University of Utah Institutional Repository
Author Manuscript

Harman's Hardness Arguments[*]

by Elijah Millgram

From *Pacific Philosophical Quarterly* 72(3), September 1991: 181-202. Electronic posting with the kind permission of Blackwell Publishing.

In *Change in View*[1] Gilbert Harman produces arguments of the following pattern: Of two competing methods of belief revision, one is *too hard*; the other must therefore be the rational method. I will call arguments of this form *hardness arguments*. Hardness arguments are not, of course, peculiar to Harman; and considerations of this kind have recently become more popular in the philosophical literature.[2] But Harman's hardness arguments provide an object lesson in the pitfalls of deploying such considerations.

I will first examine two of Harman's hardness arguments, and argue that neither is sound.[3] Examining these arguments will give us a close look at how hardness arguments are supposed to work; they will also show how hardness arguments are harder to get off the ground than might have been supposed. I will then consider whether this form of argument is valid. After arguing that it is not, I will consider what is required if one is to deploy hardness arguments nonetheless. Finally, I will raise the question of whether arguments of this form require presuppositions that prove incoherent. I will suggest that hardness arguments are often misconceived, in that they ask a question that makes sense when one is *designing* hardware, but not when one *is* the hardware.

But before doing all these things, let me address a necessary preliminary: what "hard" means here. A problem is *easy* if you can do it in a reasonable amount of time and with a reasonable amount of effort, without overtaxing your memory, and so on.[4] A problem is *hard* if it's not easy; it's *hard in practice* but not *in principle* if you could solve the problem on a bigger, faster computer that

you could buy. (I mean hardness in practice to be sensitive to what resources are actually available to you when you're trying to solve the problem; one might say *impractical for the problem solver* instead.) A problem is *hard in principle* if buying a bigger, faster computer wouldn't help. (For instance, if buying a computer as big as the known universe, with as many circuits as there are estimated to be particles in the universe, and running it from the Big Bang until the heat death of the universe wouldn't solve the problem -- then it's hard in principle.) While for present philosophical purposes the precise location of the border between hard in principle and hard only in practice may be negotiable, a rule of thumb in the Computer Science community is that problems solvable using computational resources (i.e., time, space, and number of processors) that are polynomial in the size of the input are considered not hard in principle; problems that aren't, are hard in principle.[5] Here I'm going to treat this rule of thumb as delineating, at least roughly, this distinction.

1

Harman's first hardness argument is that belief revision should not work by modifying probabilities that are explicitly assigned to all of one's beliefs.[6] The reason is that this procedure is too hard, because updating the probabilities of one's beliefs using conditionalization would be hard in principle. Harman relies on Jeffrey's treatment, which provides an equation for assessing the impact of new evidence on a proposition A.[7] What is important for present purposes is that if one's system of beliefs has n elements, the equation suggests that in order to compute the new probability of A, one must do 2^n additions.

Harman concludes that "[i]f one is to be prepared for various possible conditionalizations, then for every proposition... one wants to update, one must have already assigned probabilities to various conjunctions of [it] together with one or more of the possible evidence propositions... this leads to a combinatorial explosion..." (p. 25). That is, in order to update the probability of A, you would need to know the probabilities of all 2^n conjunctions, and the prior probabilities of A given those conjunctions. This, thinks Harman, would

require exponential memory. Doing the number of calculations suggested by Jeffrey's equation would also take exponential time. Belief revision of this kind is therefore hard in principle.

Now it's a rule of thumb in computer science, or at any rate, in the AI part of it, that the general case of any interesting problem is intractable. But it doesn't follow from this that solving interesting problems is impossible. The reason is that the general case includes the worst case; but it's often unnecessary to deal with the worst case. By making reasonable assumptions, it's often possible to find an algorithm that will turn the trick for large classes of interesting cases. In particular, there's no reason to think that it's necessary to maintain the probabilities of 2^n conjunctions in this way in order to do a reasonable job of maintaining a belief system using probabilities.[8]

The question of how best to render problems like this computationally tractable is an active area of research; to illustrate the point I'll use one of the most prominent approaches, Bayesian networks.[9] Bayesian networks are directed acyclic graphs (DAGs) whose nodes represent propositions or beliefs, and whose arcs represent the (typically causal) dependence of a proposition or belief on its parents.[10] Associated with each node is a 'link matrix' giving the conditional probabilities of the node for the possible conjunctions of its parents; this table will be small provided that the number of parents is small.

Bayesian networks are a way to model local dependencies -- the fact that, typically, most of the beliefs one has are not directly relevant to a particular, given belief: if relevant at all, they are relevant only *through* their making more or less likely some other proposition, whose truth *is* directly relevant to that of the belief in question. For example, my beliefs about the corner grocery store are relevant to my beliefs about East Asian politics only indirectly: by way of the belief that the newspaper I buy from them is an authentic copy of *The New York Times*, and not a forgery of some kind. If the probability of this belief is fixed, I need no longer consider the bearing of the probabilities of my other beliefs about the grocery store on my beliefs about East Asian politics.

Which, and how many, computations one must perform (and which probabilities one must already know) is a matter of the structure of the graph. The structure of the graph represents the relevance of my beliefs to each other. A completely connected graph would represent a system of beliefs all of which were directly relevant to every other belief in the system; Harman would be right in thinking of a system of beliefs like this that it would be impractical to maintain by conditionalizing probabilities.[11] But our systems of belief are not usually like this; and when they are not, their structural features can be exploited to make updating them computationally tractable. For example, if a Bayesian network is singly-connected (i.e., tree-structured), then it can be updated to accommodate the impact of new evidence in time proportional to the longest path in the network; this is *quite* tractable.[12]

Harman, then, argues that belief revision cannot involve updating probabilities (and that belief, therefore, cannot be a matter of degree but must be all or nothing) on the grounds that such belief revision would be too hard, because hard in principle. But the argument is unsound, because belief revision using probabilities need not be hard in principle. In many cases of interest there are likely to be tractable algorithms that will allow one to revise one's beliefs using probabilities.[13]

2

After considering and rejecting probability-subscripted beliefs, Harman invokes another hardness argument to choose between *Coherence Theory* and *Foundations Theory*. Foundations Theory holds that the justification for holding a belief is a matter of the beliefs that support it, not of the fact that one does have that belief. Coherence Theory, on the other hand, holds that the mere fact that one has a belief is (some) justification for continuing to have it; and that belief revision is to be undertaken when it would lead to an increment in coherence that is worth the cost of making the change. [14]

Harman thinks that belief maintenance and revision that accorded with the

Foundations Theory would be too hard. The reason is that "[t]here is a limit to what one can remember, a limit to what one can retrieve. It is important to save room for important things and not clutter one's mind with a lot of unimportant matters" (pp. 42f). Since the Foundations Theory requires one to remember the reasons for acquiring each of one's beliefs, it requires that one clutter one's mind with "an incredible number of mostly perceptual original premises, along with many, many intermediate steps which one does not want and has little need to remember" (p. 43). Because belief revision that accords with the Foundations Theory is too hard, the Coherence Theory is to be preferred to it.

Now while keeping track of things like this is hard (for us), it is not hard in principle.[15] It is easy to program a computer to keep track of the reasons that supported its inferences and to engage in foundationalist belief revision[16] (when you delete a belief, follow the pointers down to beliefs it supported and, if they don't have other, independent support, delete them too).[17] Again, we cannot remember all the sensory evidence that supports our beliefs, and current computers would be hard-pressed to do so as well; but it is very likely that in a few years this won't be a problem for the computers. So while belief revision that conforms to Foundations Theory is hard in practice (for most human beings, at any rate), it is not hard in principle.

It follows that Harman's second hardness argument doesn't work. That is because Harman's argument derives its preference for Coherence Theory over Foundations Theory from its view that foundationalist belief revision is *harder* than coherentist belief revision. But this is a mistake. For (I will now argue) coherentist belief revision is (probably) hard in principle, which makes it harder than foundationalist belief revision, which is only hard in practice.

The "probably" inserted in the last sentence is due in part to Harman's failure to say just what coherence is. He says that "[f]or present purposes, I do not need to be too specific as to exactly what coherence involves, except to say it includes not only consistency but also a network of relations among one's beliefs, especially relations of implication and explanation" (p. 32).[18] Since Harman does not tell us exactly what Coherence Theory would have us do, we

will have to make an educated guess, leaving open the possibility that we have mistaken his intentions.[19]

Coherence Theory does tell us the following. When faced with the possibility of revising our beliefs, we should compare our unrevised system of beliefs to our system *with* the proposed revision (and maybe to the system with other revisions). We are to determine which of the two (or more) is more coherent, and roughly how much more coherent it is. On the basis of a further comparison of the gain in coherence with the size of the proposed revision, we are to choose whether to make the revision. The revision is to be made if the gain in coherence is worth a modification of that size.[20] For our purposes, it will suffice to note that this outline of a procedure for belief revision requires one to determine and compare the coherence of systems of belief.

Now Harman has not told us exactly what coherence is, but if we represent a system of beliefs as a graph whose nodes are beliefs and whose arcs are the "relations among [them], especially relations of implication and explanation", we shall not be straying too far from Harman's own talk of a "network" (p. 32).[21] Coherence will then be at least in part a matter of graph-theoretic properties of the graph, and determining the coherence of a system of beliefs will be tantamount to determining, among other things, those properties of the graph. Therefore, the difficulty of the task posed by Coherence Theory will be at least as great as the difficulty of determining the relevant graph-theoretic properties.

Just how hard that will be will depend on what the property in question is, but I suggest that for plausible candidates for the relevant property, the problem will be hard in principle.

As Harman says, "[i]f one's beliefs are coherent, they are mutually supporting" (33). This suggests that a graph-theoretic characterization of coherence will involve solving the following problem. Consider a path between nodes *A* and *B* in a graph. Every link in the path has some probability of failing: if *CD* is such a link in which *C* is evidential support for *D*, *C* might be

true, yet *D* be false, with some probability *p*. Now we can compute the reliability of the path easily enough. But we should be interested in the general support that *A* provides *B*, through *all* paths; the presumption is that a belief system in which beliefs give each other much reliable support is more coherent than one in which they support each other less. The problem of calculating this general support is known as *Network Reliability*. This problem is *very* hard: it is in fact *#P*-hard.[22] If this is what coherence involves, then coherentist belief revision is much harder than foundationalist belief revision.[23]

It seems very likely that any plausible notion of coherence will *involve*, e.g., Network Reliability, or some variation on it. We may tentatively conclude that determining the coherence of a system of beliefs is likely to be hard in principle. If it is hard in principle, then coherentist belief revision is harder than foundationalist belief revision, which is only hard in practice. And if this is the case, then Harman's second hardness argument is unsound, since its most important premise is that coherentist belief revision is easier than foundationalist belief revision.

If Coherence Theory requires determining the degree of coherence of a graph, it is not viable, because doing this is likely to be too hard. One possible way out of this difficulty might be to avoid having to determine the coherence of graphs in the first place, by finding a simple, local operation to perform on a graph that can be proved to increase coherence. One could then comply with the dictates of Coherence Theory merely by applying the operation. But this suggestion isn't very promising. If one were to model the acquisition of beliefs by constructing a graph using some inductive procedure that preserved some (very artificial) property at each step of the way, it might be possible to prove (inductively) that performing the local operation would increase coherence. But real systems of beliefs are unlikely to have such carefully-constructed and almost certainly fragile properties. An inductive proof of this kind is extremely unlikely for anything that realistically models actual systems of belief. Coherence is a global property, and the effects of a local operation on coherence will depend on features of distant parts of the graph. If the graph has evolved messily (and whose beliefs have not evolved messily?), there is no

telling what these might be. If one's system of beliefs has evolved messily, then applying a local operation could, for all one knows, *decrease* the coherence of that system.[24]

More generally, a computationally tractable technique for determining the coherence of a system of beliefs represented by a graph can succeed only by exploiting structural features that not all graphs have. In order to make it plausible that there are such computationally tractable techniques, it would be necessary to identify the structural features that can be so exploited. Harman has not suggested what those features might be, nor has he given us reason to think that naturally occurring systems of beliefs have them.

Harman's rejection of Foundations Theory in favor of Coherence Theory on the basis of the greater difficulty of conforming to the former is, at best, premature. If anything, it seems that on a plausible construal of coherence, Coherence Theory will turn out to be *much* harder than Foundations Theory: that, unlike Foundations Theory, it will turn out to be hard in principle. Now since Harman has not given us a precise specification of coherence, it may be that he (or someone else) could produce a definition of coherence that would be both a plausible rendering of what is meant by coherence of belief systems and also computationally tractable. But I do not expect that producing a specification of coherence that meets both of these conditions will be a trivial matter (in fact, I suspect it to be impossible); I think that Harman's second hardness argument should be considered unsound pending such a specification.

There is a moral to be drawn here. Hardness arguments proceed from from premises of the form, 'Method M is too hard' (or, 'Method M_1 is harder than M_2'). But premises like these are harder to come by than they might seem.

Problems can turn out to be pragmatically easier than they had seemed: approximations may be available, or the worst case may not prove to be the normal case. It is very difficult to show that acceptable approximations to a precise problem solution are not available, if only because 'approximation' is such a flexible notion, one that is sensitive to pragmatic considerations. An

approximating algorithm might arrive at a correct answer most of the time, or in most of the cases that are in practice likely to arise, or it might arrive at an answer close enough to the precise value to be usable. What would serve as an adequate approximation will depend on the purposes for which it is required. Many very disparate techniques may count as acceptable approximations to a given technique. Moreover, problems can, on closer consideration, turn out to be harder than they had seemed; the comparative judgments needed for a hardness argument to get off the ground will often prove mistaken. If Harman's experience is any indication, sound hardness arguments are going to be hard to come by.

3

I have argued that the two hardness arguments that Harman presents are (very probably) unsound; and that this is indicative of a general problem in constructing hardness arguments. But this does not impugn the *form* of the argument; arguments can be unsound but valid, and for all I have said so far, there could be other arguments of this form that were both sound and valid. However, I am going to argue that hardness arguments, in the form that Harman gives them, are not valid after all.

To see why they are not, consider an objection that might be made to my earlier treatment of Harman's hardness arguments. It might be objected that my treatment of those argument was inconsistent. Against Harman's claim that using degrees of belief was hard in principle, I argued that there might be computationally tractable algorithms that were close enough to those Harman had in mind to serve as a reasonable approximation to them. But against his proposed replacement of Foundations Theory by Coherence Theory, I argued that algorithms that check the coherence of a graph are likely to be computationally intractable. But couldn't there be approximations to those algorithms that aren't intractable?

Lacking a definition of coherence, it's difficult to be certain; and as I remarked, it's difficult to find convincing arguments that a hard algorithm has no

approximations.[25] But the issue can be sidestepped, because the objection that there might be an approximation to an algorithm that checks for coherence is not one that someone who has embarked on Harman's project is in a position to make. Let C be an algorithm that checks for coherence: something is coherent iff (or, to the degree that) C says it is. Suppose that it is claimed that C is too hard, but some approximating algorithm A exists that is not. Presumably, for some inputs, A produces outputs that diverge from those produced by C. Define 'shmoherence' to be the property checked for by A, the algorithm otherwise characterized as approximating coherence: something is shmoherent iff (or, to the degree that) A says it is. Recall that Harman attempted to argue that a Foundations Theory algorithm would be too hard; an approximating algorithm (which checks for coherence) should therefore be used instead. Harman takes this to be an argument for Coherence Theory and against Foundations Theory. Now we've argued that an algorithm C that checks for coherence is too hard. An approximating algorithm, A, is proposed instead. But A checks for shmoherence. By parity of reasoning, this should be an argument for Shmoherence Theory, and against Coherence Theory.[26] In short, defending the view that Coherence Theory is acceptable because there are tractable approximations to the too-hard algorithms it involves is not an open option for Harman.

Why not? Hardness arguments involve what we can think of as a rule of inference, which I will call *the hardness rule*: if a method of belief revision is too hard, then it is rational to adopt an easier alternative to it. (Not, of course, *any* easier alternative; rationality demands something like adopting the *best* of the sufficiently easy alternatives.) In this way, a method of belief revision which Harman concedes to be second-best[27] can be excused on the basis of our computational limitations, and thereafter incorporated into a normative theory of rationality.

Harman's applications of the hardness rule have the following form. A problem P_1 is specified with a greater or lesser degree of formality; solving this problem in appropriate circumstances has been taken to a requirement of rationality. (Examples of such problems are belief revision using conditionalization and

foundational belief revision.) A computational model M_1 that allows P_1 to be solved is implicitly or explicitly associated with it. (The default model seems to be something like an infinitely fast Turing machine.) M_1 is rejected as being too idealized; closer consideration of problem solvers' capabilities is used to produce a new computational model M_2 of greater descriptive accuracy. It is argued that P_1 cannot be solved by M_2. Finally, the requirement that P_1 be solved is rejected, and P_1 is replaced by an alternative problem P_2. P_2 is formulated with an eye on both P_1 and M_2; in the instances we have before us, solutions to P_2 can be construed as approximate solutions to P_1 that are feasible for M_2. (For example, foundational belief revision is rejected on the grounds that a more descriptively accurate model of human capabilities, one that takes into account their small finite memories, shows that people cannot be expected to perform foundational belief revision. Coherence-based belief revision is proposed (mistakenly, I argued in section 2) as a feasible alternative.) This is what applying the hardness rule consists in.[28]

Now a rule of inference is *valid* if its application to true premises *always* produces a true conclusion. It follows that the hardness rule is not valid as it stands. For, I will now argue, uniformly applying the hardness rule would have the consequence that *whatever* one did would be rational; and this is absurd. The argument will expose two underlying problems. The first is that of the emptiness of the formal notion of a computational model, due to which uniform application of the hardness rule would prevent ascriptions of rationality from playing certain characteristic explanatory roles. The second, and perhaps the more interesting, is that uniform application of the hardness rule would prevent the theorist from distinguishing between descriptive and normative accounts of thought; but this distinction is essential to something's being a theory of rationality at all, and, moreover, is accordingly presupposed by the hardness rule itself.

The problem is this. *Whenever* an agent makes a suboptimal choice (with

respect to given goals and information), it is possible to apply the hardness rule, and thereby to construe the optimal choice (or, for that matter, *any* more optimal choice) as having been too hard for the agent. Specifically, it is always possible in principle to produce a computational model of the agent that has two features: it has greater descriptive accuracy than the background model that allows the agent to make the optimal choice; and the model holds that it is not possible for the agent to make the optimal choice.

An example may make the point clearer. Suppose that John has a math block, and doesn't think clearly when hungry or flustered; for these reasons, he fails his mathematics final. In principle, researchers could produce a computational model of John that reconstructed these shortcomings. Of course, we are far from being able to actually produce such complex and detailed models. Today, the kind of considerations Harman adduces, such as restrictions on memory size and on the ways memory can be accessed, are fairly close to state-of-the-art. But the point is that such models are available in principle. Our hypothetical model of John would have greater descriptive accuracy than a similar model that merely reflected general features of human psychology and general human capabilities. But this computational model explains why John could not solve the problems on his mathematics final. (If the model were definite enough to be turned into a computer program, the progam would fail to solve the problems on John's exam.) Applying the hardness rule, we conclude that the mathematics final was too hard for him. So far, so good; if he failed, maybe the exam *was* too hard for him. But the hardness rule is used to license the conclusion that the procedures that John used to produce his mistaken answers were *rational* (for him). This is, to put it mildly, an unintuitive conclusion.

The difficulty is more systematic than an occasional violation of our intuitions. Nothing in the example turned on the particular nature of John's failings; the point was that *whatever* the reasons for his making a suboptimal choice, these could be modeled computationally. (We can identify the reason for this as the emptiness of the merely formal notion of a computational model: a computational model of a system is merely a formal description of that system

that specifies its transitions from one state to another.) The consequence of this is that applying the hardness rule uniformly would have the result that *whatever* anybody actually did would be rational. This is another unintuitive conclusion; and again, the difficulty is more systematic than its merely being unintuitive.

First, giving up the distinction between what the agent would have been rational to do, and what the agent actually did, would prevent ascriptions of rationality from playing a characteristic explanatory role. For example, it would render unintelligible actions taken to improve rationality (such as mental exercises, taking classes in game theory, eating fish because it's 'brain food', and so on): if whatever one does is *already* rational (since any suboptimality must be due to some computational deficit), how can any action be taken to render one *more* rational? Taken to a further extreme, we will be unable to say that we are more rational than, say, clams; and this is something we want to be able to say in order to be able to provide the kind of explanations we might want to give in the context of evolutionary biology: a particular trait (say, having a bigger brain) may be adaptive because it enables organisms that have it to be more rational.

Second, theories of rationality require a distinction between normative and descriptive, if they are to be considered theories of rationality at all. This appeal to the critical or normative aspect of rationality is intrinsic to Harman's hardness arguments, which purport to tell us what we *should* or *may* rationally do. The appeal is unanswered if the response to it is: anything.[29]

Now the way I have presented the problem may give the appearance of a slippery slope argument. I characterized applying the hardness rule as involving the rejection of a comparatively idealized model of human reasoning in favor of one that more adequately reflects the capacities and circumstances of actual persons. But after any move from one level of idealization to a second, lesser level, one can ask: why stop here? Having reflected in a computational model the fact, say, that humans have small finite memories, why not, when considering a particular individual, model the peculiarities of *his* memory? The hardness rule, as it stands, does not tell us why we should not

eliminate idealization after idealization until descriptive accuracy is reached: in fact, as it stands, it invites us to do precisely that.

Arguments that turn into slippery slopes aren't *valid*. That doesn't mean they shouldn't be used. 'Removing one hair from a head that is not bald will not make the head bald' is a basis for invalid arguments. But it is nonetheless a reasonable rule of thumb, when applied with caution.

What must be required of an argument that leads onto a slippery slope? Since the hardness rule can only be applied some of the time, we should, in order to prevent its misapplication, be able to say *when* it is appropriately applied, and when not: if we cannot, we will not know whether a conclusion resulting from a particular application of the rule is correct. We must have some clear idea of what considerations would prompt us to get off the slope; and where (roughly where, if the key concept is vague) we should get off because we have slid too far. Moreover, being able to say when the rule is applicable is a test of our understanding of the considerations that underwrite the rule. If we cannot say (roughly) when the rule is applicable, we may presume that we do not really understand the reasons for having such a rule at all.

4

I have argued that because hardness arguments turn into slippery slopes, using a hardness argument requires that one should know how to get off the slope. That is, one must know (roughly) when 'hard' is 'too hard' -- and why. In this section, I will briefly discuss the kind of considerations to which one can appeal, with an eye to showing that they cannot simply be taken for granted.

In ordinary usage, 'too hard' prompts the question, 'too hard for what?' Our judgments of hardness invoke our interests and concerns. Consider, for example, a high school senior who is assigned a particularly difficult problem in his textbook (normally used in college classes) by mistake. He fails to solve the problem, but is excused: the problem, the teacher acknowledges, was too hard; he couldn't have done it; after all, he's only in high school.

The summer passes, and the same student is now a college freshman. Three months have made no appreciable difference in his talents; for the purposes of the example, we shall assume that he is as similar to his earlier self as you like. Once again, he is assigned the problem, and once again he fails to solve it. This time, however, he is not excused on the basis of the problem's being too hard. When he complains, his teacher tells him that he could have done it if he had begun studying earlier, if he had gotten a tutor, if he had tried harder, if he had thought more carefully about it. The very same problem is at one time too hard, at another time not too hard -- all without any change in the cognitive capacities of the problem-solver.

There is, of course, one clear sense in which the problem was too hard in both cases: the student failed to solve it both times. But this is evidently not the relevant sense of 'too hard', since, as examples like this one show, we support judgments of hardness by appealing to counterfactuals: 'He could have done it if only he had... '; 'He couldn't have done it even if...' Now notice *any* task is such that there is *some* contrary-to-fact assumption on which it could have been done. (He *could* have solved the problem if a futuristic brain operation had made him much smarter, or if he had applied himself to mathematics for the last ten years, or if a miracle had happened.[30]) What picks out the relevant counterfactuals, and the reference classes (high school seniors, freshmen) associated with them? In this case, the answer is clear: the expectations of the instructor, as shaped by his institutional role.[31]

Evidently, saying how one gets off the slippery slope generated by hardness arguments is in part a matter of making explicit the concerns and interests that underwrite particular judgments that something is too hard; this is what one needs to be clear about if one is deploy the form of argument. Now I do not wish to take a stand on whether Harman does in fact have a clear understanding of the concerns and interests that underwrite his hardness arguments. But he fails, in any case, to make them explicit. He does not articulate the background concerns that motivate his choices of computational model, and he does not make clear how these concerns motivate those choices. This is unfortunate, for

we cannot expect to be able to assess a hardness argument if such central considerations are suppressed. The omission is an example of a further pitfall associated with hardness arguments -- the temptation to take for granted the considerations that determine when 'hard' is 'too hard'.

Now it might be thought that when *rationality* is at issue, the background of interests and concerns *can* be taken for granted. As I have had it put to me, if a procedure is too hard for the human organism as such, it can't be what rationality demands. And I grant that there is something very plausible-sounding about the claim. But even if it is true, it is still not evident *why* it is true. And, I will suggest, there are reasons to think that 'too hard for the human organism as such' is not the place to get off the slippery slope.[32]

It is not at all clear to me why our rationality-related interests should follow species lines. There are, first of all, species-wide cognitive failures that we would not want to write into rationality.[33] There evidently *are* procedures that are too hard for the human organism that we *do* insist are demanded by rationality. Second, at the risk of entering the realm of science-fiction, we may imagine a society in which humans mingle with other species, some with somewhat greater and some with somewhat lesser cognitive abilities. Is it at all obvious that in such a society the abilities of the human organism (or of any particular *kind* of organism) would present themselves as the standard for evaluating the demands of rationality?

It might be thought that 'ought implies can', and that it follows from this dictum that if human beings can't perform a particular task, then rationality cannot insist that they ought to perform it. Indeed, I am inclined to think that this why the capabilities of the human organism seem like a plausible place to get off the slope. After all, there's nobody here but us humans, and if humans can't do, then no one can do it. And if no one *can* do it, then no one *ought* to do it. But rationality is a normative concept; that is, if it's rational to do, you ought to do it. So if you oughtn't to do it, then it can't be rational to do.

But we have seen that if we are to slide to the bottom of the slippery slope,

there are nonetheless tasks that a particular individual is unable to perform such that this failure supports the accusation of irrationality. Why does 'ought' imply 'can' at the level of the species, but not at the level of the individual? In any case, the application of the dictum is unclear. If I know that I cannot succeed at some task, then perhaps it is irrational for me to *try* to do it. But it does not follow from this that I am rational if I do not perform the task. (For example, if I know I cannot think clearly, then it is pointless for me to try. But it does not follow that I can be rational even if I do not think clearly.) Not all performances are tryings.

I do not think that I have presented a knockdown argument against the claim that the abilities of the human organism mark the point at which to get off the slippery slope. But the burden of proof is squarely on the shoulders of this position's proponent. If the line is to be drawn here, we must be given some reason to think that here is where it is to be drawn; and such reason is conspicuously absent. Because I think that judgments about what cognitive processes are too hard to be demanded by rationality are continuous with other judgments about what is too hard, I myself am inclined to the view that the line is drawn in no particular place, because there is no one line. Different lines are drawn in different places on different occasions, and consequently each hardness argument owes its audience an explicit accounting of the interests and concerns that give the judgments of hardness their content, and, when hardness arguments are being used to develop a theory of rational cognition, a further explanation of how these judgments are relevant to the theory of rationality being advanced. As we have seen, the first of these demands has not been met; and as we will see in the next section, there is reason to think that the second will not be.

5

Beyond the immediate questions about the soundness and validity of hardness arguments there is a deeper issue, that of the intelligibility of their deployment to address questions of rationality. The problem can be raised in the following way. Hardness arguments are a response to the question 'How should I think?'

when this question is construed as 'How can I most adequately solve (given) problems with my limited computational resources?' But the question 'How can I best deploy my limited computational resources?' seems to presuppose that my computational resources are something distinct from myself, something I own and use the way I might own and use a computer. The picture it evokes is one in which I am standing *behind* my mind, as it were; one in which my mind is *external* to me.[34] But of course I am not distinct from the computational resources we are considering. I am not somehow standing behind my mind, operating it the way I might operate a computer. Indeed, on the view I take to be Harman's, I *am* just the aggregate of those computational resources; that is why Harman finds hardness arguments plausible in the first place.

Whence this peculiar image? To answer this question, we should consider where hardness arguments come from in the first place. Philosophers occasionally embark on expeditions beyond the borders of philosophy departments; hardness arguments were first brought back from a foray into the neighboring discipline of computer science, where the practical need to build working machines and write working code made computational complexity an issue. In this domain, hardness considerations are used to justify design and implementation decisions. An engineer is faced with the problem of designing a device to perform a given task. We must imagine the designer beginning with a specification of the performance of his *ideal* device, the device as he would like it to work. (We may suppose this specification to take the form of a function from the device's inputs to its outputs.) Because of the limitations of the material with which he has to work (or for economic reasons, or whatever), the engineer may decide to cut corners; and here hardness considerations may play a decisive role. For example, given that he cannot build an infinitely fast device, the computational complexity of the task the device is to perform may mandate using a fast and dirty heuristic instead of inplementing the algorithm that the engineer has determined is too hard. That is, the engineer may choose to construct a device that, for some acceptably small range of inputs, produces incorrect outputs. What counts as too hard, and what counts as an acceptable substitute for the ideal device, is determined by the concerns of the engineer.

Now in such circumstances, we may assess the engineer's decision as rational or irrational. But we do not assess the *device* as being rational or irrational: the term of approbation, "rational", sticks to the engineer and not to his creation. Moreover, the deviations from the ideal specification that the engineer is rational to tolerate remain errors. They do not cease to be errors because they resulted from a rational design decision, precisely because the design decision was a decision to tolerate those errors. Hardness arguments, in short, underwrite ascriptions of rationality to engineers, not the devices they construct.

Consider now Harman's attempted adaptations of hardness arguments to underwrite ascriptions of rationality to human beings. It is the human beings who are taking the computational shortcuts, so we are to think of human beings as the 'devices.' But we saw that hardness arguments do not warrant ascriptions of rationality to the devices. Are humans also thought to be occupying the role of engineer? There may be cases for which this is a not unreasonable construal; choosing the way in which one will solve a division problem (long or short division), or choosing what habits one will attempt to acquire come to mind. But the hardness arguments Harman presents do not address cases of this kind. I do not *choose* whether or not to use probabilities in adjusting my beliefs (except in special cases); and I do not *choose* whether to keep track of the evidence for my beliefs and engage in foundational belief revision (again, except in special cases).[35] The cognitive processes in question are built in, *hardwired*; when we employ this algorithm as opposed to that, that is just the way we *do* do it.

We are not in these respects our own engineers. It follows that hardness arguments do not in any straightforward way make the term of approbation "rational" stick to us. If God designed us, perhaps hardness arguments can make the word stick to him; if no one designed us, then perhaps hardness arguments make the word stick to nobody. Hardness arguments allow one to conclude that this or that choice is rational *for the designer*; they warrant no such conclusions regarding the designed devices.

In deploying hardness considerations, it is important not to slip into imagining that one is, as it were, one's own designer; one may then conflate the rationality of the design decision with one's own rationality. 'If I were Mother Nature, I would have been rational to design myself to make these mistakes; therefore, I am rational, and they aren't really mistakes' is confused reasoning.

6

I am inclined to think that there is something importantly right about hardness arguments. The philosophical tradition from which Harman is, commendably, departing is one that understands rationality in a way that has only the most attenuated connection to human capabilities and practices. Surely it is somehow an important aspect of rationality that one's methods of reasoning and cognitive resources roughly match.

Nonetheless, hardness arguments are extremely hard to get right. Partly this is because it is hard to get their premises right. But partly it is because the argument form itself is not valid, and can accordingly not be deployed without appealing directly to the considerations that underwrite it. These considerations are poorly understood, and the matter is not helped by the tendency to treat hardness arguments directed towards what is and is not rational as of a piece with the design and implementation decisions made by computer scientists; eliding the differences between the two produces a confidence regarding the former that is entirely unwarranted.

Notes

1. Page references in the text will be to Harman, 1986.[back]

2. For example, Cherniak, 1986, is an extended instance of argument of this kind. See also Stich, 1990, pp. 28, 151-58, and Goldman, 1986, esp. ch. 6. Arguments of this kind have a longer tenure in the economics literature; cf., e.g., Simon, 1979. [back]

3. Other hardness arguments may be implicit in Harman's discussion of the Principle of Clutter Avoidance and the Logical Inconsistency Principle (pp. 12-15), and the reasoning supporting his claim that "[t]here are practical reasons to minimize change in one's view" (p. 63). Such considerations are explicit in his discussion of 'holism' in practical reasoning (pp. 98ff; cf. esp. pp. 101, 105-107, 112). But of these, only the 'holism' argument receives reasonably full treatment, and its main point (insofar as it is a hardness argument) is merely an application of the conclusion of one of the two arguments I will discuss. [back]

4. 'Reasonable' is being left an imprecisely defined notion for the present; possibly it varies from person to person and situation to situation. (More on this in sections 3 and 4.) [back]

5. In particular cases, this may need qualification. There are cases where it's rational to use NP-hard algorithms, for example, when they are suitable for the range of cases one is actually liable to encounter. (Cherniak, 1986, p. 93, presents the simplex algorithm as an illustration of this point, though here it is, importantly, not simply the *size* of the input that is the problem.) For an accessible introduction to this material, see Garey and Johnson, 1979. [back]

6. Pp. 25-27; cf. also p. 115. He excepts "a few special cases of beliefs that are explicitly beliefs about probabilities" (p. 22). [back]

7. The equation is

$$P(A) = \sum_{i=1}^{2^n} p(A \mid C_i) P(C_i)$$

where $P(A)$ is the new probability of A, $p(A)$ is the old probability of A, and the C_i are all possible conjunctions formed from the n propositions making up one's beliefs and their negations; there will be 2^n such conjunctions. (Jeffrey, 1983, pp. 172f; I have slightly modified the notation.) [back]

8. It's suggestive that Jeffrey's illustration of the technique (1983, p. 173) in fact requires not 2^n (=8) atoms but only n (=3). [back]

9. Another is the maximum entropy distribution approach. Maximum entropy distributions can be stored in space that is linear in the number of constraints, as opposed to the exponential space generally required by probability distributions (Goldman, 1987, pp. 14f). Goldman provides a technique for computing maximum entropy distributions that is efficient when the hypergraph that models the variables and constraints is not too connected (p. 32). Cf. also Goldman and Rivest, 1988. [back]

10. A *graph* is a set of *nodes* (or *vertices*), and a set of *edges* (or *arcs*), each of which is a pair of nodes. Figures (1) and (2) are representations of graphs: the dots represent nodes, and the lines between the dots represent edges. A *directed graph* can be thought of as having arrows for its edges. For a survey of this material, see Tutte, 1984.
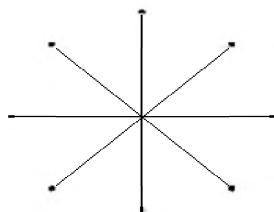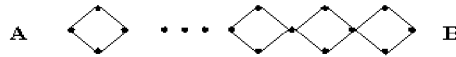
Figure 1

Figure 2

Merely associative dependencies can be represented using Markov networks, which are undirected graphs. For further treatment of Bayesian and Markov networks, see Pearl, 1988. [back]

11. Cf. Cooper, forthcoming. [back]

12. This is, of course, only an example; it is unlikely that most interesting systems of belief can be represented by singly-connected graphs. Techniques exploiting other features of graphs will be required in realistic cases. For other recent work in the field, see Beinlich *et al.*, 1989; Chavez, 1989; Chavez and Cooper, 1989a; Chavez and Cooper, 1989b; Lauritzen and Spiegelhalter, 1988. [back]

13. Employing these algorithms may still be hard in practice, even if not hard in principle; we have seen no argument on this score, one way or the other. But Harman's argument proceeds from the claim that belief revision using probabilities is hard in principle, not from the claim that it is hard in practice.

Recall that hardness in practice is sensitive to what resources one actually has available. So evidence that Bayesian inference engines can actually be constructed (see, e.g., Agogino *et al.*, 1988) does not yet tell us that belief revision using probabilities is not hard in practice for *humans*. [back]

14. Harman distinguishes the views he calls by those names from conventionally so-called foundationalism and coherence theory, but the resemblances will be evident to readers familiar with the traditional positions. [back]

15. Harman acknowledges this on p. 42; cf. also p. 10. [back]

16. Truth Maintenance Systems (TMSs) typically do this. See, e.g., Doyle, 1981. [back]

17. This procedure doesn't address the question of whether the belief is still *supportable*, i.e., if it *could* be inferred from other beliefs you have. It just checks whether you have performed an inference (rehearsed a justification) that still supports it. But the foundational procedure Harman suggests doesn't address this question either (pp. 30f). [back]

18. Just in passing: Harman's claim that coherence includes consistency suggests that one must check one's beliefs for consistency; in the truth-functional case, this is SAT, the original NP-complete problem. We may presume that checking the consistency of actual systems of belief, which are not merely truth-functional arrangements of atomic propositions, would be harder. So if coherence includes consistency, coherentist belief revision is almost certainly hard in principle. [back]

19. We may remark on what seems to be another assumption of Harman's: that Coherence Theory and Foundations Theory jointly exhaust the range of options regarding belief revision. This assumption is necessary for his inference from the hardness of foundationalist belief revision to the acceptability of coherentist belief revision. Now sometimes philosophers write as though foundationalism and coherentism were the only two epistemological options. But the two theories would *logically* exhaust the field of possibilities only if one were simply the complement of the other: if the sole content of Coherence Theory, for instance, were that Foundations Theory was wrong. If that were the appropriate understanding of Coherence Theory, however, hardness arguments would not work, if only because Coherence Theory would not then provide a *method* of belief revision: If Foundations Theory dictates use of method M_f, Coherence Theory would say no more than "Don't use M_f"; but "Don't use M_f" is not a method of belief revision. Now if Coherence Theory is not the logical complement of Foundations Theory, a hardness argument for Coherence Theory proceeding from the premise that Foundations Theory is too hard cannot exclude *a priori* the possibility that some third theory is to be preferred. So substantive argument is required to rule out alternative methods of belief revision. Harman does not give

such argument; what I wish to emphasize for future reference is the degree to which such argument would complicate the deployment of a comparative hardness argument. (Because the subject of this paper is hardness arguments, and not methods of belief revision, I will not here consider what plausible alternatives to Foundations Theory and Coherence Theory might look like.)

Moreover, for hardness arguments of this kind to work, two further premises must be available. The first is comparative, of the form 'M_1 is harder than M_2'; the second, to the effect that 'M_2 is not too hard'. Arguing that the method dictated by one of two theories is too hard does not by itself settle the question in favor of the second method; it may be too hard as well. (There is no *a priori* guarantee that *any* method is acceptable: it may be that some problems we face cannot be adequately handled by *any* technique.) [back]

20. Cf., e.g., pp. 30, 32. Harman discusses how to estimate the sizes of proposed modifications at pp. 59ff. [back]

21. We can also accommodate Harman's claim that "some explanatory connections are immediate, whereas others are indirect" (p. 66), in this way: one belief's immediately explaining another can be represented in the graph by connecting the nodes that represent those beliefs by an edge. [back]

22. Problems which are *#P*-hard are thought to be harder than those that are merely NP-hard. For a discussion of *#P* and a survey of some problems in it, see Valiant, 1979. [back]

23. I will review a few more possibilities to give some of the flavor of the problem. A large number of paths between nodes might be suggested by remarks Harman makes on p. 67; this is computationally tractable, but allows graphs like that in figure 2: there are *many* paths between *A* and *B*, but the graph represented in figure 1 does not seem to be a representation of a highly coherent system of beliefs. (High connectivity -- requiring a high number of internally node-disjoint paths between any two nodes, instead of requiring a large number of possibly overlapping paths -- would rule out the counterexample in figure 2. But connectivity as usually defined is insufficiently robust: on the standard definition, a very coherent graph (one with high connectivity) could be reduced to very low connectivity (1), just by adding a 'tail' (a node connected by only one edge), but the addition of a single belief that this represents surely does not so drastically reduce the coherence of a system of beliefs. And even if this problem is somehow circumvented, the disjointness requirement means that of two paths that even partly overlap, only one will be counted; so that high connectivity

does not adequately reflect the notion of coherence we are trying to capture. Finally, the strength of the connections is not taken into account; when this is done, the problem turns into Network Reliability.)

Alternatively, one could try to minimize distance between pairs of nodes. This is not hard in principle: it can be done in, at worst, $O(n^3)$. But this would end up classifying graphs like the one in figure 1 as very coherent (no two nodes in it are more than two arcs apart); and this is not what we mean by coherence. (While I will not try to specify just what it is about these graphs that makes it seem implausible that they represent highly coherent systems of belief, it is curious that both seem to contain bottlenecks.)

Yet another possibility might be to try to capture the notion of relative coherence in terms of minimum covering by cliques. A clique is a set of nodes such that each is connected to every other. A graph is covered by a set of cliques if every edge is in a clique in that set. In looking for a property that would characterize coherent systems of beliefs, one might notice that they are typically made up of large groups of beliefs whose members are very tightly related to each other, while the groups of beliefs themselves are more loosely related. Comparing two graphs to see which of them more closely matches this pattern would involve finding a graph-theoretic property to express it. Cliques look promising; the fewer of them it takes to cover the graph, the better. Unfortunately, the problem of covering a graph with the smallest possible number of cliques is hard in principle: it is NP-complete (Garey and Johnson, 1979, p. 194). Of course, not all beliefs in these tightly related groups need be 'immediately adjacent' to each other. To deal with this, one can extend the notion of covering a graph with cliques to that of covering it with subgraphs of low diameter. The diameter of a subgraph is the worst shortest-path distance between any two nodes in it. So, for example, a subgraph of diameter 2 contains nodes each of which can be reached from any other in two steps. It is easy to reduce covering by cliques to this problem, which is thereby shown to be hard as well.

What this very brief survey of candidates suggests is that they fall into two categories: the computationally tractable candidates, which are not even remotely what we mean when we talk about the coherence of systems of beliefs, and the candidates which are on the way to capturing (even if they do not entirely capture) what we mean by coherence, but which are computationally intractable. [back]

24. It might be suggested that we should be trying to increase *local* (rather than global) coherence. (For pressing this point, I'm grateful to Gary Ebbs and Ed Stein.) Since the subject of the present paper is hardness arguments rather than coherence theory, I will refrain from assessing this possibility in detail; we can remark in passing that there will be

difficulties in making out the relevant notion of locality, and problems with motivating the view: why should one *want* to increase specifically local coherence? For present purposes, it suffices that this is not Harman's view: "Whether... a belief is justified depends on how well it fits together with *everything else* one believes" (pp. 32f, my emphasis). [back]

25. I suggested a moment ago that coherence almost certainly involves Network Reliability, and to the best of my knowledge, no one has figured out a way to approximate it. But that doesn't mean that there's any guarantee that someone can't come along with an approximation tomorrow. Nevertheless, the proof of the pudding is very much in the eating: it's unclear how seriously we have to take an objection that runs, there *might* be an approximating algorithm -- but does not actually propose any such thing. [back]

26. A move structurally similar to this one is in fact made by Harman, for his Principle of Positive Undermining and against the Principle of Negative Undermining; foundationalists take the former to be an approximation to the latter. Cf. pp. 39f. [back]

27. Harman presents Coherence Theory as an approximation to Foundations Theory. If one had unlimited computational resources, it would always be right to behave in accordance with the dictates of Foundations Theory; but because our resources are limited, we must get by with mere coherence. "If one had unlimited powers of record keeping and an unlimited ability to survey ever more complex structures of argument, replies, rebuttals, and so on, it would be rational always to accept things only tentatively as working hypotheses, never ending inquiry" (p. 50). But, "something like the foundations theory applies to what one tentatively accepts as a working hypothesis" (p. 49). (For further textual indications that this is in fact Harman's view, see his description of "extensive use of probabilities" as "too complicated for mere finite beings" (p. 27), the implication being that *other* beings might do better to rely on probabilities; his remark that "[b]ecause of one's finite and limited powers, [53] one cannot escape the first three commitments except in a few instances when one is able tentatively to accept certain propositions merely as working hypotheses" (p. 52f); and the attribution of the necessity of Clutter Avoidance to "the limitations of finiteness" (p. 56).) [back]

28. It might be suggested that a precise formulation of the hardness rule is in order here. But I will forego producing one, for two reasons. First, as we will see, it's not necessary for the argument. Second, an artificial precision would involve picking a specific version of the view I'm attributing to Harman -- an unnecessary exegetical risk, and one that would tend to turn a stalking horse into a straw man. [back]

29. In *Change in View* this difficulty remains in the background. First, Harman does not apply the hardness rule everywhere one might; consequently, the question of what would happen if it were applied everywhere can be overlooked. Second, Harman is unenthusiastic about distinguishing descriptive from normative theories of rationality; he "find[s] it hard to say whether the theory [he] want[s] is a *normative* theory or a *descriptive* theory": the two kinds of theory "are intimately related... it is hard to come up with convincing normative principles except by considering how people actually do reason... [and] any descriptive theory must involve a certain amount of idealization, [which is] always normative to some extent" (p. 7). [back]

30. There are counterfactuals of this form even when NP-hard problems are at issue. One can imagine physical laws that would permit the construction of Huffman's lamp devices; such devices could solve NP-hard problems rapidly. Even for an NP-complete problem, it is possible to produce the counterfactual: if the laws of physics had been different, he would have been able to do it. [back]

31. This does not mean that there is an 'institutional' sense of 'too hard' that is to be contrasted to a more central 'capabilities' sense. Suppose I am attempting to assess a person's capabilities by giving him a task which he fails to complete because a tornado sweeps through the town. Normally, I do not conclude that it was beyond his abilities, that it was too hard. Rather, I say that (perhaps) he could have done it if the tornado had not prevented him. This is because the task is conceived in one way rather than another: if the task had been conceived as an exercise in performing under extreme stress, I might have said instead that he couldn't do it, that it was too hard for him. [back]

32. I have also had it suggested to me that what matters is whether a procedure is too hard for a *rational* being. But this is a red herring. One cannot determine what it is rational to do by determining what is too hard, if one's determinations of what is too hard require already knowing what is rational. Notice also that it is unclear that the appeal to rationality to determine what is too hard would classify as too hard problems that are merely computationally difficult, since it might be said that *had* the agent been more rational (something which would have involved having greater computational abilities) he would have gotten the right answer. [back]

33. See Nisbet and Ross, 1980, for a depressing survey of some of these. [back]

34. Candace Vogler's critique of Elster's proposals for managing *akrasia* and self-deception

prompted this discussion; she calls the person so conceived an 'angel in disguise'. I'm grateful to her for allowing me to read a draft version of her paper "Imperfect Rationality and the Ghost in the Globally Maximizing Machine". [back]

35. Harman concurs: he does not "want to suggest that one ever makes *conscious* use of principles of revision... it may well be that reasoning is a relatively automatic process whose outcome is not under one's control" (p. 2). [back]

References

Agogino, A. M., Srinivas, S., and Schneider, K., 1988. "Multiple Sensor Expert System for Diagnostic Reasoning, Monitoring and Control of Mechanical Systems." *Mechanical Systems and Signal Processing*, 20(2).

Beinlich, I. A., Suermondt, H. J., Chavez, R. M., and Cooper, G. F., 1989. "The Alarm Monitoring System: A Case Study with Two Probabilistic Inference Techniques for Belief Networks." Stanford University Knowledge Systems Laboratory, Working Paper KSL-88-84.

Chavez, R. M., 1989. "Hypermedia and Randomized Algorithms for Medical Expert Systems." In *Proceedings of the Thirteenth Annual Symposium on Computer Applications in Medical Care*. Stanford University Knowledge Systems Laboratory, Report KSL-89-14.

Chavez, R. M. and Cooper, G. F., 1989b. "A Randomized Approximation Algorithm for Probabilistic Inference on Bayesian Belief Networks." Stanford University Knowledge Systems Laboratory, Report KSL-88-72.

Cherniak, C., 1986. *Minimal Rationality*. MIT Press, Cambridge, Mass.

Cooper, G. R. (forthcoming). The Computational Complexity of Probabilistic Inference Using Bayesian Belief Networks. *Artificial Intelligence.*

Doyle, J., 1981. A Truth Maintenance System. In Webber, B. L. and Nilsson,

N. J., editors, *Readings in Artificial Intelligence*, Morgan Kaufmann, Los Altos, California, pp. 496-516.

Garey, M. and Johnson, D., 1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, New York.

Goldman, A., 1986. *Epistemology and Cognition*. Harvard University Press, Cambridge, Mass.

Goldman, S. A., 1987. "Efficient Methods for Calculating Maximum Entropy Distributions." Master's thesis, Massachusetts Institute of Technology. Laboratory for Computer Science, Technical Report 391.

Goldman, S. A. and Rivest, R. L., 1988. A Non-Iterative Maximum Entropy Algorithm. In Lemmer, J. F. and Kanal, L. N., editors, *Uncertainty in Artificial Intelligence 2*, Elsevier Science Publishers, North-Holland, pp. 133-148.

Harman, G., 1986. *Change in View: Principles of Reasoning*. MIT Press, Cambridge, Mass.

Jeffrey, R. C., 1983. *The Logic of Decision (2nd edition)*. University of Chicago Press, Chicago.

Lauritzen, S. L. and Spiegelhalter, D. J., 1988. "Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems." *Journal of the Royal Statistical Society (Series B)*, 50(2), 157-224

Nisbett, R. and Ross, L., 1980. *Human Inference: Strategies and Shortcomings of Social Judgment*. Prentice-Hall, Englewood Cliffs, NJ.

Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, Calif.

Simon, H., 1979. "From Substantive to Procedural Rationality." In Hahn, F. and Hollis, M., editors, *Philosophy and Economic Theory*, Oxford University Press, Oxford, pp. 65-86.

Stich, S., 1990. *The Fragmentation of Reason*. MIT Press, Cambridge, Mass.

Tutte, W. T., 1984. *Graph Theory*. Addison-Wesley, Menlo Park, Calif.

Valiant, L., 1979. "The Complexity of Enumeration and Reliability Problems." SIAM Journal of Computing, 8(3), 410-421.

Rendered into HTML on 12/29/02.

Elijah Millgram
Philosophy Department
University of Utah
Salt Lake City UT 84112