Edited good cause notes, started collecting Oct. 4, 2002

A Re-introduction to Philosophy

Why Post-Fregean Methods Have Not Reduced Philosophical Paradox and Disagreement

John C. Cahalan

2. Causality

It might appear that causality has been a much discussed topic in recent philosophy. In actuality, we have discussed it very little. Treatments of causality have discussed such logical and epistemological topics as law-like and nonlaw-like universals, deterministic predictability, contrary-to-fact conditionals, and the modality of necessity. I will deal with such topics only after I have analyzed causality. Causality is something ontological, not epistemological or logical; it is a relation of thing to thing, not of thought to thing or of thought to thought.

Our acceptance of Hume's critique of causal knowledge is, of course, the reason why we think that a discussion of causality is first and foremost a discussion of certain kinds of knowledge, with their specific logical properties. Hume's critique of causal knowledge is multifaceted, but I will not have to deal with all of it. I intend to establish as a knowably necessary truth that every change has an efficient cause; a change without an efficient cause would be a contradictory reality. Our knowledge of that truth follows from our understanding of terms; but the truth concerns a condition holding in reality, if change occurs in reality. It happens to be the case that, contrary to Hume, the understanding of terms leading to knowledge of this necessary truth comes from sense experience; and I will try to explain how. However, the important issue is not where the concepts come from but this proposition informing us about reality is necessarily true.

To establish this, I will first show that there can be necessary causal connections, then explain what necessary causal connections are, and then show why there must be necessary causal connections.

2.1 How There Can Be Necessary Causal Connections

Imagine two billiard balls, balls A and B, at rest and separated by some distance. Assume that these balls have the properties of being able to undergo locomotion and of not being able to occupy the same space at the same

time. Now imagine that ball A begins to roll in the direction of ball B until it comes in contact with ball B. Call ball A's changing of place change 1. In the imagined circumstances, A's contact with B is a situation in which another change, call it change 2, must occur.[1] When that contact occurs, it is contradictory for both A and B to remain what they are in all respects. At least one of them must cease being what it is, in some respect, and become something else.

For example, at the contact, A may simply cease moving and remain where it is. If so, A has ceased to be a thing in motion relative to B; A ceases to be what it is in that respect. Or A may bounce off B, in which case A ceases to be a thing moving in one direction and becomes something moving in another. But if A neither stops nor changes direction, B must begin to move. If so, B has ceased to be something at rest relative to the surroundings it shares with A; it ceases to be what it is in that respect and becomes something else, something in motion. We can, of course, imagine A continuing to move in the same direction but passing through B "magically." In that case, B would not cease being something at rest, but we would not have eliminated another change required by the first. For we assumed that A and B cannot occupy the same place at the same time. For A to pass through B, that characteristic must change; so change 2 would be the change from their not being able to occupy the same place to being able to do so.

If a second change did not occur in this situation at least one of the things in the situation would both be and not be what it is. Specifically, if neither A nor B changes, A both is and is not what it is. By hypothesis, A is something moving in the direction of B's space and something that cannot occupy the same space as B. If B does not move, A is either no longer moving in the direction of B's space or no longer something that cannot occupy the same space as B. The assumption that a second change does not occur would require us to hold that A both does and does not possess at least one of its properties.

This example shows the following: a prior change can bring about a situation such that, given that the things in the situation are what they are and are related to each other as they are in the situation, at least one of the things cannot remain what it is. That situation would not exist without something else coming into existence, another change. Or, to get rid of the subjunctive, if that situation exists and another change does not occur, something is and is not what it is. When such a situation occurs, I will say that sufficient causes of the second change exist. Sufficient causes are things such that, given what they are and the relation they have to each other in the situation, the situation could not exist without something else coming into existence. Of course, the same things could exist but not be in the relation to each other that requires something to undergo a change. In that case, we can say sufficient causes of the change exist but sufficient conditions do not, the causes being things, the conditions being a relation holding between things, for example, A's contact with B. The first change, then, brought into existence sufficient conditions for the second change. It brought into existence sufficient causes for the second change by making one of the causes, A, something other than it was (something moving instead of something at rest) and by bringing the causes into a relation with one another (A's contact with B while continuing to move in the same direction) that necessitated the second change.

The concept of sufficient causes just stated includes the otherness of causes and effects. Whatever else a cause may be, it is something really distinct from its effect, where a real distinction between things can be established by the fact that at least one of them can exist without the other existing. Each of A and B can exist without the other; A can exist without its state of change existing, and all three can exist without the existence of the second change, the effect. The otherness of causes and effects is what creates the problem with their being necessary connections between them. The necessary is that whose opposite is contradictory, and contradiction is a

denial of a thing's identity with itself. But a connection between cause and effect is a connection between a thing and something other than itself. Yet our example shows that a connection between distinct things can be not only necessary but knowably necessary. We know a second change must occur, since, if it did not, something other than the second change is and is not what it is.

2.2 What Necessary Causal Connections Are

From the assumptions we have made, we cannot predict what the second change will be; nor can we state that such-and-such a second change as opposed to all others is necessitated by the sufficient causes. All we can say is that some second change is necessary. To further analyze this necessity and its knowability, however, it will be helpful to have a specific example in mind for the second change. So let us assume that the second change is B's moving out of A's way. The analysis will leave aside what we know about change from other parts of science, except to illustrate specific points. For instance, we know that A's "action" on B requires an equal and opposite action of B on A, but nothing in the imagined situation alone would tell us that. What the argument has shown so far about causality and what the remainder of the argument will show is capable of being known prior to the empirical parts of science, that is, prior to any particular knowledge verified by the experience of contingent events. But the argument is not prior to experience since all the concepts the argument has used and will use are drawn from experience. The experience from which these concepts are drawn includes knowledge of contingent events, but no such events are not what verify the results of the argument.

A necessary causal connection holds between distinct realities if at least one of them could not exist without the other(s). (That is, there is a necessary causeal connection between A and B if it is true that if A and B do not co-exist, something is and is not what it is.) Three different kinds of connection can satisfy that description. In our example, the sufficient

conditions for the second change could not exist without that change occurring; "could not exist without" describes the forward-looking relation from causes to effect. But is it true that the second change could not exist without the existence of a set of conditions distinct from itself sufficient to bring the second change into existence? If so, "could not exist without" describes the backward-looking relation of effect to cause(s). In the first case, the necessity is from antecedent to consequence; in the second case, the necessity is from consequence to antecedent. The third possibility is that a set of conditions is both sufficient and necessary for the coming about two distinct effects. In that case, the two effects "could not exist without each other," not because there is a causal connection between them, but because there is a causal link between each of them and a third thing, the set of necessary and sufficient conditions, such that if one of them exists, the other must exist.

So far, the analysis has described causal connections in terms of necessity, the sufficiency of causal conditions being defined by the necessity of the effect. But the analysis does not exclude contingency. A's moving through a certain place is a cause of B's moving out of that place. But although A's motion is a cause of B's, B's motion is contingently related to A's motion. The A's moving through that place does not require B's motion unless B is occupying that place. If B were not in that place, A could move through it without requiring B to move. And A could be what it is, a thing in motion through a certain place, without B being in that place. Nothing is contradictory about A being what it is and B not being in that place. So B's motion is contingently related to A's being what it is. B's motion is not contingently related to the entire set of sufficient conditions being what it is, but is contingently related to any element or any collection of elements in that set of conditions short of the entire set itself. They can all exist without B undergoing any change.

Likewise, if every change requires a set of conditions containing

sufficient causes for the change, that necessity does not rule out the possibility of there being many differing sets of conditions, sets of conditions uniting different entities in different configurations, sufficient to cause the change. In our example, billiard ball A causes B to move, but why couldn't the same motion in B have been caused by billiard ball C or a gust of wind or a push from a hand, and so on. Although necessity would link B to some set of sufficient conditions, contingency would link it to this individual set of conditions or that. B could not exist without the existence of some things in a configuration requiring B to occur, but B can exist without A's moving through the place B occupies.

The concept of necessity (in the ontological sense of if X is, then Y is and not is what it is) is useful for describing causality but does not answer all our questions. I have spoken of causal "connections" and "links." These words can lead us to look for some third element standing between A's motion and B's that constitutes the causality of A's motion on B's. In terms of our example, such a third element would be mysterious indeed; for the only things we have hypothesized that require B's motion are A's contacting B and continuing to move the same direction. If there is no "action" at a distance, contact between cause of motion and thing moved is crucial. But contact can hardly be the element we are looking for, since all sorts of things are in contact which do not put each other in motion.

In fact, no such third element is required for causality. We may be tempted to look for something that "goes out from A" answering to the phrase "A's causality." But for causality to hold between A's motion and B's, the only thing new that need exist is the effect itself, B's motion. One of the reasons causality has appeared to be so mysterious is that we don't look in the right place for a further understanding of why change 2 occurs. We need to look at what B is, not just at what A is. B is what is traditionally known as the subject of the change, the subject of a change being that which undergoes the change, that to which the change occurs. The sufficient

conditions for the change include not only A's being what it is A and B having a certain relation to each other, but also B's being what it is. If, for instance, we had assumed that the law against cohabitation of the same place did not apply to B, A's motion through B's space would not require a second change. Or if we had assumed B to be immovable, the second change would have to be different. For example, the second change might be that B shatters, or that A stops moving or bounces off B, and so on.

To illustrate this point with a less far-fetched example, we can use an example from our empirical knowledge. A fist with a signet ring and moving with a certain momentum reaches a point in space. At that point, there is some wax. The fist continues moving, and the wax does not remain what it is; the wax undergoes a change in which it acquires the impression of the signet ring. The fist could have moved through the same place with the same momentum and, if the wax had not been there, that second change would not have taken place. Perhaps some other change would have occurred, but not that one. To understand that change, what we need to understand, in addition to the fist's and the ring's being what they are, namely things moving with a certain momentum, we need to understand what wax is. If it is the nature of wax that it cannot exist in this configuration with the fist and ring without ceasing to be what it is with respect to the shape of its surface, the nature of the wax will tell us why that change occurred. For the fist and ring to be causes of that change, the only thing that need "go out from" the fist and ring is the effect, the change in the wax's surface.

This is not a proof that there is no third element to be called the fist and ring's causality. Nor is it a proof that the nature of wax does necessitate this change and no other. But it is a counterexample showing that the postulation of such a third element requires reasons other than the reasons which, as seen in the example of the billiard balls, make the occurrence of some second change a necessary consequence of the first. This point deserves further consideration, due to the tenaciousness of the

inclination to think of the causality of A or of the fist and ring as something other than A's or the fist and ring's being what it is when it contacts the subject of the second change. From our experience, we expect different second changes to occur if the fist and ring encounter wax, steel, water, glass, sand, and so on. Each such situation is a causal situation like A's contacting B while A is in motion; for in each of these situations, the contact of the fist and ring with something else requires, under penalty of contradiction, a second change; one of the things in the situation must cease being what it is in some respect. A's motion will be one of the causes of that second change. Yet A's motion is the same in each of these cases, while the second changes are all different. Although A's motion is a cause of each of the second changes, A's motion does not cause any of the second changes to be this change as opposed to that. Shall we say that the second change's being this as opposed to that is not caused, that is, is not necessitated by the sufficient conditions necessitating the second change? Perhaps, but should we not then look for some real distinction between, say, B's moving out of its place or the wax's acquiring the impression of the seal being a change and being this change, since its being a change follows from the causal situation while its being this change does not?[2] (If there were no real distinction corresponding to being a change and being this change — e.g., at this speed in this direction, beginning at this point in the ball's duration — causing something insofar as it is a change would also cause its being this change.)

More fundamentally, however, we are asking whether the causing of the second change requires more than A's being what it is and B's being what it is when A contacts B. The second change is required by the fact that, at the moment of contact, A and B are what they are. Their being what they are in this circumstance requires that at least one of them cease being what it is. In other words, in addition to A's being what it is when it contacts B, the only thing that necessitates the change is B's being what it is. If the

change that results happens to be a change occurring B, why must causality involve anything more than that B's nature cannot remain what it is in a certain respect given that A is what it is and is in contact with B? And since B's being what it is is one of the causes of the change and different natures for B would vary the requirement of or the possibilities for the second change, why cannot the nature of B require that this change rather than that (wax's acquiring the impression of the seal rather than glass's shattering) be the change resulting from the situation?

It might seem that all I have done here is replace one mystery with another. To avoid postulating a mysterious tertium quid between A and its effect, I have given the nature of B the ability to impose a kind of necessity outside the mind, and "natural necessity" is an equally mysterious notion. But the example of the billiard balls has already shown that a configuration of things would both be and not be what it is, if a change did not occur in what one of the things in the configuration was prior to the configuration. Why can it not be the case that the what the subject of the resulting change was up to the moment of the configuration's existence necessitates that this change, with its distinguishing characteristics, be the change that occurs. Here, "necessitates" means that the subject would both be and not be what it is up to the moment of the configuration if the change that resulted were some other change. Situations like that of our hypothesized billiard balls constantly occur in our experience. Changes constantly bring into existence situations necessitating some further change happening to some subject of change. And in our experience, different types of things are regularly subject to different types of changes, even though those changes follow the coming into existence of circumstances otherwise the same, circumstances requiring the occurrence of some further change. Wax regularly undergoes one type of change when moving signet rings encounters it, water another type of change, glass another. Again, causality can obtain between distinct realities; situations can occur that necessitate a change in one of the things

making up the situation. And since the subject of the resulting change is one of the change's causes, one of the elements of the situation that would not be what it (the situation) is if some change did not occur, it is reasonable to believe that the nature of the subject of the change is what determines the kind of change the subject will undergo, given what the other factors in the situation are.

And if reasonableness constrains us from postulating more than the evidence calls for, it is unreasonable to postulate more in the other causes of a change than their being what they are. A's causing of B's motion is the fact that B's motion results from a conflict between A's being what it is and B's being what it is. The conflict is such that not both of them can remain what they are. If B is the one that does not remain what it is, B's ceasing to B what it is results from A's being what it is.[3] On any analysis, the only new element the evidence so far presented requires is the change resulting from the conflict between A and B remaining what they were up to the moment of their contact. Words like "action," "causal connection," "causal link," and so on are best understood as referring to resulting changes (effects) and characterizing changes by their resulting from situations that necessitate the occurrence of a further change. A's "causing" of B's motion, or "production" of B's motion or "action" on B is B's motion itself viewed as having the relation to A's being what it is that I have described; it is B's not remaining what it is under the relation of being required, in this circumstance, by A's being what it is. In other words, A's action is B's motion as arising, emerging, or deriving from the conflict between A's being what it is and B's continuing to be what it is. Whether this is what is in the mind of people using causal words in ordinary speech is another matter. But this interpretation can save what needs and deserves to be saved about the situations ordinary speech intends to communicate by means of causal language.

For example, when a hammer drives in a nail, the action of driving the nail, if by that we mean something other than the hammer's being what it is,

is the motion undergone by the nail with the relation of resulting from the hammer's being what it is. And for understanding causality, it is best that we mean by "action" something other than the hammer's being what it is. For the hammer can be what it is, that is, something moving with a certain velocity and momentum and a certain place, without the action of driving in the nail occurring, since there need not be any nail at that place. And if there is no nail at that place, there is no action of driving in a nail, even though what the hammer is is the same in both cases. What can make this reading confusing is the fact that we rightly think of the motion of our arm holding the hammer, which is the same in both cases, as an action of ours. But what makes the change undergone by our arm an action of ours is the fact that we cause that change by previous changes which bring into existence sufficient conditions for that change to occur in our arm. Thus, to call the change undergone by the arm our action is best understood as a description that relates that change to ourselves, or parts of ourselves, as containing sufficient conditions for the occurrence of that change.

Since the biological conditions for such a change more complex than are those for ordinary mechanical changes, it will be easier to understand the point if we think of the hammer being swung by an arm on a machine. We have taken away the nail, so there is no action of driving in a nail. There is an action of swinging a hammer attached to an arm. But detach the hammer, and there is no action of swinging the hammer even though the arm is still moving. The motion of the arm is an action of the machine, but to call it such is to say that that motion is an effect produced in the arm by other parts of the machine, for example, a drive shaft. Detach the arm from the drive shaft and the action of moving the arm no longer exists, even though the drive shaft is still rotating. The rotation of the drive shaft is an action of the machine. But the same analysis of action applies to the rotation of the drive shaft and to all the preceding changes in the machine that bring into existence sufficient conditions for subsequent changes.

Another supposed dilemma concerning causality concerns whether cause and effect are simultaneous. Hume thought that if cause and effect were simultaneous, all time would collapse. All subsequent effects in any causal chain would occur at the same time as the initial causing, so no causal chain could produce a temporal sequence. On the other hand, if cause and effect are not simultaneous, the cause would no longer exist when the effect was produced, and the effect would have a nonexistent for its cause. The billiard ball example allows us to know that the coming to be of the effect can be simultaneous with the existence of sufficient conditions for the effect. B's beginning to move out of its space is simultaneous with A's beginning to move into B's space. The same simultaneity would hold for whatever change fulfilled the requirement for some change 2 to result from change 1. For example, if what permitted A to move into B's space was B's shattering or ceasing to be a body whose space could not be occupied by another body, that change could occur no later than A's beginning to enter B's space.

On the other hand, what is there to prevent the motion that A causes in B from being a motion that continues in B when B is no longer in contact with A? If B's motion can continue, A has produced an effect remaining in existence through time. And that effect can itself become a cause of further change at a later time. For example, sometime after being put in motion by A, B can hit another billiard ball C, causing C to undergo a motion that can continue through time. Thus, both the causing of a change can require a previous change that takes time, and the effect of a previous change can remain in existence through time. Yet, the beginning of the later change can be simultaneous with the coming into existence of sufficient conditions for it as a result of the prior change. Experience abounds with examples of sequences of this sort. The weight of the straw gradually increases until another change occurs; the camel's back breaks. The camel falls until it hits a lamp. The lamp rolls away until it meets some wood. The wood catches fire. The heat of the fire grows until some water begins to boil. And so on.

All of this supposes that the sufficient conditions brought about by a prior change can cause a change that can continue when those conditions have ceased to exist. Ancient and medieval philosophers thought this was impossible. They interpreted their belief that every change must have a cause to mean that an extended process of change must have a cause as long as it exists. Consequently, they searched for ways to explain why projectiles remained in motion after their contact with their projectors had ceased. We cannot rule out a hidden cause for a projectile's motion a priori. But the principle that every change must be caused does not require them.

To the extent that a motion is continuous, the motion is remaining the same is some respect. For example, if a motion continues in existence with the same speed an the same direction, it has not changed with respect to speed and direction. If there is no change in these respects, the principle that every change has a cause does not call for any additional causality to account for a body's continuing to move with that velocity. The body's change from being at rest to being in motion at that velocity would require a cause, for example, billiard ball A's striking billiard ball B. But since that change brings into existence something that remains constant, the principle of causality does not of itself give us a reason to look for a further cause maintaining the existence of that which is constant.

Empirically, we know that bodies do not maintain constant velocity with no further causality, since the universe is an environment of forces of acceleration and deceleration. But since acceleration is a change in that which would otherwise be constant, the principle of causality requires that changes in velocity be caused. There are instantaneous changes and continuous changes. The principle of causality requires that the instantaneous change from rest to motion be caused and that the instantaneous change from one velocity to another be caused. Of course, an instantaneous change in velocity can be the beginning of an acceleration that continues at a constant rate. Could we not conceive an acceleration at a constant rate as something that

continues in existence without any additional causality, the way motion at constant velocity continues in existence? To the extent that the rate was constant, the acceleration would be something unchanging and not covered by the principle of causality. This is another possibility which we have no a priori way of ruling out. But before we knew empirically that constant acceleration requires continuous causality, the principle of causality would tell us that the instantaneous change beginning a new rate of acceleration must be caused.

For the principle of causality, we could use something like the formula "Any instantaneous change to a new rate of acceleration must be caused." But that would rely too heavily on empirical knowledge for my purposes at the moment. I am inquiring how we know a cause is needed for a change bringing into existence something that can remain in existence after the change. Any change brings into existence new state of affairs that can remain in existence after the change. If every change has causes, those causes are also causes of the state of affairs resulting from the change. For if sufficient conditions for a change cannot exist without the change occurring, they cannot exist without the existence of the state of affairs the change brings about. It turns out that, were it not for the presence of other accelerating and decelerating causes, the motion billiard ball A causes in B would be a constant state remaining in existence after B's instantaneous change from being at rest to being in motion. So it turns out that A causes B's continued change only in the sense that it causes the instantaneous change from which B's continued change comes. But we acquire knowledge such that it is the presence of other causes that prevent B's motion from being a constant state by looking for causes on the basis of the belief that instantaneous changes must be caused (and I will argue below that verifying the results of such research also presupposes the principle of causality). The results of looking for causes of change refine our belief that they must exist but do not contradict it. For the causes of instantaneous change by that fact are causes

of the result of the change, even if the "action" of the causes exists only when the instantaneous change exists, since their action is identical with that instantaneous change.

The principle of causality, then, requires an efficient cause for the beginning of a state of affairs, when that state of affairs comes from a change in a previous state of affairs. Sometimes the new state of affairs is a continuous process of change, for example, something's moving with constant velocity or accelerating at a constant rate. The principle of efficient causality tells us that the beginning of such a process must be caused but does not tell us whether or not its continuation requires to be continuously caused. Pursuing the implications of the principle of efficient causality, we learn that a constant rate of acceleration requires continuous causality. But we also learn that, were it not for such causes of acceleration, motion at a constant velocity would not require continuous causality, would not require the continued existence of a state of affairs other than the motion such that this state of affairs cannot be what it is without the motion's existing.

Any instantaneous change puts its subject in a new state. Its subject gains or loses one or more qualities or characteristics. As a characteristic of a subject, the state is something that would not exist without the subject. Without billiard ball B, B's motion would not exist. Some other change might exist, but as the result of a situation that requires one of its members to cease being what it is, that other change would still be something happening to a subject and hence something that would not exist without its subject. Since the subject of the change and of the state coming into existence through change is something distinct from then without which they would not exist, the subject is one of their causes. Aristotle called the subject a material cause, the state a form, and described change as the acquisition or loss of a form by a material cause. This terminology has its drawbacks, and I have elsewhere suggested an alternative terminology which, inevitably, has its own drawbacks. Here, I will use the Aristotelian terminology, not out of piety,

but in order to over burden you as little as possible with unfamiliar jargon.

"Material" or "matter," of course, do not here refer to physical nature in general. Aristotle used them analogously to the sense in which we speak of a artist's materials, the things an artist works on to make something new out of them. Likewise, "form" does not mean something contentless, an empty structure or outline gets its content from somewhere else. Rather form gives specific content to an otherwise undetermined matter, the way a sculptor makes a lump of clay into a beautiful or useful object by the form he puts in it. Thus, billiard ball B's motion is one specific determination out of an infinite number of possibilities for speed and direction.

The Aristotelian terminology is pertinent because of the next point. The state brought into existence by a change relates to the subject of the change as act to potency. The subject only undergoes changes it is capable of undergoing; it only acquires new characteristics it is capable of having. For example, if we had assumed ball B was immovable, the second change resulting from A's change could not be B's motion. While the subject of a change, by hypothesis, is something after the change that it was not before, the subject of the change also remains the same in some respects; what exists after the change is in some ways the same as what existed before. If not, we are not dealing with change, where what is new comes out of what already exists, but with total annihilation of the old and its replacement with something having no connection with the old. In the latter case, we could not say that something has changed into the newly existing thing. And since something remains in existence throughout a change, the new state the subject acquires must be compatible with the features the subject already possesses. Therefore, the subject must be capable of the changes it undergoes; its nature must give it the capacity for the forms those changes bring about.

For example, a billiard ball is capable of changing in many ways while remaining a billiard ball. It may change in temperature, go from rest to motion or motion to rest, go from motion in one direction or speed to motion

in another, etc. But a billiard ball cannot undergo a change making it a drop of water or lump of coal while remaining a billiard ball. The atoms or subatomic particles composing the billiard ball might undergo changes making them parts of a drop of water or lump of coal, but they would cease being parts of a billiard ball in the process.

Today, Aristotle's potencies would be discussed, along with other things, under the heading of "dispositions." There is nothing wrong with doing so as long as we do not attempt to define dispositions solely in logical or epistemological terms. For example, Ryle tried to reduce beliefs about dispositions to beliefs about whether certain contrary-to-fact conditionals were useful as "inference tickets." At that rate, esse, for dispositions, would be percipi, since conditionals and inference tickets are features of our knowledge of things, not features of the extramental things known. That a contrary-to-fact conditional is true or an inference ticket holds may be a necessary effect of a thing's having a certain disposition, but its having the disposition cannot be identical with the fact that a contrary-to-fact conditional is true or an inference ticket holds. A spherical thing's having, and a cubical thing's not having, the capacity to roll does not result from these things being made objects of knowledge.

Instead, a thing's potentialities are identical with its nature, that is, with the extramental features it possesses. What a billiard ball is is a capacity for being something hotter or colder, at rest or moving, moving faster or slower, this way or that, and so on. If the thing ball A struck was not spherical but cubical, the second change could not be this thing's rolling out of A's way. The features making a thing what it is are actualities, achievements, since they already exist. But what is an accomplishment in one respect can be a disposition for further accomplishment in another. When we mold clay into a particular shape, we fulfill a potency, but we also give the clay a new set of potentialities it would not have if we had given it another shape.

Dispositions, capacities, potentialities, in other words, are concepts that can express what things are in their extramental existence. Specifically, they can inform us about extramental causal connections. For aside from considerations of efficient causality, we know that change and the states of affairs resulting from change are caused realities. We know this because we know they would not exist without something other than themselves, the subject of the change. And the result of a change, for instance, ball B's being in a state of motion, would not exist without two things not identical with itself, B and the state of motion. Although it could not exist without B, the motion is not identical with B, since B can exist without undergoing the motion. Therefore, the event of B's moving is a union of distinct realities without which that event would not exist.

The concepts of potency and act, of capacity and fulfillment of capacity are necessary to understand that causal union. Contrast the union of B with its motion to another kind of union. Assume that instead of causing B to move, A had simply rolled up to B and stopped when they were in contact. We might call the result of the first change a union of A and B, AB. Does the whole, AB, contain any reality over and above the distinct realities of its components, which realities were already in existence before the change? Or, is their constituting a whole a logical construct imposed on their distinct realities by a knower, the way the set [A,B] is a logical construct out of A and B? What is new in the situation of A and B being in contact is their spatial relationship. But many would want to deny reality to relations as such, for reasons that are far from trivial.

I will return to the issue of spatial relations. But for the moment, notice that in the case of B's undergoing motion, neither the existence of a new extramental reality, the motion, nor the extramental character of its union with B is in doubt. Questions about the extramental character of the whole constituted by the union of B and its motion do not arise as they can arise in the case of AB. B's motion is something of B; for it is a

fulfillment of a capacity of B to be something in motion. Someone who thinks that AB's being a whole is something in the mind of the beholder might want to say that outside the mind, AB does not constitute a unified whole but a mere juxtaposition of unrelated things. In that sense, the union of B and its motion is not a mere juxtaposition of unrelated things. When we describe B as having the capacity to be in motion, we are describing what B is. Before that motion exists, that way of describing what B is makes use of reference to something not as yet real, the motion. But if the capacity so described were not something real, the motion never would exist.

The fact that we can describe a feature of B as a relation to a nonexistent term simply means that this relation is not a supervening reality distinct from what B is, the way A's relation of being next to B would be a supervening reality, if there are such realities. It may make no sense to say that A acquires a new characteristic which is nothing but a way what A otherwise is now respects something that does not exist. B's capacity for motion, on the other hand, is more than something relating things nonidentical with itself, what B is and something else that does not exist; it is identical with a feature or features constituting what B is, features which are actualities and are not just ways of relating things distinct from themselves. (E.g., B's feature of being cubical instead of spherical is an actuality and not just a way of relating to B's sliding instead of rolling, but it also what constitutes B's capacity for sliding instead of rolling.) And if it made no sense to say that what B is related B to nonexistent things like motions by the kind of relation we call "capacities" or "potentialities," those motions could never become existent.

Likewise, the motion's relation of dependence on B, its incapacity for existing without there being something distinct from itself that undergoes it, is not a supervening reality different from the motion. Unless we can rule out such a tertium quid on other grounds, there might be a relation supervening between B and its motion when they are united. But that relation

would not constitute the reason why the motion needs a subject. The necessity for the motion to have a subject consists in the fact that a motion would not be what it is without it. Without a subject, A motion undergone by something would not be a motion undergone by something. (If the subject does not exist, a motion undergone by the subject does not exist.) Therefore, a motion's relation of dependence on a subject is one with what the motion is. And if a motion acquires a supervening relation to its subject, when it is united with its subject, we can ask whether the motion's acquisition of that relation is necessary. For it to be necessary, the motion could not be what it is without acquiring that relation. On any hypothesis, the motion is identical with a relation of dependence on something other than itself, its subject.

Obviously, there is an important difference between the capacity for the motion that is identical with what the subject is and the dependence on the subject that is identical with what the motion is. The identity of the subject's potencies with its nature allows its potencies to relate it to what does not as yet exist. The identity of the motion with its relation of dependence implies the opposite. A motion's relation of dependence on a subject is a requirement that there really exist something undergoing the motion. The motion's dependence on a subject is a relation to a really existing term. The real existence that is the term of the relation is not the existence the subject had before it came to be in motion, but the existence it has at the time of the motion. Possibly, a subject and its motion could come into existence at the same time, so that the subject would have had no preexistence. What is not possible is that the motion exist at a time when a subject for it did not exist.

Apart from considerations of efficient causes, then, a thing's being in motion is a caused reality, a reality that would not exist without other realities, the thing and its motion, non-identical with the whole they make up. What distinguishes the union of these realities from a mere juxtaposition of otherwise unrelated things are relations that each member of the union has

to the other by its nature. While A might be in no way affected by occupying the space next to B, the motion affects its subject, is a modification of its subject. Specifically, the motion fulfills a potentiality of the subject, so that the subject is now different from what it was before in that respect. And unlike A and B whose existences may be independent of one another, the motion's existence is entirely something of the subject, something in the subject, where "in" does not denote spatial encompassment but a relation of dependence one with the fact that the nature of the motion is to be a fulfillment of a potency constituting what something other than itself is.

Before turning to the question of a change's need for an efficient cause, it will be best to deal with one more difficulty concerning change. Again, assume that ball A moves in the direction of B and stops when it arrives at B. Since a change has occurred, something new must exist, something that did not exist before the change but that the change brought into existence. But the only thing new in our hypothesis is the spatial relation of A and B. Does the fact of locomotion, then, require us to assert the existence of spatial relations? When something moves next to or away from another, must there be realities corresponding to the words "next to" and "away from"? Unquestionably, the distance extending between any two physical objects is a reality of some sort. But if A is at a spot five feet from B, what is it for A to be "at" that spot. Is "at-ness" a quality supervening on what A is otherwise, specifically a quality relating A to this place is space and no other?

Perhaps it is. But for those who are understandably reluctant to admit the real existence of supervening relations, I want to show how the admission of causal connections can help. Since the new states of affairs resulting from change are realities, the problem is to find some reality that can result from change other than new spatial relations between things. In fact, we know from science that all things exercise causal influences on their environments, where "causal influence," again, denotes changes in the environment resulting

from the things in the environment being what they are. At a minimum, when anything changes place, the electro-magnetic and gravitational fields in the universe become different from what they were and different from what they would be, ceteris paribus. These changes are extramentally real and result from the fact that things change place relative to one another. Again, I am not denying that places, in the sense of areas of the universe's extension, are realities. I am only suggesting an alternative candidate for what comes to be and ceases to be as things move from one place to another. Instead of the new realities being relations answering to words like "in" or "at," they can be changes, like changes in fields of force, caused by what the things in particular places are.[4]

This account might appear to beg the question, since "force" is a causal term accounting, ultimately, for changes in place. But the reality of things being in motion from one place to another is not in question. What is in question is whether the new state of affairs that exists at each moment of the motion is a distinct entity describable as a spatial relation. When a force puts something in motion, the thing that is changing place is itself something that has causal influences on its environment. Those influences can cause the change in place of other things that also have causal influences on their environments. At bottom, what changes place are elementary particles that bear different kinds of forces. A change in place of such a particle requires a change in the force fields of the universe. For particles and fields to be bearers of forces is for what they are to require that other things not remain what they are, when particles and fields exist in certain configurations with those other things. And since particles always exist in configuration with a surrounding spatial environment, particles continuously cause a change in their surrounding area, the result being the continuous existence of a field of force in that area, a field such that, when other things are in the area,[5] they undergo certain changes.[6]

This way of avoiding spatial relations assumes the existence of causal

connections. The argument I will use to show the necessity of causal connections will not depend on whether or not there are spatial relations. It will only depend of the fact of change. On the other hand, those who wish to avoid spatial relations have no alternative but to admit causal connections of some kind, since change in place requires that something come into being (or cease being) from the change; otherwise, nothing has changed. And for change in place to require that something other than itself come into existence is for there to be a necessary causal connection.

2.3. Why There Must Be Necessary Causal Connections

The argument by which I have defended forward-looking causal connections is far removed from the fact of regular concomitances. Our belief in the necessity of causes for change, backward-looking causal connections, is equally far removed from the occurrence of regular concomitances. As children, we learn we can expect to find that a change resulted from a state of affairs brought into existence by a previous change. But children do not believe that all changes have causes as a result of their observation that all changes obey universal laws. In fact, for the vast majority of changes we experience, uncontrolled observation alone does not show that they are instance of universal laws. (And we use controlled observation because we assume the existence of causal connections; we use controls to identify which factors have a causal connection to the phenomena under study by screening out factors that may exist but do not have a causal connection to it.)

Most changes result from situations in which multiple causes, each obeying their own universal laws, combine to produce individual effects not covered by those laws. I now know that everything I see, as I look out my window, resulted from the combined action of causes that obey universal laws. But no universal law tells me that there should be a tree at that place on my lawn, that there should be a brown patch of grass three yards to its left, that a robin should be landing on its lowest limb at the same time that a dog starts barking, that the sky behind the tree is cloudless today, and so on.

Our daily experience is composed of such unique events, events that fall into patterns only after considerable reflection or investigation. Yet we believe that the changes bringing such states of affairs about have causes well before we have reflected sufficiently to discern many patterns, certainly well before the concept of cause we believe in could be analyzed by regular concomitance. And the reason that we undertake investigations that will later yield regular patterns is that we already believe that there are causes for those investigations to find.

Until we are conditioned to think otherwise by our philosophic education, we believe that every change has a cause? What, if anything, is the basis of that belief, since it is not based on the observation that changes universally embody universal laws? When their belief in the necessity of causes is challenged, students of philosophy sometimes reply with an argument like the following. A billiard ball is capable of beginning to move at an infinite number of times and is capable moving in an infinite number of directions at an infinite number of speeds. There must be a "reason" why the ball began moving at this time and with this velocity, since infinite other times and velocities were possible, and that reason must be something other than what the ball's being what it is. Something other than the ball must explain why the ball began moving at this time in this direction at this speed, since what the ball is is compatible with all other beginning times, speeds, and directions as well.

The informed teacher will answer that indeed something other than the ball must explain the ball's motion, if we assume that the ball's motion needs an explanation. But that is the point under dispute. Why must there be any explanation or reason for the ball's motion, where "explanation" and "reason" mean a cause of the kind we have illustrated by billiard ball A. The teacher will grant that, if a ball's beginning to move at this time with this velocity or that needs a cause, the ball is not that cause, if its nature is equally compatible with all other beginnings and velocities. But the teacher will not

grant that we know a ball's motion needs any cause.

More than that, the informed teacher will argue not only that we do not know that a cause is necessary the ball's motion, but also that we cannot know this. Necessity is shown by the contradictoriness of the opposite. Contradiction affirms and denies that a thing is what it is, that A is not A. But when we deny the existence of a cause, we are not denying that A is A, we are denying the existence of something than A. A's identity with itself can show us only that when A exists, A exists. To deny the existence of something other than A is not to deny that A is A. But when we deny the existence of a cause of A we are, by definition, denying the existence of something other than A.

****

(Here the change is named M, the subject S and the efficient cause A.)

The fact that M would not exist without S is the fact that M would not exist without S's having become M's component cause after only potentially being M's component cause.

There are only two possibilities: 1) S's being M's component cause is a necessary effect of the features constituting what A and S are without M; or 2) S has no cause making it M's component cause. In that case, M either has no component cause or M is that about its component cause, that feature of its component cause, that constitutes what it is for its total cause to be its total cause.

M is that which differentiates that which is not something M has for a cause by its other features from what M now has as a cause, from what now has the intrinsic feature of being M's cause.

Both with A and without A, M is included in that which M has for a component cause. But if M only depended on S, M would be its own cause by being that about S that constitutes S the only thing M depends on. But with A, M is not included in that which it has for a component cause in a way that constitutes its total cause to be its total cause. For it has a sufficient total cause in what A and S are apart from M.

M is caused to make S its component cause by something which is not constituted a

cause by having M as a feature.

The fact of the subject's being M's component cause has an ensemble of causal factors, of which the subject's being M's component cause is not a part, sufficient to cause that fact.

Without an efficient cause, the total cause of M is the total cause of M only if it includes M.

But if we did not analyze the situation any further, we might consider the union of the change and the subject to be only a mereological sum, and mereological sums are logical constructs just as relations of reasons like "desired by" are.

1. Assuming that A and/or B remain in existence and are not completely annihilated. Creation ex nihilo and total annihilation would not be changes. What I mean by "change" requires something that undergoes the change and, hence, continues to exist after the change but in a new state.

2. More on this possibility below.

3. In fact, we know that, in nature, both A and B undergo changes. But that simply means that B is not the only element of the situation that cannot remain what it was up till the moment in which sufficient conditions for the change exist. B has an equal and opposite causality on A. And B's causing of A's change is the fact that A's motion results from a conflict between B's being what it is and A's being what it is. A's ceasing to be what it is results from B's being what it is.

4. It might seem that causal connections cannot replace spatial relations as the results of locomotion, since different locomotions can lead to the same causal connection. For example, the closer an iron filing comes to a magnet, the greater is the attraction of the

magnet on it. But that change in distance and attraction between them could be achieved by different locomotions, the movement of the filing toward the magnet or the movement of the magnet toward the filing. Since those changes in place differ, they must bring about different realities, for example, the filing's being at a place where it was not before or the magnet's being at a place where it was not before. Consequently, the results of these different changes in place cannot be replaced by the common change in the attraction between the filing and the magnet. The answer is that locomotion is change in place relative to the entire universe, not just to this or that thing in the universe. The causal connection between the filing in the magnet is the same for each of the locomotions, but not the causal history of the universe as a whole. If the filing changes place, the universe's force fields undergo one series of changes; if the magnet changes place, the universe's force fields undergo another series of changes. Moreover, although empirical knowledge cannot distinguish motion relative to the universe from the universe's motion relative to something in the universe, from the point of view of ontological knowledge there is a difference between, say, a billiard ball's motion relative to the universe and the universe's motion relative to the billiard ball. If the efficient cause causes the ball to move, the motion of the ball has a relation of dependence on the ball; if the efficient cause causes the universe to move, the motion of the universe has a relation of dependence on the universe. The fact that empirical evidence does not allow us to determine which motion is actually occurring is an epistemological fact, not an ontological one.

5. The phrase "in the area" may appear to reintroduce spatial relations at a level prior to the change caused as a result of something's being in the area. But the reality of places in the sense of specific areas is not in doubt; nor is the reality of something's being at or in an area, as long as its being at or in an area does not require a distinct entity describable as a spatial relation. Since the reality of extension is not in doubt, we can replace phrases like "in the area" with phrases like "when the spatial location of something relative to the rest of the universe is the same as the location of this area relative to the rest of the universe" or "when the spatio-temporal distance of something from some point of reference is this place in this field of force."

6. When the forces on a body are in equilibrium, the body does not change place. Given

that the body is what it is and that it exists in a certain configuration with other things that are what they are, it is not the case that the body must change in place; instead, it is the case that the body cannot change in place. If it did change in place, then either it, some other thing in the configuration, or the configuration would not be what it is. So if a previous locomotion results in the forces on the body being in equilibrium, we may have no reason to say that the previous results in any further change to that body. But it does not follow that the only change resulting from the locomotion is change in spatial relations. As mentioned in note 4??, we have to consider the results of locomotion relative to the entire universe. Even if the state of the universe were exactly the same at the beginning of a locomotion and the end (think of circular motion, for instance), the causal history of the universe would be different, ceteris paribus, since force fields of the universe would be changing throughout the motion.