The infinite regress of optimization

Philippe Mongin

Delta, Ecole Normale Supérieure, 48 Boulevard Jordan, 74014 Paris, France

Electronic mail: *corsec@bucllw11.bitnet*

The critique of the optimizing account of individual decision-making has generally emphasized either allegedly empirical failures ("in the real world businessmen do not maximize") or more subtle methodological difficulties of the sort usefully discussed by Schoemaker. This commentary deals with an altogether different class of problems that are best referred to as *logical* ones. As Schoemaker also briefly indicates, there is an infinite regress lurking behind the use of maximizing concepts by decision theory as soon as the latter relaxes the tacit assumption of zero informational and computational costs. Maximizing a constrained objective function is indeed a costly procedure on the agent's part. Supposing that one's search and algorithmic costs can be assessed, these should be made part of one's choice situation. The result is a metaoptimal decision that may or may not upset the optimal one. Supposing that the costs of the procedure used to reach the metaoptimal decision can in their turn be assessed, they should be made part of the agent's choice situation, and so on *ad infinitum*.

There are two ways of understanding the infinite regress just sketched: (a) as a threat to the consistency of the agent's decision criterion, or (b) as a threat to the consistency of the observer's decision theory. It is not clear to what extent agents can assess higher-order costs. A reasonable guess is that more often than not they discover them once they have been incurred; hence there is little sense in engaging in higher-order decision making (except in the case of repetitive decisions). The safe interpretation of the infinite regress is in line with (b) rather than (a): Supposing that the decision theorist knows all the relevant costs, is there a logical level at which his recommendation is *reflectively consistent*, that is, is not upset by the underlying logic of the theory?

Now, the sort of stability property with which the optimizing theory of decision is or is not endowed can be made precise in two ways: (1) There is a logical level n where the $(n + 1)$-optimal decision endorses the n-optimal one; (2) there is a logical level n where the $(n + 1)$-optimal decision endorses the use of optimization at the n-level. That is, the $n + 1$-optimizing decision may be concerned with either n-level pointwise decisions and their costs or n-level decision criteria and their costs. If (1) is violated, the optimizing theorist's recommendation will oscillate in the action space as $n \to \infty$. An example of a violation of (2) is when the theorist shifts back and forth between optimization and satisficing when $n \to \infty$ (the induced recommended action will also no doubt oscillate in this case).

The above discussion should to some extent be familiar to computer scientists, because they obviously have to face the problem of optimal computations and can give a relatively nonarbitrary meaning to the elusive concept of higher-order costs. Strangely enough, economists have not paid due attention to the infinite regress of the optimizing theory of decision, despite the fact that they are the main proponents of this theory. (The few references in the economics literature are Göttinger 1982; Mongin & Walliser 1987; Winter 1975.) This is all the more surprising given that economists have worked out a "theory of search" (Stigler 1961) on the grounds that optimal decisions are, indeed, costly to make.

Discarding objections that could be raised against the shaky variant (a) of the problem, the following seemingly powerful counterargument remains: The infinite regression critique is irrelevant because it affects optimization and alternative theories in exactly the same way. This argument equivocates on the meaning of the critique. It is true that, for example, Simon's (1983) "satisficing" model gives rise to an infinite regress of its own. For any theory T that recommends d, to know that there were decision costs to d and that T is a theory of *rational* decision-making (rather than a nonnormative one) is enough to raise doubts as to whether d should have been recommended after all. Hence the infinite regress *itself* is by no means limited to optimization, but it may or may not converge, depending on the particular theory at hand. The results of the economist's "theory of search" as well as sheer common sense would suggest that convergence in the sense of (1) is more difficult to secure with optimizing than with nonoptimizing theories.

To be specific, Stigler's search model exhibits metaoptimal solutions that are typically different from the optimal solutions. This occurs because the model's monotonicity and regularity assumptions make it worthwhile for the agent to incur a "suboptimality cost" to lower his search cost. The simple trade-off argument can be repeated at any higher logical level. Clearly, it would not apply in the same way, or would not even apply at all, in the context of satisficing. In another example discussed by Mongin and Walliser (1987), a simple assumption connecting the complexity of decision rules with the cost of applying them is enough to destabilize the optimizing theory; that is, the infinite regresses to which it gives rise in this model typically do not converge in the sense of (1).