

Causality and Belief Dynamics à la Gärdenfors

Introduction

Gärdenfors (1988) defends a notion of causal beliefs within a general framework of belief dynamics. In the current paper, I shall first discuss his views on this in the style of a general, section by section, critical commentary. Second, I shall attempt to ferret out the origins of his possible errors and oversights in an attempt to bring up some positive suggestions for the project. These limited positive suggestions shall occupy the third portion of the present paper.

Part One- Critical Discussion of Gärdenfors on Causal Beliefs

I shall examine each section of chapter nine of Gärdenfors (1988) individually, though since each section in his work generally presupposes those that occur before it, my analysis shall generally be so dependant as well.

§9.1 - Background

This section starts out by alerting the reader to the commonly assumed connection between causation and explanation. He mentions regularity theses to begin with, and discusses some of their common problems, and a resolution to these problems offered by Lewis. Both these accounts, according to Gärdenfors, presuppose what he calls a "deterministic" universe. This choice of terminology is common but unfortunate, as it tends to cloud noncausal forms of "becoming" that may nevertheless be lawful, and further, need not be probabilistic either.

Bunge (1979) has made the preceding point with some forcefulness. I shall make use of his analysis of what he refers to as "general determinism" later in this work, so it will bear some exposition here. In short, Bunge starts by defending two principles of general determinism, namely the principle of lawfulness, and the genetic (or Lucretius) principle. The principle of lawfulness is the thesis that all events occur in objective regularities. This can also be put negatively: there are no miracles. Bunge suggests that this principle of lawfulness is required for science, and indeed, for action generally. In a lawless universe one could never be sure one's actions would ever be efficacious in any way. The genetic (Lucretius) principle is that nothing comes out of nothing and nothing goes into nothing¹. Causation then

¹ A clarification of this principle in the light of modern physics is sometimes necessary. Virtual particle creation does not violate this principle

becomes one specific case of general determinism, and is characterized by constant production. It is suggested that there are perhaps no genuinely causal laws, only laws with causal ranges. Laws here are the objective regularities mentioned previously. This is important, as we shall see, part of Gärdenfors' (too hasty?) dismissal of nomological causation stems from his Kantian interpretation of laws.

A notion of causal beliefs based on this framework could go in two distinct directions. One could have beliefs about the degree of causality involved in a particular class of cases. One could also have beliefs about events (putatively) involving causality.

What Gärdenfors calls "probabilistic causation" is similar to what Bunge calls instances of stochastic laws. We shall see in due course how there is a potential confusion created by this label. (See part II of the present paper for this discussion.)

Gärdenfors then finishes this section by telling his readers that he is going to elucidate causal beliefs by dealing with them in terms of his previously discussed notions of belief contraction. (For this notion see especially §3.4, all of chapter 4, and §5.7-10 of his book.) He then gives his criterion of causation: C is said to cause E if the occurrence of C raises the probability of E.

§9.2 - The Information Problem

Gärdenfors introduces this section by discussing the views of Suppes (1970). He makes the point that Suppes views have been criticized extensively and that the common thread to the problems raised is that the point out that having a single probability function is not enough to differentiate genuine and spurious causes and other similar worries. This is fine as far as it goes, however, we shall see later that Gärdenfors' use of a probability function is open to similar kinds of objections as he points out have been raised against Suppes. He finally suggests that more knowledge of the situations described in the supposed counterexamples would make the counterexamples disappear. This is promising, as typically Humean or Kantian views on causation do indeed leave out many steps between "cause" and "effect".

§9.3 - An Epistemic Analysis of Causal Beliefs

This section introduces Gärdenfors' Kantianism; he holds that causal relations as the "vacuum" in modern physics always contains electromagnetic fields.

get "imposed" on chains of events, rather than objectively occurring between events themselves. He is correct to point out that nothing necessarily hinges on this choice. However since it appears at least *prima facie* plausible that beliefs about something would vary depending on whether they are about things and their changes² in the world as opposed to simply "something in our heads".

After briefly comparing his view to that of Lewis, Gärdenfors proceeds to formally state his causal belief principle. I quote it for reference (pg. 195):

In the epistemic state represented by the probability function P , C is a cause of E iff (i)
 $P(C) = P(E) = 1$ and (ii) $P_C^-(E/C) > P_C^-(E)$.

In words, C is a cause of E iff the "probability" of both C and E are unity and the "probability" of E given C in a state where C is rejected is greater than that of E alone. As I have intimated previously (Douglas 1999) this use of the term "probability" is somewhat dubious. See below in this section and section two of the present paper for the "roots" of this worry as it pertains explicitly to causation and causal beliefs.

After some discussion of minimality in terms of revisions which are central to the notion of causal beliefs being developed, Gärdenfors then discusses the temporal nature of his account briefly, suggesting that we take P_C^- prior to an event, say, the rolling of a die. This section is curious on several grounds. Since this probability notion is central to understanding Gärdenfors' notion of cause, it is vital to examine it closely. I shall do that below.

First, it is odd because he seems to confuse events and beliefs about events. This can be shown if one looks in a little detail at a rolling of a an ordinary fair die. Then, conventionally, the probability of rolling, say, a three, is 1/6. Let the event of rolling a three by "A". Then $P(A) = 1/6$. But on the subjectivist interpretation, one has to figure out $P'(P(A)=1/6)$ is, where P is the objective probability function and P' the subjective probability function. There is no reason why these have to be the same. Subjectivists like Howson and Urbach (1989) would no doubt contend that the two have to be identical to avoid "Dutch Book" arguments. Perhaps so, but numerically equal does not entail that one can dispense with this notion which amounts to the distinction between the chance of the world actually coming out

² This presupposes that it makes sense for a Kantian to talk about changes of things in themselves. I am not sure that it is, but that is another story for another time.

with a die with three on top, and one's credence that event³ will happen with probability 1/6th. Before going on here, I would like to forestall an objection: nothing hinges on whether the event is ontologically random at the lowest level. One often overlooked fact about randomness is that there are "levels" to it (see Bunge 1998b). All that it is necessary for this objection to fail to do justice to what I am claiming is to note that "objective randomness" here means "large insensitivity to initial conditions". However, if that is insufficient, the reader is asked to consider the outcome of some event with what is scientifically recognized (at present) to have *fundamental* randomness, or full propensity to "come out" either of two or more ways⁴. With that out of the way, let us look at what is wrong with the failure to distinguish these two kinds of "probability" at work in one event.

This points to a horrible confusion because it "levels the playing field" between random events and events that are epistemically uncertain. (Now, of course, one may not think that there are random events in the way I have described above, but this amounts to denying some very solid results from modern science.) I object to this levelling, as it is not clear how one compares events of different kinds of credence. Suppose I believe a proposition Q with "subjective probability" 1/6th (e.g.: my friend Robin is a bad speaker of Hebrew.). The subjectivist would claim that I ought to take bets at equal odds on "three comes up" in the situation described previously and "Robin is a bad speaker of Hebrew". How does this work? How is my uncertainty related? In science, we are able to tell why certain quantities are equal; for example, two equal accelerations are produced if equal forces are applied to equal masses. Of course, the subjectivist will simply answer that "it just feels right", or something. This is grossly unsatisfactory, because it does not allow us "double check" our figures. When I calculate the acceleration on a body using Newton's laws, I can check to see that the observed result is reasonably in agreement and further agrees with other accelerations produced by forces of other quantities. No such opportunity manifests itself here.

Second, as hinted at above, this definition of cause makes propositions or sentences causes of other propositions, which is odd. This is not merely a quibble about the Kantianism mentioned previously. To me, this appears to be a barbarism of language, as we say "the moving eight ball caused the nine ball

³ Here I am using "event" as it is normally used in "conventional" metaphysics.

⁴ For example both radioactive decay of atoms and genetic shuffling are regarded this way.

to move", not "'the eight ball moving' caused 'the nine ball to move'". To be fair, Gärdenfors does appear to waffle (pg. 197, 1988):

"Assume that we have thrown a die (event E) ..."

But then we get the problem of the probability functions alluded to above over again. How can one function be defined over events and propositions ABOUT those events? Alternatively, if "E" is supposed to be a belief about an event⁵, then again the language seems to be a bit strange. I shall assume hereafter that Gärdenfors means by "event E" something like "event represented by event E." This leads back into the first worry mentioned above as it seems to require what might be called a "second order probability measure". He does seem to almost recognize this need in chapter 8 (1988), but it is not discussed at all in the material on causal beliefs for some reason.

Third, there is a worry that the definition of cause consecrates some events as causes that intuitively we would not consider causes. One of these is that it appears that it makes everyone's birth a cause of their death⁶, and indeed, stronger claims, like "The accretion of the Earth out of interstellar dust and rock is a cause of the construction of Tienamen Square", which I imagine strikes many of us as very odd - it certainly does to me. Gärdenfors attempts to deal with "spurious causation" in the next section, so I will elaborate on this point in my commentary thereon.

Finally, it does not appear that the definition of causality provided allows for time correctly. Much literature on causality (e.g. Brand 1976) is concerned with the notions of time and temporal priority involved. Consider the event: "Robin and I having dinner". This event makes it more likely that "Robin and I met", but under most conceptions of causality, later events do not cause earlier ones⁷. (This analysis does assume that both events have

⁵ One weakness of Gärdenfors' account generally is that it does not allow us to pick out the kinds of events in the world we would want to have causal beliefs about. The obvious Kantian move is "all of them" (phenomenal world) or "none of them" (one reading of the noumenal world) (see Kant 1929 [1781] in the Analytic of Concepts where Kant attempts to demonstrate universal causality), but this (as is well known) is fraught with difficulties.

⁶ This claim is subject to a very large revision later, when I consider the subjectivity of beliefs.

⁷ For a striking and novel denial of this common sense claim, see Price 1996.

happened, so they have unit probability in one sense at least⁸.) It may appear that this case may be dealt with by his "screening off" principles, as developed in the next sections, but this is not clear. More on this in the following subsection of my paper.

§9.4 - Analysis of Some Examples

In this section Gärdenfors deals with some examples from the literature, and sees how his conception of causation holds up under scrutiny from typical cases.

The first case he considers concerns the falling of a barometer's reading, stormy weather, and low air pressure. He points out that normally we regard the notion that the falling barometer readings cause the stormy weather as wrongheaded. He explains that his account justifies this by looking at the belief that the barometer was falling, and what minimal change to beliefs is required to keep the rest of the data of the example the same. He suggests that we would want to say that the barometer was malfunctioning or disturbed, etc. This is rather problematic, as without a way of elucidating similarity, this move does not appear to do any work. How is what he is doing any different than enshrining our pretheoretical concerns?

It may be objected that the next sentence in the section, concerning how the "low pressure" datum is more entrenched than that of the "falling barometer" datum solves this puzzle. But we have no general notion of entrenchment to make use of. Nowhere does Gärdenfors provide us with postulates for ranking beliefs according to entrenchment. Chapter 4 (1988) discusses this at great length; I shall briefly discuss this "way out" in section two below. But note that this also makes use of an unanalyzed notion of similarity. One can use the "subjective probability" values, but they are equally ad hoc - to see this, just consider an agent who's "subjective probabilities" of both events are reversed from Gärdenfors' intuition. Does this mean that causes are subjective somehow⁹? He also gives no explanation of why the subjective probability of two of his events is unity¹⁰.

⁸ They may not have unit probability to some particular subjectivist, of course. More on this in section II, below.

⁹ A Kantian could presumably claim that humans are using the same categories and so any humans would always judge the same way in each case. This does seem very unlikely, especially considering we can (as I have just done) counterfactually judge in a different way.

¹⁰ This worry is deeply connected with my earlier worry concerning the connection between the two kinds of probability I mentioned in passing in the

I have nothing to say about his next example, so will hence next discuss his flagpole example which follows it. This example is interesting as it is one he claims provides problems for regularity analyses of causation. As my usual notion of causation (much like Bunge's, described briefly above) is one kind of regularity thesis, so his claims do bear careful scrutiny.

He asks us to consider a flag pole of height h which casts a shadow at noon on a sunny day of length l . His first awkward remark concerns most people's intuitions. While it may be true, as he writes, that most people's intuitions are such that they think that the position of the sun and the height of the flagpole cause and explain the length of the shadow. I do not share this intuition. I do not regard the sun (or its location), etc. as explaining anything. The facts of the matter may be USED in constructing an explanation by an agent, but only agents can explain. In other words, explanation is semantic¹¹ and psychological, and not purely formal or factual. This, however, is just a quibble, though it is one that can be run through many discussions of explanation. The next point Gärdenfors considers, that this example provides problems for regularity accounts of causation, is far more interesting. He claims that the regularity accounts cannot explain our intuitions on this matter because all regularity accounts which derive the length of the shadow from the height of the flag pole are of the same logical form as ones that would derive the height of the flag pole from the length of the shadow. This is no doubt correct, but the conclusion that he draws from this, namely (1988 pg. 199),

"... there can be no laws or logical connections that can distinguish the height of the flag pole causing the length of the shadow from the converse relation."

is altogether too hasty. To see why, we need to both investigate Gärdenfors' notion of cause and his notion of law. The latter is held rather tacitly, and is the greater source of the problem here.

It appears that what Gärdenfors has in mind by "law" is what Bunge (1979) calls law statements. I take this from the fact that he says "laws or logical connections". If he did not have in mind statements, but instead (say) objective regularities, the statement would not be semantically homogeneous. discussion of "second order probabilities."

¹¹ Here I do NOT mean semantic in the sense that it is sometimes used in logic and model theory.

Hence in the interests of charity I shall assume he meant law statements¹².

Gärdenfors has also (and more problematically) assumed that only the form of a causal explanation (or generally, one by means of law statements) matters. This is false; the content of such explanation matters as well (Bunge 1998b), which is not determined merely by logical form. Thus it is at least *prima facie* plausible that one could develop a regularity analysis of causation that is not merely formal. Perusal of the various modern papers in Brand (1976) suggests that this purely formal look at causal explanation indeed fails to capture regularity. But how could it? Formal properties are logical or (rarely) semantic properties; the causal (or more generally, lawful) regularities in the world are not logico-semantic properties of things. We thus must capture these nonlogical properties of things by creating correspondance rules (Bunge 1998a).

To avoid begging any questions against Kantian accounts of causation here, I must stress that I am not defending the various regularity theses here directly, but am instead merely pointing out that Gärdenfors appears to have too hastily dismissed them.

With that out of the way, we can next examine Gärdenfors' own causal explanation of the flag pole example. Gärdenfors tells us (pg. 200) that:

“the smallest variation of the situation to change the height of the flag pole would also change the length of the shadow. On the other hand, the smallest variation to change the shadow would not involve any change to the flagpole.”

This is almost in agreement with a regularity analysis. The regularity analysis would presumably in addition provide a mechanism¹³ by which the flagpole produced its shadow. But there is a problem with this - the above passage is the nontechnical explanation of the preceding one. This is interesting, to say the least. Let us compare it to the preceding paragraph in

¹² Another reason to suppose this is Gärdenfors' Kantianism. It is plausible that to a strict Kantian, there are no laws (i.e. the “noumenal world” is lawless). At least they are fundamentally unknowable, law statements being “imposed on phenomena”, as laws would belong to things in themselves, if to anything at all.

¹³ Note that here proposed mechanisms are regarded a (deep) form of explanation. However, what they refer to is something in the world, usually involving transphenomenal properties and things. This latter characteristic is likely why empiricists (and Kantians, who are empiricists in a weird sort of way) mistrust mechanistic explanations.

order to draw out the differences. The preceding paragraphs read (1988, pg. 199-200):

"My solution to the flagpole example is to investigate the appropriate *contractions* of the outlined epistemic state in the examle. Let K be this epistemic state, and let H be the fact that the height of the flagpole is h, L be the fact that the length of the shadow is l, and S the fact that the sun is shining (from its particular position). If we now want to show that H causes L, we should, according to (Def Cause), consider K_A . If we give up our belief in H, this would not affect our beliefs in S, and consequently we would give up the belief in L, that is, $K_H(L) < 1$. Because $K_H(L/H) = 1$ according to (P5) (which is valid here because the situation is deterministic), it follows that H causes L in K according to (Def Cause).

In order to determine whether L also causes H, we must consider K_L . Now when giving up L, we must also give up H or S; that is, we must imagine a situation in which one of these events does not occur. The simplest way is to give up S (that the sun is shining), for example, by assuming that a cloud temporarily covers the sun so that the shadow disappears (this is a smaller change than assuming the height of the flag pole has been altered). In this situation H would still be accepted, that is $P_L(H) = 1$, which is sufficient to show that L does not cause H."

These two paragraphs are a veritable minefield of strangeness. First, Gärdenfors seems to beg the question when he says "If we give up our belief in H, this would not affect our beliefs in S, ..." I agree with this statement, but Gärdenfors' model has no room to allow for it. This is because the only connections he is considering between statements are logical ones. For all he knows, the content of the statements does indicate there is some odd connection between H and S. How does the Kantian get around this? This is another lace where semantic aspects of explanation (i.e reconstruction¹⁴) of causation, etc. is necessary¹⁵.

Further, the remark that this case is deterministic is odd. In the light of what I have said previously (see my remarks on §9.1, above) on determinism, of course that's true, but Gärdenfors is using "determinism" another way, to mean "causal", as opposed to "probabilistic." On this model of causation, how does one make the distinction without begging the question against an agent who

¹⁴ Reconstruction is the epistemological operation of building a mental model of something in one's head; of representing a system and its parts and relations. (Here model is again not being used in the sense it has in model theory.)

¹⁵ For a simple demonstration that correct logical form is necessary but not sufficient for explanation, consider the following explanation why I should have Jell-o™ for dinner. If marijauna is legal in the Netherlands, then I should have Jell-o™ for dinner. Low and behold, marijauna IS legal in the Netherlands. Hence I conclude that I should have Jell-o™ for dinner. This silly example is to show that at least what might be called semantic closure is needed. This is all the more the case in scientific explanation. (See Bunge 1998b for this.)

denies the unit subjective probabilities involved? Consider the case of two agents examining a particular case. One insists that she accepts (in Gärdenfors' terminology) certain propositions thus and so and so a situation "becomes causal" on this Kantian-like account. The other agent insists that, no, something is merely "probable" in the situation. Thus the agents will disagree on whether a situation is causal. Since humans notoriously disagree very often, it is quite reasonable to suppose that some people will find Gärdenfors' account of causal explanations in a particular case to be wrong. Since the attribution of causation thus becomes ideoyncratic, at best, it thus appears that the account that Gärdenfors has given us is somewhat diminished in its usefulness. But let us turn to the next large section of the chapter to see if anything further can be salvaged.

§9.5 - A Comparison with Granger Causality

This section contains Gärdenfors' brief discussion of how his model relates to that of Granger. I am not familiar with Granger's work, and so my comments are limited to two very brief remarks. One is that Granger's account is *prima facie* more plausible to me than Gärdenfors' because it explicitly includes the time variable. This is important, as Gärdenfors is discussing something temporal (the very word "dynamics" in his 1988 title is temporal in meaning¹⁶). Second is that Granger makes use of the concept of an event, though perhaps unfortunately he has not elucidated it. However, to evaluate its merits and to do a proper comparison with Gärdenfors would take me too far afield, so I will simply note that these two differences make comparison difficult. It appears that Granger is more of a realist based on his use of "event" than a Kantian¹⁷, it thus seems strange that Gärdenfors thinks the two accounts are compatible.

§9.6 - Causation and Explanation

In this section causation is tied together with explanation. I feel that I have dealt with Gärdenfors' earlier remarks related to this subject adequately above and do not need to dwell on this section much. I make the following general note, however. Since Gärdenfors' notion of causality is not

¹⁶ Dynamics is the study of change, which is generally regarded as "requiring time", or more recently, time itself is the "unfolding of change". See Bunge 1977; Price 1996; McCall 1994 for very different views on this.

¹⁷ It is possible that "events" could be in the world of phenomena alone for a Kantian, or derived from the categories and thus imposed on the world. But this just suggests that "anything" at all is possible imposed by us, and Kant's treasured transcendental idealism becomes subjective idealism really fast.

ontological but epistemological, it does stand to reason that there is a very close connection between his concepts of explanation and of causation. To a realist, of course, causal statements may be involved in explanations, though probably never alone. Furthermore, causes do not belong in explanations according to the realist. Of course, I cannot insist that by saying that causes are explanatory in the way Gärdenfors does is wrongheaded alone without begging the question against the Kantianism. It hopefully is clear from the previous sections why I regard this move as mistaken.

§9.7 - Some Further Aspects of the Causal Analysis

This section contains some rather interesting remarks. One is that the causal relation is not transitive in general. I agree, however, as we have seen, Gärdenfors cannot rule out an agent who insists that "her causal relation"¹⁸ is completely transitive, which is a bit odd. (However, if we regard events that have already happened to have unit probability, whole "historical chains" quickly become sequences of transitive causal chains.) I imagine the consistent Kantian must simply bite the bullet on this one; I do not regard this last comment as an objection but merely a puzzling spot to point out that may work to "undermine" the project by showing one of its more implausible features.

This is dealt with somewhat in the next paragraph, that concerning what Gärdenfors calls causal overdetermination. His example does have another bizarre conclusion that he seems to accept gleefully, namely that "A or B" is the cause of something, which again only makes sense if one thinks that causes are propositions. As might be gathered, I regard this result of having "disjunctive" causes as another unacceptable, though not thoroughly self-refuting result of Gärdenfors' theory. A realist about causation (and presumably also a non-Platonist) finds "disjunctive causes" to contain a category mistake. Things in the world do not have logical properties - our reconstructions of them do. Note also that on a realist model not even our causal beliefs can be "disjunctive". This is the difference between $C(A \otimes B)$ and $C(A) \otimes C(B)$, where C is some sort of "causal belief function" that has domain events - events have no logical properties. Of course, the Kantian can simply deny that there is anything being reconstructed.

But again later in this section, Gärdenfors waffles. He writes (pg. 207):

¹⁸ Since, as I have tried to argue, Gärdenfors is committed to "causes being in our head" and there is thus nothing to stop one from relativizing the notion to every individual, there is no reason not to call it thus.

"According to the analysis presented here, the causal relation between single events is the fundamental relation."

The above statement is very odd. We have no elucidation of the notion of an event given to us or even referenced (the index to Gärdenfors' (1988) text contains no entry "event"), so it is difficult to know exactly what he means here. But if causation is reduced to causal beliefs¹⁹, how can there be such a relation between events, unless of course the events are "epistemic" in some interesting respect. I feel that because of this potential confusion, Gärdenfors "owes" the reader an elucidation of event, or at least a reference to the literature on the subject²⁰.

The rest of this section continues on these lines; I see no need to pursue them further.

§9.8 - Limitations of the Analysis

In this section, Gärdenfors "owns up" to some of the shortcomings of his account of causation presented earlier in the chapter. He admits that the temporal nature of causation is left out, though as I remarked earlier this oversight is perhaps understandable within the broad (pseudo?)Kantian framework he is working with. He claims that one can "build in" time by postulating certain features of the "probability functions" he is using. Perhaps, but it appears that it would require putting time as an epistemological concept "or categories", rather than an ontological one, which as is well known is fraught with difficulties. Of course, this is simply the Kantian card being played again, this time with an even more (apparently) basic notion. This section ends the chapter, and indeed the main parts of Gärdenfors' (1988) book.

Section II - Roots of Errors and How to Save as Much as Possible

¹⁹ Or one might say, rather awkwardly, "beliefs-that-are-causally-structured".

²⁰ I assume that there IS a literature on the subject. I have not done enough work in this area of metaphysics to know. Two works referenced earlier briefly mention it though in a much larger context (Price 1996; Brand 1976). Other possibilities include Bunge 1977, McCall 1994 or even Whitehead 1929, though at least the latter two apparently have serious flaws. I regard this oversight on Gärdenfors' part as producing the general lesson about having to do metaphysics before epistemology. It may be countered that the sort of logicism he is using is a metaphysics of sorts. Granted; however to explain why this is unsatisfactory would take me too far afield; let us just agree that this is an area of possible weakness and leave it at that. To make his argument more complete, this "logicism" needs at least a passing defense. (See also Bunge 1977 for more on "why logic is not a metaphysics".)

In this second section of the present paper I shall take the specific oversights discussed in section I and repackage them into a few general ones, and attempt to come up with a way of rescuing Gärdenfors' project without being too violent to his approach²¹, which shall occupy the final section of the present paper.

I see three broad roots of what I take to be Gärdenfors' mistakes. These are ontological, epistemological, and semantic confusions. I shall touch on each in turn, then in the final section of the paper, offer a suggestion in each of the three categories.

First, ontological confusions. As I have stressed repeatedly, there are two main aspects to Gärdenfors' ontology that bear some examination. One is the Kantian aspects; the other is the curious lack of contact with the nature of the agents he is supposedly modeling beliefs of. Both of these infect the discussion of causality extensively.

As intimated previously, a Kantian if pressed becomes a subjective idealist. If that route is taken as an "escape" I shall ignore it, as being thoroughly incompatible with everything we know (and take for granted). So the Kantian has to "touch base" with the real world once in a while to avoid falling down that well.

It is correct that to defend his Kantian view of causality would take Gärdenfors too far afield, however, that simply begs the question against alternative explanations, even taking into account the belief dynamics he has sketched elsewhere in the (1988) book.

We have seen that the Kantian aspect of his account seems to require him to owe us elucidation of several important concepts which conspicuously missing from his work. These are: *event*, *second order probability*²², *cause*, *effect*. I put the last two in, as we have seen there is some confusion over whether the latter are '*something in the world*' or propositions (here taken to be

²¹ As an aside, I have come to realize, by reading Gärdenfors' work and other literature on this general subject that the idea that we carry beliefs sententially in our heads is profoundly and horribly mistaken. But that is another story for another time.

²² Of course, Gärdenfors does not use this term in this chapter. I have argued in section I of the present paper that it is needed to make sense of some of his claims. If this is not the case, at least it appears he has to defend his reasons why such a notion is not needed here, especially in the light of the admission in §9.6 that there is a strong connection and explanation.

'sentences in the head') of some kind. This ties into the second general ontological worry. I mentioned that his account does not 'touch base' with real epistemic agents much. This is fine, except as it pertains to causation and in particular, causal explanation.

This is because a failure to do so results in unexamined and consequences concerning the subjectivism I have noted previously.

Second, Gärdenfors indulges in several semantic confusions. I have hinted earlier that there is a running conflation between causality, causal explanations, and causal beliefs. I trace that to the Kantianism, discussed above. It appears that the general lesson is that the Kantian cannot "move causality" out of the world as much as she would like to. Let us examine the flagpole example again to this end.

He writes that (1988, pg 199-200):

"Let H be the fact that the height of the flagpole is h , L be the fact that the length of the shadow is l , and S the fact that the sun is shining (from its particular position). If we now want to show that H causes L , [...]"

If as he writes earlier (*ibid*, pg. 194) causation is in our heads, then the "facts" alluded to above are also in our heads in order to retain semantic closure of the above sentence. This is an "ordinary language" use of the word "fact"; however here it seems a tad out of place for that reason.

Further, "belief" is missing from the above account. It is implicit in his use of " H causes L ", which we can read as " H produces belief in L ". But this causes a semantic version of Hume's problem of causation. What necessary connection is there, even if we relativise the notion to each agent as I have suggested earlier? His story about there credences ("subjective probabilities") being raised in such and such circumstances is purely descriptive. Nothing about the states of the agent involved are mentioned. This semantic point thus straddles ontology and epistemology, and is brought out most forcefully above, where I mention "second order probabilities".

Finally, another semantic point is as follows. If "talk of causes" is short hand for "talk of causal beliefs", then talk of "the cause" and even "caused" itself become a bit linguistically awkward. The latter because it is really short hand for "made someone believe that". The term "the cause" is equally awkward, as causes fail to be singular across different agents. Further, as beliefs change in agent "the cause" of certain other propositions may change

or cease to be one.

Epistemologically confused aspects here straddle both the semantic and ontological areas, which is why I have chosen to deal with them last, as they are usually merely corrolaries to the confusions of the two previous sorts.

We have seen one of these in my discussion of semantic problems. Another concerns notions of belief contractions. As mentioned above, I will spend some time here explaining why Gärdenfors' notions concerning what he calls belief contractions and entrenchment cannot help with notions of similarity. Let us look briefly at the earlier section of the book where this is discussed to this end. Page 80 of Gärdenfors 1988 introduces the definition of *contraction function* he is going to use: "(Def Part) $K_A^- = \mathbf{IS}(K \setminus A)$ " where $K \setminus A$ is (ibid, pg. 76) "the set of all belief sets K' that are maximal subsets of K that fail to imply A [...]." S is a selection function that "picks out" the most entrenched beliefs in K . He then discusses several models for determining which beliefs are more entrenched, including Grove's "system of spheres" and a "possible worlds" representation. His postulates of ordering and so on are (reasonably) well founded, and I will take them as given. However, the only remark concerning the origin of degrees of entrenchment (rather than their changes) is found on pages (ibid) 91-92, where he briefly mentions an information theoretic approach to this problem. As I have mentioned briefly elsewhere (Douglas 1999), this makes entrenchment rather subjective²³. It thus cannot ground similarity and hence not causation, if indeed it is desirable to avoid having "subjective causes."

I note that even a Kantian should be worried about subjective causes, for at least two reasons. First, as noted previously, the more "things" become mere concepts, the more likely the Kantian is to get on the slippery slope to subjective idealism. Second, it leaves unanalysized the origin of our causal beliefs; why do we have them? We clearly cannot be caused to have them. The whole notion of cause becomes somewhat superfluous. This is because one can always imagine that we should be "cautious" and never admit non-tautologies with unit "probability".²⁴ Since Gärdenfors' criterion requires that the

²³ Lest I be misunderstood here: there are possibly two kinds of subjectivity at work here. I do not mean that it is problematic to have agents with different epistemic entrenchments. I mean that that what is found to be entrenchments thus become subject.

²⁴ I note in passing that this motivates another reason to suppose that there are two accounts of "probability" being used by Gärdenfors without realizing it. Tautologies and "statements about events that definitely happen in situations thus and so" have probability of 1. This brings out the semantic

proposition "cause" and the proposition "effect" have unit probability, it thus follows that this particular individual has no causal beliefs at all (note that tautologies do not "cause" other tautologies as it is impossible (see ch.4) for a Gärdenforsian agent to contract her belief set with respect to a tautology). It seems plausible to state that many an individual who tries to avoid being dogmatic will not hold any beliefs with unit "probability" and hence not hold any causal beliefs either.

This subjectivism also infests some of the more "traditional" cases of "probabilistic causation". Is a dice toss now "reduced to" our beliefs about the toss? It is not clear at this stage what Gärdenfors would want to say, in light of the semantic homogeneity issues I have raised earlier. The two different notions of probability I have mentioned thus suggest two different sorts of "causation" or notions of "causal belief". I suggest that this is a source of some confusion on the part of Gärdenfors; in the next section I will explain briefly how it might be avoided.

Section III - Cleanup

In this section, I clean up Gärdenfors' notion of causation and causal beliefs and give a begining sketch of a realist version using his previous framework. I shall do these by means of suggestion of two postulates of a more semantic nature. First is an axiom of semantic closure; the second concerns postulation of a mechanism, which also helps to deal with ontologico-episetic components. Let us take each of these in turn.

Introduce the *partial meaning* of a belief B relative to an belief set K as the intention of B within K^{25} . Two beliefs A and B are said to be *semantically closed* just in case the intersection of the *partial meanings* of A and B in K is non-empty. The postulate that helps Gärdenfors is thus that cause and effect must be semantically closed. This rules out causes of the kind Gärdenfors and I regard as spurious, without either sacrificing his epistemically-based notion of causality or without the difficulties we have seen earlier

Then event C is a cause₁ of event E just in case there is a semantically inhomegenity quite forcefully. A similar point was made by Seetzen (1999). I am indebted to his comments on this matter for some of the "origin" of this analysis.

²⁵ For one suitable definition of intention, see Bunge 1999. Nothing hinges on exactly this choice; another theory of reference could be substituted without difficulty assuming that intentions can be modelled set-theoretically.

closed mechanism²⁶ available to the agent in K such that:

- 1) C has temporal priority to E
- 2) $P(E/C) = 1$ where **E** is the reconstruction of the event E,
and, the same, *mutatis mutandis* for C and **C**.

None of this means giving up Kantian notions. If Gärdenfors wants to continue being a Kantian, the probability issues I have raised must be solved - I cannot do that here for lack of time or ability. Assuming that is accomplished P then is a subjective probability function of this new sort. Cause₁ thus is a subjective cause, or causal belief; what I have given is a begining sketch of what would be a causal belief story if a subjectivist interpretation were to be pursued. If one wants to know the actual cause of E, one must verify that $P(E/C) = 1$, rather than "discovering that" $P(E/C) = 1$.

Conclusion

I have critically analyzed Gärdenfors' claims about causation and causal beliefs and found them wanting. I have systematized these perceived oversights, and presented a minimalist way in which his general project can perhaps accomodate a perspective that is hopefully more satisfactory in several ways.

²⁶ Mechanisms do not necessarily involve causation. Instead, mechanisms assure what Hume would have called "constant conjunction" and this is what will do the work in "just a moment".

Bibliography

- Brand, M., ed. 1976. *The Nature of Causation*. Urbana: University of Illinois Press.
- Bunge, M. 1977. *Treatise on Basic Philosophy, Volume 3: The Furniture of the World*. Dordrecht: Reidel.
- Bunge, M. 1979. *Causality and Modern Science* (3e). New York: Dover.
- Bunge, M. 1998a. *Philosophy of Science I: From Problem to Theory*.
- Bunge, M. 1998b. *Philosophy of Science II: From Explanation to Justification*.
- Bunge, M. 1999. *Dictionary of Philosophy*. Amherst: Prometheus Books.
- Douglas, K. 1999. Gärdenfors' Belief Dynamics. Unpublished paper for University of British Columbia course PHIL 323, Winter Term.
- Gärdenfors, P. 1988. *Knowledge in Flux: Modelling the Dynamics of Epistemic States*. Cambridge: MIT Press.
- Howson, C. & Urbach, P. 1989. *Scientific Reasoning: The Bayesian Approach*. Open Court: La Salle.
- Kant, I. 1929 [1781]. *Critique of Pure Reason*. Houndsills: Macmillan. N. Kemp-Smith translation.
- McCalll, S. 1994. *A Model of the Universe*. Clarendon Press: Oxford.
- Price, H. 1996. *Time's Arrow and Archimedes' Point*. Oxford University Press: New York.
- Seetzen, H. 1999. Unpublished comments from the proceedings of University of British Columbia Course PHIL 323, Winter Semester.
- Suppes, P. 1970. *A Probabilistic Theory of Causality*. North-Holland Publishing Company: Amsterdam.
- Whitehead, A. 1929. *Process and Reality*. Macmillan: New York.