must arise in a more subtle fashion.

I suggest that the threat should be described as follows.

*Universal Irrationality*: the regress problem entails that there is no action which it is rational to perform.

Two arguments can be given in support of this threat. First, it can be argued that no act is rational, since there is always some superior act that it would have been more rational to perform, namely acquiring information about the act in question. But this is equally so when the act in question is an act of acquiring information. Thus *no* act is rational. Second, it can be argued (following Resnik) that to avoid an infinite regress, agents must make some initial decision without reference to relevant decision-guiding principles. Since such an unguided decision is irrational, so are all the subsequent decisions and actions that flow from it.

In what follows I shall investigate the extent to which Universal Irrationality represents a genuine threat to rational decision-making. In examining this threat, I shall ask whether solving the problem requires us to utilize satisficing rules, which would otherwise constitute less attractive decision guides.

It is helpful to distinguish (as I just did in rough form) between two

different types of cases in which the regress problem may arise. In *nondeferral* cases, the agent possesses a set of alternative target actions performable only at a given time. The agent does not have the option of deferring any of these actions to any subsequent time. Any one of them must be performed at the given time; since the set of alternative acts is exhaustive, he must perform at least one of them at that time. In *deferral* cases at least some of the agent's set of alternatives can be deferred for future performance. In this paper I shall consider nondeferral cases only.

## 2 THE THREAT TO RATIONAL DECISION-MAKING WITH SIMPLE DECISION GUIDES

Agents' beliefs and information about their prospective actions vary enormously. Hence no single decision-guiding principle is adequate to the task of guiding every agent in every decision.<sup>3</sup> Some agents have fairly full and complex beliefs at the time of decision about their prospective actions. Such an agent could make a decision by utilizing, for example, the principle of maximizing expected utility (MEU). But some agents' beliefs are insufficient for them to derive any prescription from this principle. For example, an agent may lack beliefs about the probabilities of his actions' consequences.<sup>4</sup> Such an agent must employ a decision guide that requires a less comprehensive set of beliefs. He might do best to employ a satisficing rule or a maximin rule. The cognitive situation of every agent will be covered only if there is a hierarchy of decision guides, the joint range of which is sufficient to accommodate all the possible cognitive situations in which agents find themselves. If an agent's beliefs enable him to use either a higher or a lower decision guide in this hierarchy, then he ought to employ the higher one. For example, he ought to maximize expected utility rather than maximin. We can summarize this by saying that it is rational to choose an action if and only if it is prescribed by the highest decision guide that one is capable of using to make one's decision.<sup>5</sup>

Since to use a decision guide is to derive a prescription directly from it and one's beliefs, whether or not an agent can use a given decision guide depends on her actual beliefs at the time of decision. In determining what decision guide an agent can use, one must be strict about what beliefs an agent is counted as possessing. Suppose that proposition P = "Act A has an expected utility of 10 while its sole alternative, act B, has an expected utility of 8" and proposition Q = "Act A would maximize expected utility," and MEU states that an action is choiceworthy if and only if it would maximize expected utility. Further suppose that, at  $t_i$ , an agent

believes P but not Q. (Perhaps she has not yet seen that P entails Q.) Because this agent would not derive a prescription from MEU in a one-step inferential process, she is not able at  $t_i$  to use MEU in deciding between A and B. Before she can use MEU, she must derive Q from P. This is a (very short) process of deliberation. She must perform it before she can use MEU.

Since an agent's beliefs can change over time, the highest usable decision guide available to her may change as well. In such a case the action that it would be rational for her to choose may also change. Suppose that an agent must choose between doing A or doing B on Wednesday. On Monday she calculates (correctly) that A would maximize her minimum gain. The maximin rule (MM) is the highest one she can employ: hence she would be rational on Monday to choose A for Wednesday. Suppose that she comes to believe on Tuesday that B would maximize expected utility. This belief renders the superior rule of maximizing expected utility usable by her as a decision guide. Assume that this rule prescribes B. Then on Tuesday, by contrast with Monday, she would be rational to choose B for Wednesday. Normally in this sort of case the agent's information improves. Of course it can also degenerate, for example if she forgets previously held beliefs. Either form of belief change shows that our previous meta-principle should incorporate temporal indexing in the following manner.

- M. It is rational to choose an action at  $t_i$  for performance at  $t_n$  if and only if that action is prescribed by the highest decision guide that the agent is capable of using at  $t_i$  to make her decision.

Note that this meta-principle is Janus-faced: it implicitly provides two evaluations. On the one hand it evaluates a choice at  $t_i$  as the one it would be rational to make, while on the other hand it evaluates an act at  $t_n$  as the one it would be rational to choose.

It is often thought that what it is rational to choose depends, not on the action's satisfying the relevant decision guide, but rather on the agent's justifiably believing that it does. This view is best interpreted as a view about the appropriate content of decision guides. On such a view, for example, the correct statement of MEU says that an action is choiceworthy if the agent is justified in believing it to maximize expected utility. I shall call such decision guides "justified-belief decision guides." The impact of the regress problem on decision-making utilizing justified-belief guides will be considered in section 3. For the remainder of this section and the next, I shall confine the discussion to what I call "simple" decision guides, ones that merely, for example, prescribe an act as choiceworthy if it maximizes expected utility.<sup>6</sup>

An agent who is attempting to decide which action to choose may have excellent reasons for engaging in deliberation. This form of information-gathering activity may be valuable in a variety of different ways. First, it may supply the agent with more accurate beliefs by which to judge, relative to a given decision guide (such as MEU), which options it prescribes as choiceworthy. Second, it may supply the agent with beliefs that will render a superior decision guide usable in his decision (for example, it may enable him to switch from using MM to using MEU). Third, it may improve the agent's assessment of which decision guide is superior. For example, deliberation may lead him to believe that maximizing is superior to maximaxing. In all these ways it may enhance his chances of performing a superior act. Deliberation may also be valuable in ways not directly connected with the acquisition of information relevant to the problem at hand; for example, in using one's computer to generate spreadsheet information for a business decision, one may also discover that one's software should be replaced by a more powerful program.

Counterbalancing these possible gains are significant possible losses. First, some forms of deliberation possess intrinsic disvalue (for example, arithmetical calculation may be inherently unpleasant for the agent). Second, deliberation may involve opportunity costs (for example, tying up the computer in a prolonged statistical analysis prevents it from being put to other uses). Third, deliberation may have disadvantageous side-effects (for example, prolonged sessions in front of a video display terminal may cause the subsequent development of cataracts). Finally, deliberation may actually lead one to perform a less desirable act than one would have performed if one had thought less (for example, careful thought about the long-run consequences of a certain chess move blinds one to the immediate trap that it opens up).

Let us spell this out in terms of a concrete example in order to see how a regress problem might arise. Suppose that a physician must, at  $t_n$ , order one of two different types of chemotherapy (A or B) for a cancer patient. The physician believes that the outcome of this choice will be determined by the physiological origin of the cancer. He ascribes a probability of 0.6 to the cancer's having originated in the lungs (state  $S_1$ ), and a probability of 0.4 to the cancer's having originated outside the lungs (state  $S_2$ ). He believes that, if  $S_1$  obtains, therapy A would produce 800 utiles, while therapy B would produce 600 utiles, and further that if  $S_2$  obtains, therapy A would produce 700 utiles while therapy B would produce 1,200 utiles. He has not yet compared these figures to arrive at a belief as to which therapy would minimize the worst possible outcome, or which has the highest expected utility. It is now  $t_1$ , and the physician

is choosing which therapy to prescribe. Part of his representation of the choice may be depicted as follows:

	$S_1(0.6)$	$S_2(0.4)$
A	800	700
B	600	1,200

Let us confine our attention to two possible decision guides, MEU and MM. I shall assume that MEU is superior to MM and ought to be used by any agent capable of using either. In this example I assume for simplicity that MM and MEU each select a unique act, but of course this is not generally the case.

At  $t_1$  itself neither of these two decision guides can be used by the physician. However, if he engaged in suitable brief mental deliberation  $D_1$  starting at  $t_1$  he would shortly acquire beliefs sufficient for utilizing MM; if he engaged in slightly longer deliberation  $D_2$  starting at  $t_1$  he would acquire beliefs sufficient for utilizing MEU. Use of either decision guide would lead to his performance of the act actually prescribed by it. His options for  $t_n$ , and the acts that would follow them, can be represented as in figure 6.1. In this figure "D<sub>1</sub>" represents the deliberative process that would enable the physician to apply MM to his options at  $t_n$ , and "D<sub>2</sub>" represents the deliberative process that would enable him to apply MEU to the same options. "(MM- $t_n$ )" and "(MEU- $t_n$ )" represent the agent's deriving prescriptions for  $t_n$  from MM and MEU respectively. The dotted lines represent the duration of deliberation, while the broken lines represent the agent's "activities" following his derivation of a prescription. To simplify discussion of this case I shall assume that these activities (which of course may not involve overt actions at all) have zero utility. Thus in option 2 the physician calculates ( $D_2$ ) the expected utilities of A and B, derives a prescription (MEU- $t_n$ ) for B as having the highest expected utility at  $t_n$ , and finally does B. Note that this figure represents the physician's possible choices as a very well-informed observer might understand it, not necessarily as the physician does. The physician, for example, does not realize at  $t_1$  that his calculating which act would maximize expected utility would lead to his performing B at  $t_n$ . His mental representation of his choice

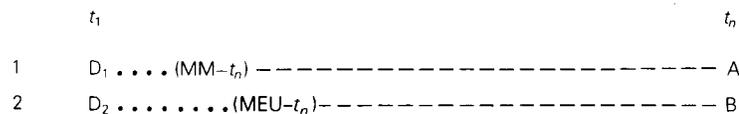


Figure 6.1

problem might refer to the act at  $t_n$  as "whatever terminal act is prescribed by MM."<sup>7</sup>

Because deliberation and immediate choice have their costs and benefits, which can be measured in utilities, we can ask which alternative at  $t_1$  it is rational (according to meta-principle M) for the physician to adopt.<sup>8</sup> For example, suppose that MEU is the highest decision guide that the physician is capable of using with respect to his options at  $t_1$ . Does MEU recommend that the physician deliberate in order to apply MM (as in option 1), or does it recommend that he deliberate at greater length in order to apply MEU (as in option 2)? The physician himself may ask this question as he seeks guidance in choosing which form of deliberation to utilize. But this choice, too, may be assisted by deliberation at a higher level. For example, if he does not already know the expected consequences of deliberating in order to apply MEU, he may need to calculate these in order to apply MEU to *this* choice. But this calculation also has its costs and benefits according to various decision guides, and so a choice must be made as to whether or not to carry it out. And the agent may need to deliberate before making this choice. Any of these forms of deliberation could be carried out at  $t_1$  as alternatives to  $D_1$  and  $D_2$ .

Taking this into account, we can more fully represent the physician's set of options as in figure 6.2. This diagram can be read on the model of figure 6.1. For example, in option 3, the physician begins by deliberating about deliberation at a lower level. He calculates ( $D_1^*$ ) the worst possible outcomes of (a)  $D_1$ , that is, calculating the security levels of A and B and then applying MM to his choice of A versus B, and (b)  $D_2$ , that is, calculating the expected utilities of A and B and then applying MEU to his choice of A versus B. On this basis he derives a prescription ( $MM-t_2$ ) from MM to perform  $D_1$ . He then derives a prescription ( $MM-t_n$ ) to perform act A at  $t_n$  and finally performs act A itself. (Note that I am here assuming that MM, when applied to the physician's choice

of what type of deliberation to engage in, prescribes deliberation relevant to the future application of MM itself. However, there is no reason to suppose that a decision guide will always prescribe deliberation leading to its subsequent own use. Option 6 reflects this fact. Note also that in order to carry out, say,  $D_1^*$ , the physician need not actually carry out the deliberations about which he deliberates.) In option 5, the physician calculates ( $D_1^{**}$ ) the worst possible outcomes of performing  $D_1^*$  or  $D_2^*$  at  $t_2$ . On this basis he derives a prescription ( $MM-t_2$ ) from MM to perform  $D_1^*$  at  $t_2$ . He then performs  $D_1^*$  as just described.

### 2.1 Universal irrationality: the first argument

We construed the threat posed by the regress problem as a threat of Universal Irrationality, according to which every action is irrational to perform. Two arguments were given for this threat. First, it was argued that no act is rational, since there is always some superior act that it would have been more rational to perform instead, namely acquiring information about the act in question. Second, it was argued that, to avoid an infinite regress, agents must make some initial decision without reference to relevant decision-guiding principles. Since such an unguided decision is irrational, so are all the subsequent decisions and actions that flow from it.

Let us examine the first of these arguments. In terms of figure 6.2 we can see that the argument assumes that (for example) it would not be rational for the physician to perform  $D_1$  at  $t_1$ , because it would be more rational for him instead to deliberate about performing  $D_1$ . That is, it would be more rational for him to perform, say,  $D_1^*$  at  $t_1$ , deliberating about whether to perform  $D_1$  at  $t_2$ . But by parallel reasoning it would not be rational for him to perform  $D_1^*$  at  $t_1$  either, because it would be more rational instead to deliberate about performing  $D_1^*$  - for instance, to perform  $D_1^{**}$  at  $t_1$ , deliberating about whether to perform  $D_1^*$  at  $t_2$ . And similar reasons would be introduced for rejecting any deliberation at  $t_1$  in favor of deliberating at  $t_1$  about performing that deliberation later.

In order to conclude that no act at  $t_1$  is rational, the argument requires that each agent possess an infinite number of alternatives for any given time, each alternative incorporating a higher level of deliberation about lower-level deliberations. (Otherwise the alternative incorporating the highest level of deliberation would emerge as the rational act, since no superior deliberation concerning it would be available to the agent.)

The descending dots in figure 6.2 suggest that the physician does indeed possess an infinite number of appropriate alternatives at  $t_1$ . But is this

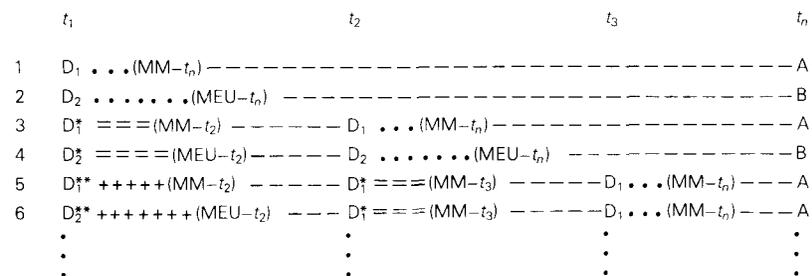


Figure 6.2

really the case? It is sometimes argued on metaphysical grounds that every agent, on every occasion, possesses an infinite number of alternatives. Even if true, however, this would not give rise to the problem of Universal Irrationality. This problem is only raised if the agent has a nonterminating ascending series of alternatives in each of which the agent deliberates about deliberating at the next lowest level. If the physician's alternatives, depicted in figure 6.2, ascend infinitely in the same pattern, this fact would provide a necessary condition for success of the first argument for Universal Irrationality of actions (necessary, but not sufficient, since the first argument for Universal Irrationality assumes *both* that there is an infinitely ascending set of deliberation alternatives *and* that each alternative involving a higher level of deliberation is superior to the alternatives involving any lower level of deliberation).

However, there is good reason to think that the alternatives available to the physician (or any other agent) cannot ascend infinitely in the same pattern. Any act must take some finite amount of time.<sup>9</sup> Hence an agent cannot, within a finite amount of time such as  $t_1$  through  $t_n$ , perform a sequence of acts having infinitely many members. But this is what the existence of an infinite ascending series of deliberation alternatives on the model of figure 6.2. would require. For any option N, however densely packed with deliberations about deliberations, there would have to exist a higher-level option including an additional deliberation about those deliberations. At some point this would no longer be possible, since there are only so many acts of deliberation that an agent can squeeze into a finite span of time. When this point is reached, the terminal option has been reached in the agent's series of alternatives of the figure 6.2 type.

Unfortunately, even if the number of deliberations that can be packed into a finite time span is finite, we cannot infer from this that there is an upper limit on the *level* of deliberation that an agent may perform, for an agent might possess alternatives which involve "gappy" sequences of deliberation. In these options an episode of deliberation at one level is not followed by deliberation at the appropriate next-lowest level. Thus in our current example the physician might have the following gappy option:

$D_1^{**} + + + + (MM-t_n) - - - - - A$

Here, although  $D_1^{**}$  consists of calculating the security levels of performing  $D_1^*$  and  $D_2^*$ , the physician does not appropriately proceed to either  $D_1^*$  or  $D_2^*$ , but rather leaps directly to applying MM to the choice of A versus B. Since it seems possible to perform such a sequence, the previous argument establishes no upper limit on the level of deliberation that an alternative might involve. For example, suppose that the agent can only

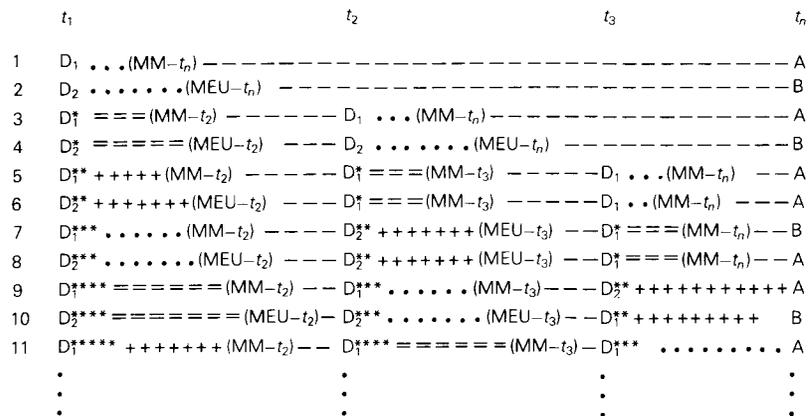


Figure 6.3

perform three deliberations between  $t_1$  and  $t_n$ , as shown in figure 6.3.<sup>10</sup> Even so, if he performs gappy sequences, he would have the depicted ascending series of options available. Notice that (after option 6) as the level of the initial deliberation performed at  $t_1$  goes up, lower-level deliberations drop out of the sequence in order to make room for the initial deliberation.

With the availability of gappy options of this sort it will be true that, for any given level of deliberation, the agent might perform a gappy option incorporating that level of deliberation as the initial act. Hence, for any given level of deliberation, the agent will have a (gappy) alternative initiated by a higher-level deliberation about that deliberation. Thus the threat of Universal Irrationality reasserts itself.

We should not forget, however, that human beings have severe constraints on their cognitive processing abilities. These constraints undoubtedly rule out any possibility of an agent's genuinely possessing the option of deliberating at a very high level, despite the possibility of performing gappy sequences. It already strains our capacities to grasp the physician's options as I have shown them in figures 6.2 and 6.3; any option much more complicated than  $D_1^{**}$  probably exceeds what any normal agent could actually perform.

But suppose that we waive this point. Let us grant, *arguendo*, that agents possess infinitely ascending series of (gappy) deliberation alternatives. Does this drive us to the conclusion that no act is ever rational to choose, because for any given act, some alternative act involving a higher level of deliberation would have been superior? The answer is no.

I shall argue for this answer in detail, but it is worth noting the intuitive plausibility of this answer. Deliberation, as we have seen, may have beneficial effects. But it may also have significant costs, and there is no reason to suppose in advance – as the first argument for Universal Irrationality supposes – that the benefits always outweigh the costs.

According to meta-principle M, two conditions must be met in order for an act to be rational for the agent to choose and perform. First, there must be at least one legitimate decision guide that the agent is capable of using to make his decision. Second, some act must be prescribed by the agent's highest usable decision guide. (This act is rational for him to choose and perform. Notice, as was pointed out in note 5, that the relevance of a decision guide in determining which act is rational for an agent is partly a function of the decision guide's usability for the agent. Its usability depends on the agent's stock of beliefs. But the rationality of the act is determined by its being prescribed by the decision guide, not by the agent's believing that it is prescribed by the decision guide. He may not believe this.) Are these conditions satisfied in the physician's case – is there an alternative at  $t_1$  that the physician would be rational to choose immediately prior to  $t_1$ , say at  $t_0$ ? I shall argue that there is such an alternative; it provides a counter example to Universal Irrationality. Indeed, I shall argue even more strongly that there is a significant class of such counterexamples. Let us suppose that the physician believes at  $t_0$  that the maximin action for  $t_1$  is  $D_1$ . This belief renders MM usable by the physician at  $t_0$  with respect to his alternatives for  $t_1$ . Hence the first condition is met. Let us further assume that MM is the physician's highest usable decision guide. The only remaining question is whether MM prescribes any action for  $t_1$ . Given that the physician has – and, let us stipulate, believes that he has – an infinite number of alternatives for  $t_1$ , can MM prescribe any action for  $t_1$ ?

Under certain plausible assumptions, MM does prescribe an alternative for  $t_1$ ; in particular, it prescribes  $D_1$ . Moreover, these assumptions are ones that may frequently hold in cases having the same structure as the physician's present case. Any cases having this structure, and for which these assumptions hold, will be ones in which MM prescribes the analog to  $D_1$ . That is, it prescribes calculating what terminal action MM requires, and then performing that action. In these cases, contrary to the first argument for Universal Irrationality, meta-principle M does not imply that, for any alternative, it would always be better for the agent to deliberate about that alternative.

What are these assumptions? Examining the physician's case we can see that the security level (utility of the worst possible outcome) of any act at  $t_1$  is determined by two variables: (a) the intrinsic security level of the

physician's deliberative process between  $t_1$  and  $t_n$  (for example, the process  $D_1^* = = (MM-t_2)D_1 . . . (MM-t_n)$  in option 3), and (b) the intrinsic security level of the terminal act (for example, terminal act B in option 2) that he would perform at  $t_n$ . By "intrinsic security level of his deliberative process" I mean the security level it possesses *apart from* the possible terminal acts to which it may lead; by "intrinsic security level of the terminal act" I mean the security level it possesses apart from the processes leading up to it. Of course, in a realistic case a deliberative process may have side-effects that influence the act's security level. However, since the presence of such effects is immaterial to our concern, I shall assume that the only possible effects are those inherent in deliberation itself.

If a given alternative at  $t_1$  involves (a) a deliberative process having an intrinsic security level at least as high as that of any alternative, and (b) a terminal action having an intrinsic security level at least as high as that of any alternative, then that alternative is prescribed by MM. There are plausible assumptions under which  $D_1$  meets both these conditions. (Of course,  $D_1$  might be prescribed by MM even if it met only one of these conditions. But I shall argue that it meets both. This fact makes it unnecessary to quantify and compare the values of the security levels in question.)

Let us first consider the intrinsic security level of the deliberative process involved in  $D_1$ . Suppose that deliberation in itself always has disutility for the agent, that longer deliberation has greater disutility than briefer deliberation, that higher-level deliberation has greater disutility than lower-level deliberation of the same duration, and that all alternatives to  $D_1$  involve either longer or higher-level deliberation than  $D_1$  itself. It follows that  $D_1$ 's deliberative process has a higher intrinsic security level than the deliberative process involved in any of its alternatives, since  $D_1$  (being the shortest and lowest-level deliberation) involves, for certain, less intrinsic disutility for the agent.

Let us next look at the intrinsic security level of the terminal action that the physician would perform if he performed  $D_1$ . Suppose that if the physician performs  $D_1$  he will then perform, at  $t_n$ , the action that maximizes the minimum possible gain. The terminal action that follows any alternative to  $D_1$  cannot have a higher intrinsic security level than this.<sup>11</sup>

Since the intrinsic security level of  $D_1$ 's deliberative process surpasses those of its alternatives, and the intrinsic security level of its terminal act at least equals those of its alternatives,  $D_1$ 's overall security level exceeds theirs. It is prescribed by MM as the maximin act.

The assumptions that led to this conclusion may hold in a number of cases. Moreover, the physician himself may have noted their truth in

previous cases, and concluded that in all cases of this sort the maximin act involves calculating the security levels of the terminal acts. This conclusion may be what leads him to believe, as we assumed, that  $D_1$  would maximize the minimum possible gain. (Recall that his having this belief makes MM usable by him for deciding what to do at  $t_1$ .) In describing the case I simply assumed that he had this belief. We can now see that the belief need not be unreasoned or arbitrary.

I have described assumptions under which MM prescribes an act for the physician, and under which the physician would be rational to choose and perform that act. These assumptions are not bizarre or outlandish. On the contrary, they are likely to hold in many cases having the structure of the physician's case. Hence we have undermined the threat of Universal Irrationality. There are at least some cases, and possibly a great many, in which some particular act would be rational for the agent to perform. It is notable that this act does not involve deliberation at a very high level. The quest for rationality need not drive agents to higher and higher realms of deliberation about deliberation. We may also note that this demonstration does not depend on idiosyncratic features of MM as a decision guide. Similar conclusions can be reached about MEU, an exercise that I shall leave to the reader.<sup>12</sup>

Of course there are other kinds of cases in which the agent cannot derive any prescription from MM. Suppose, contrary to our previous assumptions, that the physician derives substantial positive utility from deliberation, and greater utility from higher levels of deliberation. Then MM rates any act involving more and higher deliberation as superior to any act involving less and lower deliberation. If the physician genuinely possessed an infinitely ascending set of deliberation alternatives, then MM could not select any alternative at  $t_1$  as the most rational.

But this kind of case does not pose a problem for the rationality of decision-making that needs to be taken seriously in the context of this discussion. First, such conditions are highly unlikely – much less likely than those previously laid out. Normal cases will resemble our previous case more than this one. Second, the underlying phenomenon that produces the untoward result in the revised case has nothing to do with the hierarchical structure of deliberation. The result only arises because of the bizarre character of the agent's preferences. Similar problems can arise in any case involving an infinitely large set of (non-deliberation) alternatives for which the agent has ever-increasing utility values. For example, if an agent derives higher utility the faster she drives her car, and for each speed at which she can drive her car there is another faster speed at which she could drive it instead, then (assuming all other possible outcomes remain stable across her alternatives) MM can select no action as

the most rational one for her. The problem we see here has nothing to do with the hierarchical structure of deliberation. Moreover, the fact that in the unusual conditions just described MM prescribes no act for the physician only shows that he is unable to apply MM as a decision guide. There may be other decision guides that he can apply, and the highest of these will dictate what it is rational for him to do. A satisficing decision guide is one example of the kind of principle that can cope with this sort of situation. But the bottom line is this: we have found a general class of cases in which an action can be identified as rational, and so we can reject the claim that the regress problem shows that *no* action can ever be rational.

In general, there is no reason to think that an agent's highest applicable decision guide will require very high levels of deliberation. In normal cases (ones where the agent does not derive positive utility from the process of deliberation itself), the primary reason to deliberate is to enable oneself to make a better choice about the terminal acts (acts A and B in the physician's case). As we have seen, the only way in which an agent can engage in options involving very high levels of deliberation is for him to perform "gappy" sequences. We saw, however, that the proper sequence of deliberations in such options is broken, so that "pre-gap" deliberations do not exercise proper normative control over "post-gap" deliberations and acts. It is unlikely that the agent gains anything by pursuing such gappy options, since they are unlikely to enable him to make a better choice about his terminal acts. If any act of his is rational, it would normally involve a sufficiently low level of deliberation to avoid gappy sequences.

## 2.2 Universal irrationality: the second argument

Two arguments for Universal Irrationality were described. We have now examined the first of these and found it wanting. The second argument states that, to avoid an infinite regress, an agent must make some initial decision without reference to relevant decision-guiding principles. Since, according to the argument, such an unguided decision is irrational, so are all the subsequent deliberations and acts flowing from it. (Since, in conjunction with meta-principle M, we have already implicitly introduced the term "irrational" to express another concept, let us label unguided decisions "arbitrary" rather than "irrational.") This argument is mistaken. Consider again the case of the physician as described in connection with figure 6.2. In assessing the second argument we must consider two distinct scenarios. In the first the physician has never made a decision prior to his decision at  $t_0$  with respect to his options at  $t_1$ ; this decision is an

“initial decision” in the fullest sense. In this scenario, any decision he makes at  $t_0$  cannot be tainted by previous defective decisions. (Of course it is implausible that an adult agent should never have made a decision, but ignore this.) Despite his lack of decision-making experience, at  $t_0$  the physician correctly believes that MM is his highest usable decision guide. Moreover he possesses enough beliefs at  $t_0$  to enable him to derive a prescription from MM to perform  $D_1$ . Thus MM is usable by him. In deriving this prescription and choosing to perform  $D_1$ , he performs an “initial decision.” But, *contra* the above argument, this decision is guided by a relevant decision-guiding principle, namely MM. The mere fact that the decision is the first decision that the agent has ever made in no way shows that it cannot be guided by a correct decision guide. Without making any prior decisions, the agent may nonetheless have acquired (say, through early education and through a combination of perception and natural induction) the beliefs about decision guides and about his circumstances necessary for making a guided decision. So his decision is not arbitrary, and it does not taint any of the acts or deliberations flowing from it. It need not be the case, then, that an infinite regress can only be avoided by engaging in arbitrarily unguided initial decisions.

Now let us consider the second scenario. In this scenario the physician does make an arbitrary decision, one unguided by any decision guide. Need this arbitrary decision taint the subsequent decisions and acts that issue from it? Suppose that the physician reads an article about decision theory at  $t-1$ . Reading this article influences his beliefs about his choice problem for  $t_0$ , but the act of reading the article itself is performed completely spontaneously, without reference to any decision-guiding principle. The decision to read the article is arbitrary. The decision to read the article results in the physician’s deciding at  $t_0$  to comply with MM by performing  $D_1$  at  $t_1$ . Does this mean that the decision to perform  $D_1$  is in some way defective? As far as I can see, the answer is no. Although the decision to read the article is arbitrary, and this decision leads to the decision to perform  $D_1$ , the decision to perform  $D_1$  is nevertheless nonarbitrary – it is guided by a relevant decision guide. Moreover, if the highest decision guide usable by the agent at  $t_0$  dictates his choosing  $D_1$  at  $t_1$ , then that choice is rational (as is the performance of  $D_1$  itself). In this circumstance both the decision to perform  $D_1$  and the performance of  $D_1$  itself are dictated by the Janus-faced meta-principle M. The fact that an arbitrary choice prior to a given decision provided information that influenced that decision has no tendency to show that the decision itself is arbitrary or irrational. Nor does there seem to be any other substantive sense in which this decision is irrational. We can conclude that the second argument for Universal Irrationality has even less foundation than the first.<sup>13</sup>

### 2.3 The role of satisficing principles

Let us note an important implication of our discussion. Many theorists such as Elster and Hardin appear to believe that the solution to the regress problem requires agents to utilize satisficing principles to guide their decisions (Elster, 1983, p. 18; Hardin, 1988, p. 4). Since such rules appear on other grounds to be inferior to traditional maximizing decision guides, the necessity of invoking this solution may show that the regress problem demonstrates an important deficiency in the rationality of decision-making.

But we must be careful before drawing this conclusion. On the one hand, we should remember our previous admission that, quite apart from the regress problem, agents need a large range of decision principles to guide their decision-making. Agents often lack sufficient beliefs about their terminal actions to use such decision guides as MEU or even MM, and so must turn to lower-ranking decision guides. Indeed, to assist all agents making all decisions, however cognitively impoverished those agents may be, a very large number of decision guides may be required. Some of these principles will undoubtedly be satisficing principles. Hence, even if there were no regress problem, satisficing principles would still be an important weapon in our arsenal of decision guides. Principles such as MEU cannot do the work alone. If the regress problem forces decision-makers to turn to satisficing principles, its doing so hardly introduces a new element into the practice of making decisions.

Second, we have seen that satisficing principles are *not* required to solve the regress problem. We have described a case in which the decision-maker’s use of MM avoids any problem arising from potential regression in deliberation. In such cases the agent need not, and should not, follow a satisficing principle when a superior decision guide will serve.<sup>14</sup> Of course satisficing principles often enable agents to derive prescriptions for action quickly and on the basis of sparse information. Hence the rational agent will often deliberate in order to apply a satisficing principle rather than some other. Satisficing principles have their place in the hierarchy of decision guides. But the regress problem often can, and should, be defused without invoking them.

### 3 JUSTIFIED-BELIEF DECISION GUIDES

In section 2 we considered whether the regress problem threatens the rationality of decision-making with “simple” decision guides such as “An act is choiceworthy if and only if it would maximize expected utility.” We

noted that many theorists hold that the rationality of choosing an action depends, not on the action's satisfying the relevant decision guide, but rather on the agent's justifiably believing that it does. This view was interpreted as a view about the content of proper decision guides. On such a view, the above version of MEU should be rejected in favour of something like "An act is choiceworthy if and only if the agent *is justified* in believing that it would maximize expected utility." In this section we consider whether such "justified-belief" decision-guides are more vulnerable to difficulties raised by the regress problem than are simple decision guides.

We should note that not every decision-guide can be a justified-belief rule. For a justified-belief guide to be usable by an agent, the agent must believe that her beliefs about the features of her prospective acts are justified. (According to such rules, an act counts as choiceworthy if the agent has certain justified beliefs regarding it. Hence for her to be able to *use* the rule, she must believe that she has these justified beliefs.) However, an agent may have plenty of justified beliefs about the features of her prospective acts but be uncertain whether *any* of these beliefs is justified. (Perhaps she does not recall the source of her beliefs, or perhaps a philosophical skepticism has rendered her uncertain whether any beliefs are justified.) In such a case she could not apply any justified-belief rule to her decision. Yet she still needs to make a decision, and could do so by invoking a pertinent simple decision guide. Since this seems a rational strategy, it appears that the set of appropriate decision guides must include simple as well as justified-belief guides.

But, for the sake of argument, let us explore the view that *all* decision guides must be justified-belief rules. We may stipulate that no choice is rational for an agent who cannot apply any actual-justification rule. Does the regress problem raise difficulties for decision-making with justified-belief rules that do not arise when all decision guides are simple? I shall argue that it does not.

If justified-belief rules raise any difficulty, it would be a version of the problem of Universal Irrationality. One could argue as follows for the existence of such a difficulty. To avoid an infinite regress, agents must make an initial decision to perform some act without any prior deliberation about that decision. Suppose that this initial decision is the decision at  $t_1$  to deliberate at  $t_2$  about an act to be performed at  $t_3$ . Since, by hypothesis, the decision at  $t_1$  is made without prior deliberation, then the beliefs in virtue of which the agent makes his decision at  $t_1$  are unjustified. But if they are unjustified, then the decision is irrational, since it is not prescribed by any justified-belief decision guide. If it is irrational to decide at  $t_1$  to deliberate at  $t_2$ , then (by Janus-faced meta-

principle M) deliberating at  $t_2$  is also irrational. If deliberating at  $t_2$  is irrational, then it cannot provide justified beliefs for the decision to perform any act at  $t_3$ . Hence the choice to perform an act at  $t_3$ , and the chosen act itself, are similarly irrational. This argument can be generalized to show that every decision-making process is similarly flawed, and hence that every action is irrational.

Let us call this argument the "zipping-forward" argument. The argument has a persuasive appearance, but it is incorrect. The first flaw to notice is that it would not be accepted by anyone who accepts a current time-slice theory of epistemic justification – for example, foundationalism or coherentism.<sup>15</sup> According to such theories, a current belief is justified just in case it stands in the correct relationship to other current beliefs or quasi-doxastic states of the agent. Hence the agent's justification in believing something does not depend in any way on what kind of deliberation preceded or generated that belief. For example, coherentism holds that, if the agent makes a decision at  $t_1$ , and the decision rests on beliefs that are justified by their relation to other contemporaneous beliefs, then lack of prior deliberation in no way shows that the decision is irrational. For the coherentist, or other current time-slice epistemologist, the zipping-forward argument cannot get off the ground, and the rationality of decision-making is not impaired.

The argument looks more persuasive, however, for theorists who accept a *historical* account of justification. According to such accounts, the justificatory status of a current belief depends entirely on what prior processes generated it. For example, a simple "reliabilist" theory might state that a belief is justified just in case it is produced by a reliable belief-forming process such as deduction from true premises. Beliefs formed by unreliable processes (such as wishful thinking) are unjustified.<sup>16</sup> In the context of reliabilist accounts of justification, it appears as though the necessity for an agent to make some initial decision without prior deliberation validates the zipping-forward argument.

However, this appearance is an illusion. Although reliabilism stipulates that a belief must be generated by a reliable process, this process need not be a form of *deliberation*. Nor need it be any other process which the agent *chooses* to perform. To take a dramatic example, consider an agent who is kidnapped by terrorists and strapped to a chair. While strapped in, he is unavoidably exposed to certain sounds. These sounds initiate a perceptual process resulting in the agent's forming various beliefs ("Now my captors are leaving the building"). These processes may be reliable, and so the beliefs will be justified, even though they did not result from deliberation or from any other process that the agent chose to initiate or could have avoided. Hence, contrary to the assumption of the previous

paragraph, an agent's beliefs can be justified even though they do not result from any previously chosen deliberational process. And if those beliefs form the basis for a decision (say, the decision to try to escape while one's captors are absent), then the decision can be rational, because it can be in accord with a suitable justified-belief decision guide. Of course, few of us fall into the hands of terrorists, but for most of us the processes that produce our beliefs as infants and young children are ones we do not choose and about which we do not deliberate. This fact does not prevent our beliefs or subsequent decisions from being rational. Even for historical accounts of justification, the zipping-forward argument is incorrect.

It is important to see that the zipping-forward argument goes wrong in an additional way. The argument begins by assuming that the beliefs in virtue of which the agent decides at  $t_1$  to deliberate are unjustified because they do not result from any process of deliberation. We have just seen that this is a mistake; justified beliefs may arise from involuntary nondeliberational processes. But let us consider a case in which the beliefs on the basis of which the agent decides at  $t_1$  to deliberate at  $t_2$  are *not* justified. Since the decision is not prescribed by an appropriate justified-belief decision guide, it is irrational, and the agent's deliberating at  $t_2$  is likewise irrational. The zipping-forward argument assumes that an irrational deliberation at  $t_2$  cannot provide justified beliefs for the decision to perform any act at  $t_3$ . Hence any decision with regard to  $t_3$ , and any act at  $t_3$ , are themselves irrational. But this assumption is a mistake, arising from an illegitimate confusion of the evaluations of epistemology with those of practical decision-making. The deliberation at  $t_2$  may have been irrationally chosen and performed. But as an epistemic process it may be perfectly reliable, and confer justification on any beliefs arising out of it.

To see this, consider an agent whose inordinate love of arithmetical calculation leads her to decide at  $t_1$  to deliberate at  $t_2$  about the expected utilities of her prospective acts at  $t_n$ . She is not justified in believing at  $t_1$  that this form of deliberation would be best, since this belief arises as a result of an unreliable process in which her love of calculation plays an inappropriate role. Her evidence indicates that she ought instead to calculate the security levels of her prospective acts. Hence her decision to calculate expected utilities is irrational, and deliberating itself is an irrational act. However, the process of deliberating – her calculating the expected utilities – may be perfectly reliable as an epistemological process, and it may deliver the justified belief at  $t_3$  that act A would maximize expected utility. Suppose that the justified-belief version of MEU is the agent's highest usable decision guide at  $t_3$ . If act A satisfies this decision

guide, a decision at  $t_3$  to perform A is rational, and so is the act itself. The mere fact that the agent irrationally chose to perform this deliberation on the basis of unjustified beliefs does *not* taint all her subsequent decisions and acts with irrationality.

What this example shows is that it is important to separate epistemological assessments from those of practical rationality. A deliberational process that is irrational from the perspective of practical rationality may nonetheless be justification-conferring from the point of view of epistemology. For this reason a single irrational decision in a sequence of decisions need not infect all the subsequent decisions with irrationality. Thus even if we were to accept the view that all decision guides must be justified-belief rules, and even if we accepted a purely historical account of what makes beliefs justified, we need not worry that a version of Universal Irrationality undermines practical rationality. As we saw before, some justification-conferring processes are not themselves the subject of decisions, and even those that are, and that are chosen irrationally, do not necessarily taint the decisions and acts that follow them.

#### 4. CONCLUSION

Because information-gathering processes, and in particular deliberation, are themselves activities that can be chosen and about which an agent can deliberate, it has often been feared that decision-makers may be vulnerable to some kind of vicious regress undermining the rationality of decision-making. This fear is best articulated as the concern, supported by two distinct arguments, that decision-making is subject to Universal Irrationality: all acts and decisions are irrational. I have argued that this fear is ill founded, whether decision-making is guided by simple decision rules or by justified-belief decision rules. I have not attempted to rule out the possibility that the regress problem renders some acts and decisions irrational. But I have argued that in a substantial range of normal cases this threat can be dismissed.

#### NOTES

I am grateful to Robert Cummins, Alvin I. Goldman, Keith Lehrer, John Pollock, and the editors of this volume for helpful comments on an earlier version of this paper.

<sup>1</sup> Resnik himself rejects this objection, advocating the employment of "policies" instructing us when to deliberate and when not. He admits that the use of such

policies may itself require reassessment, and so raise the same problem all over again. More detailed examination of such policies and their re-examination is contained in Bratman (1987, Chs 5, 6).

- 2 See Bales (1971, pp. 257-65) for references to holders of this view.
- 3 See Smith (1988, pp. 89-108) for development and defense of this claim, and discussion of issues in this section.
- 4 In some accounts, to have a subjective probability assignment for a given proposition is to have a certain disposition to act on that proposition (for example, to bet on it). These accounts often assume that such dispositions are always in existence, so that (contrary to the text) the agent never lacks beliefs about the probabilities of his actions' consequences. This simplistic functional account of belief states seems to me inadequate. However, even if true, there is no guarantee on such an account that the agent has knowledge (or even beliefs) about his own relevant behavioral dispositions. Thus it may take him time and a variety of thought experiments - that is, deliberation - to determine what his subjective probability assignments are. (See Raiffa (1968, ch. 5) for a description of a method of eliciting subjective probability judgments.) In some accounts, an agent's choices always reflect his subjective probability and utility assignments, even though he himself may be unaware of the content of these assignments. These accounts appear to leave little room for agents to reflect on, or need guidance in, decision-making.
- 5 The criterion for rank-ordering decision guides will not be discussed in this paper; the reader is invited to provide his or her own criterion.

Through out this paper I assume that the relevant notion of "capable of using" does not entail that the agent makes, or can make, the *correct* derivation of a prescription from the decision guide. All it requires is that she be able to derive a prescription. See Smith (1988) for further discussion. Thus the agent might believe that A would maximize expected utility, and so be able to use MEU. However, if MEU actually prescribes B, then it is rational for her to choose B, not A.

- 6 An alternative interpretation of this view is possible, according to which the content of the decision guides remains simple (in my sense), but what enables an agent to apply a guide is his justified beliefs, rather than his actual beliefs. I shall not discuss this alternative interpretation here, but the conclusions reached in section 3 would remain roughly the same.
- 7 In this case I am assuming, for brevity of presentation, that the agent has only two alternatives available at  $t_n$ . In most cases of course the number of alternatives would be greater. Maximizing decision guides (such as MEU and MM) are often criticized on grounds that there are infinitely many alternatives for a given time. (This presents a *theoretical* difficulty only if, for any alternative act, there is another alternative that has a higher security level or expected utility. It presents a *pragmatic* difficulty for the usability of such decision guides if the agent does not have time or resources to discover, among the infinite set of alternatives, the one (if there is one) having the maximum security level or expected utility.) I shall set this alleged difficulty aside, since

it is distinct from the difficulties raised by the opportunity to deliberate. In any event, most agents do not consider themselves as having an infinite number of alternatives about which they *should* deliberate; their attention is restricted to a few salient alternatives.

- 8 I shall assume for the purposes of this paper that the rationality of a given option at  $t_1$  is determined, in part, by the actions that would actually follow it. Thus the rationality of the physician's performing  $D_1$  at  $t_1$  is partly a function of the value or disvalue of his then performing A. This assumption is necessitated by the fact that agents' subsequent acts are often not the ones it would be rational for them to perform. Because of this assumption, I do not use the traditional decision tree to represent the agent's choice at  $t_1$  in figure 6.2. For discussion and defense of this view, see Smith (1976, pp. 449-87).
- 9 If time can be divided into infinitely small intervals, then any bodily movement consists of nested shorter and shorter parts (for example, moving your hand for 1 second, moving your hand for 0.1 second, moving your hand for 0.01 second, . . .). In this way movements can be defined for any interval of time, however small. However, these arbitrarily small movements do not count as acts. In waving your hand for a minute, you also wave your hand for 0.0000001 second. But since you cannot effectively choose to wave your hand (as an independent motion) for 0.0000001 second, doing so does not count as an action, anymore than falling off a collapsing building counts as an action, even though it is a motion of your body.
- 10 I am assuming in this example that deliberations at the second level or above take no more time than those at the first level. If deliberations become longer as their level increases (a highly plausible thesis), there will be a highest-level deliberation (the highest-level deliberation just short enough to fit between  $t_1$  and  $t_n$ ) that could be performed by the agent between  $t_1$  and  $t_n$ .
- 11 Alternatively, it may be uncertain whether the physician's performing  $D_1$  would lead to his performing the maximin act at  $t_n$  - perhaps there is a possibility that he would miscalculate the security levels, or perhaps there is a possibility that he would fail to carry through by doing at  $t_n$  what he identifies at  $t_2$  as correct. In this case, we must ascribe equal security levels to the terminal acts following each of the physician's alternatives at  $t_1$  - since any of those alternatives may fail to lead to the maximin act at  $t_n$ .  $D_1$  would still be prescribed as choiceworthy by MM, since the security level of its terminal act is no less than the security level of the terminal act following any alternative.
- 12 It is easiest to demonstrate this result for MEU by concentrating on the kind of case in which deliberation provides the agent with better information for applying the same decision guide, e.g. MEU. For a helpful discussion of the rationality of obtaining information by research (as opposed to deliberation), see Raiffa (1968, chs 3 and 7).
- 13 At this point a pressing question might be raised, and has been raised by the editors of this volume. Suppose that an agent has to perform either action

X or action Y at time  $t_n$ . She has beliefs that would enable her at  $t_1$  to apply either MM or MEU to the choice between X or Y. MM prescribes X, while MEU prescribes Y. The agent is familiar with both MM and MEU. We have assumed that MEU is a higher decision guide than MM. However, this agent falsely believes that MM is higher than MEU. According to meta-principle M, she would be rational at  $t_1$  to choose Y, since Y is prescribed by the highest decision guide (namely MEU) that she is capable of using at  $t_1$  to make her decision. But can it really be true that she would be rational to do Y, since she believes MM to be the higher decision guide, and MM prescribes X rather than Y?

The answer has to be both yes and no. The choice to perform Y is the most rational choice for her in one clear sense – a sense somewhat analogous to the sense in which an act Z is the right act for an agent to perform because it is prescribed by the correct moral code. It is correct in this sense even though the agent believes some alternative code to be correct, and the alternative code prescribes act W instead of Z. We can contrast the correctness (in this sense) of performing Z with the correctness, in a different sense, of performing the act W prescribed by the moral code that the agent believes to be correct. We might call this second sense the “putative” sense of correctness. Similarly we can contrast the rationality of choosing act Y in our original case with the rationality of choosing X, the act prescribed by the decision guide that the agent believes to be correct. In a different, equally clear, sense, it would be rational for the agent to choose X. We might call this the “putative” sense of rationality. There is, as far as I can see, no way to avoid the divergence between the first and the putative senses of rationality (and correctness in the case of codes). In this paper I have restricted myself to the first sense, rather than the putative sense, but it is clear that the latter has an important role to play. Part of what is interesting about the divergence between these senses of rationality/correctness is that the bifurcation cannot be avoided (as some theorists might hope) by shifting from “objective” norms, such as moral codes, to decision guides.

Note that in the putative sense of rationality/correctness there may be *no* act which it is rational to choose or correct to perform. If the agent does not believe of any moral code that it is correct, then there is no act which it would be correct (in the putative sense) for her to perform. Similarly, if the agent does not believe of any decision guide that it is the highest usable one for her, then there is no choice which it would be rational (in the putative sense) for her to perform. I have explored this problem, from a somewhat different point of view, in Smith (1988) where I argue that its existence places a limit on how widely usable we can expect moral codes to be, even when they are augmented with auxiliary decision guides of the sort represented by MM and MEU.

Does the possibility that an agent may be uncertain which decision guide is highest raise an infinite regress problem that has not been addressed by the arguments in this paper so far? I would argue that it does not. Suppose that our agent is at  $t_0$ , and suppose that she does not yet have the beliefs necessary to apply either MM or MEU, although she can obtain them by

engaging in deliberation  $D_1$  or  $D_2$  respectively. Assume that she also has the option at  $t_0$  of deliberating whether MM or MEU is the highest decision guide. Then her situation at  $t_0$  might be represented as follows:

	$t_0$	$t_1$	$t_2$	$t_3$	$t_n$		
1	$D_1$	.....(MM- $t_n$ )	-----		X		
2	$D_2$	.....(MEU- $t_n$ )	-----		Y		
3	$D_3$	.....( $D_2$ - $t_2$ )	-----	$D_2$	.....(MEU- $t_n$ )	-----	Y

If she engages in  $D_3$ , she comes to believe that MEU is the highest decision guide and so ought to be applied via  $D_2$ ; she then engages in  $D_2$ , calculates whether X or Y has the greatest expected utility, and eventually does Y. Merely adding this option does not commit her to any infinite regress. If we ask “Which option at  $t_0$  is rational for her to choose at  $t_{-1}$ , just before  $t_0$ ?”, the answer is given by meta-principle M, for our guiding sense of “rationality.” What it is rational for her to choose at  $t_{-1}$  depends on which option is prescribed by the highest decision guide usable by her at  $t_{-1}$  for her options at  $t_0$ .

I have not tried in this note to spell out the conditions of rational choice for putative rationality. However, there is no apparent threat of infinite regress there either. What it is putatively rational for her to choose at  $t_{-1}$  depends on what she believes at  $t_{-1}$  to be her highest usable decision guide. If she had no belief about what her highest usable decision guide is, then there is nothing at  $t_{-1}$  that it is putatively rational for her to choose at  $t_0$ . She has a problem, but it is not an infinite regress problem.

- 14 In the case described in the text, a satisficing rule was not considered as one of the physician’s possible decision guides. However, even if such a rule had been included, we cannot conclude that using it, or deliberating further in order to use it, would have been the rational choice for the agent. Employing a satisficing rule typically takes less deliberation than employing a decision guide such as MM, but its potential benefits may not be nearly so advantageous. Theorists often assume that a decision-maker maximizes overall expected utility by utilizing a satisficing rule in deciding how much information to obtain, but this need not be so. For recent expressions of this assumption, see Nozick (1981, p. 300) and Hollis (1987, ch. 8). Robert C. Richardson argues against this mistake in his unpublished manuscript, “Satisficing and optimizing.”
- 15 Detailed versions of coherentism are provided by Lehrer (1974) and Bonjour (1985). Foundationalism is advocated by Chisholm (1977), and is discussed by Pollock (1986, ch. 2). The terms “current time-slice” and “historical theories of justification” were introduced by Goldman (1979).
- 16 A detailed reliabilist theory is developed by Goldman (1986).