A does not cause C.

Consider the following example of a causal chain: three dominoes stand in a line; the first domino is caused to topple (event A); the first domino then strikes the second domino causing the second domino to topple (event B) the second domino then strikes the third domino, causing the third domino to topple (event C). We say that A causes B and B causes C. The question that arises is: does A cause C? If it is true that event A causes event C then the causal relation is transitive.

There are several reasons for thinking that causal relation is transitive. These include: our ordinary use of causal language, David Lewis's counterfactual theory of causation, and determinism.

The debate is not completely one-sided; early in the 21st century philosophers started to put forward counterexamples that purport to show that causes are not transitive.

In this essay I argue that causes are NOT transitive.

Following the introductory sections I shall discuss reasons for and against believing that the causal relation is transitive; in the second section I attempt to nullify these reasons; this is to clear the decks for arguments that show that the causal relation is not transitive. It could be argued that only the last section of this essay is necessary. I have structured in the way that I have with the hope being persuasive.

General problem

The significant problem facing anyone writing about a property of causal chains is that philosophy has not been able to establish an adequate account of causation. In particular, philosophy has not been able to establish what a cause *is*. In fact there isn't even agreement as to whether the relata of the causal relation are immanent (in space and time) or transcendental. The Stanford Encyclopedia of Philosophy entry entitled: The Metaphysics of Causation (Schaffer 2016) describes the full horror story.

Most of the Introduction to the Oxford Handbook of Causation consists of reasons why there is still very little agreement on the most central question concerning causation: what *is* it?

In addition to the problem of 'cause' not being well defined, there is, in my view the additional difficulty relating to applying the concept of transitivity to causal chains. The concept of transitivity is usually introduced by making use of mathematical relations such as 'greater than' or 'equal to'. These mathematical concepts are applied to numbers on a line. It is easy to see that if value X is greater than Y, and Y is greater than Z, then it must be true that X is greater than Z. The difficulty is, that causes in a causal chain are not points on a line. Causation may have a number of dimensions.

One characteristic of causal chains is that causes occur in temporal order; consistent with Hume's regularity account of causation. Temporal relations such as 'before' are transitive. But there is more to the concept of causation than temporal ordering. The temporal ordering of causes in a causal chain does not prove that the causal relation has the property of being transitive. But perhaps this temporal ordering does contribute to a belief that causes are transitive; especially if a causal chain is of a type that exhibits regularity.

For transitivity: causal language

Normal use of causal language suggests that causal chains are transitive. We might say, for example, that moving the switch causes the lights to come on. We do not usually give any thought to power stations, distribution systems, electrical circuits or the flow of electrons. Someone defending this

A does not cause C.

position might say that our normal use of causal language gives us the meaning of causation and that this entails that causes are transitive.

For transitivity: counterfactual theories of causation
Counterfactual theories of causation rest on the idea that the meaning of causal descriptions can be explained in terms of counterfactual conditionals that have the form if A had not occurred then C would not have occurred. In other words A is causally dependent on C. For example, if the first domino had not toppled the third domino would not have toppled.

The key feature of these theories is that the relation of counterfactual dependence does capture the fact of a causal relationship; e.g. the fact that if the ball had not been thrown, then the window would not have broken - does suggest that the ball did cause the window to break.

Counterfactual theories also have the supposed advantage that they can accommodate causes that are things that are absent and events that did not happen. For example, 'the car stopped because it ran out of petrol' is amenable to counterfactual analysis; i.e. the car would not have stopped if it had not run out of petrol.

It can be seen that counterfactuals are transitive. Early in the development of the analysis of causation by counterfactuals by David Lewis, this result was welcomed, because it was thought at the time that causes are transitive.

For transitivity: determinism
Determinism is the thesis that a state of the world in conjunction with natural laws necessitates all other states of the world.

The state of the world that included an event A (first event in a causal chain) determined a later state of the world that included event B, and that this state determined a later state that included event C. But also note that the state of the world that included event A also determined (independently of the state that included event B) the state of the world that included event C. Two things follow: a) determinism is a transitive relation; b) the fact that determinism encompasses causal chains in conjunction with the transitivity of determinism, at least suggests that causal relation is transitive.

It might be argued that the world is not completely deterministic, but the literature suggests that philosophers are mainly interested in causation as it applies to normal life. Our experience of causal regularity and the fact that there are respected accounts of causation, including Hume's – that rest on regularity, shows that most events are determined.

Against transitivity: counterexamples
So far the weight of evidence is in favour of the causal relation being transitive; however, philosophers have put forward a number of examples of causal chains, believed to show that causes are not transitive (counterexamples with respect to causes being transitive). These examples are important because, as far as I am aware, these examples are the only reason philosophers have for believing that the causal relation is not transitive. If these examples of causal chains did prove that the causal relation is transitive, the debate would be over and there would be no point to this essay.

A does not cause C.

Nullifying support from causal language

Causal language is employed to provide explanations. The problem is that nobody knows what a cause is. To make use of causal language we have to skip round this problem. One way of doing this is to enquire: what is it that made a difference under the circumstances that prevailed. It does not seem to matter in our use of language what this thing is. And because we don't know what a cause is, we have no way of verifying that the thing identified as a cause, is anything of the sort. An example will illustrate the point I am trying to make. A car that was behaving normally runs out of petrol; the car comes to a stop. The thing that is different between the car moving forward and the car stopping is the absence of petrol. So the cause of the car stopping is that it has run out of petrol or at least this is what we say. This appears to be an acceptable explanation.

But what are the metaphysical implications? What is being asserted is that something that does not exist is causally effective. Surely this analysis indicates that what we sometimes say, when we use causal language, is nonsense. Unfortunately there are many that do take such utterances as being meaningful, and additionally take it as important that we should do so. Apparently Lewis amended his counterfactual theory to accommodate absences.

The requirement that communication be effective and efficient results in some of our causal utterances being untrue. At least some of the causes we refer to are not genuine causes. Therefore our use of language is not a reliable guide to the transitivity of the causal relation.

Incidentally, if causes are transitive, in the case of the car that ran out of petrol, it would be correct to say that starting the car engine caused it to stop! I suppose a supporter of the transitivity of causation might decide that the best response would be to agree.

Nullifying support from counterfactual theory

There is no doubt that counterfactual theory has attractions; it is coherent with our use of causal language – we do make claims such as "the car stopped because it ran out of petrol". But as indicated in the section above there is doubt as to whether such claims are true. It can hardly be an advantage that the theory is coherent with utterances that are not true.

In the next section I discuss determinism. If determinism is true then, in principal all changes in the world can be understood in terms of events in time and space involving physical objects and natural laws. To resort to counterfactuals to explain change in a physical world does seem somewhat desperate. In his 1973 essay 'Causation' Lewis says: *"I have no proof that regularity analyses are beyond repair… Suffice it to say the prospects look dark. I think it is time to give up and try something else"*.

I do not believe that any criticism I might have of counterfactual theory counts for much, but professional philosophers have raised numerous difficulties. Amongst these difficulties is the possibility that the causal relation is not transitive.

Initially, the property of transitivity of counterfactuals was welcomed, because it was thought that causality is transitive, but from the beginning of the 21st century the transitivity of the causal relation had become controversial. Philosophers have put forward examples of causal chains that purport to demonstrate that causation is not transitive; however these examples are themselves controversial. I believe the controversy is justified; my view is that these examples are defective.

I believe that my own analyses of uncomplicated causal chains (see below) show that the causal relation is not transitive. This is possibly a fatal problem for counterfactual theories.

A does not cause C.

Nullifying support from determinism
What I am attempting to do here is compare and contrast the concepts of 'determinism' and 'causation' with the aim of showing that while determinism is transitive, the causal relation is not

The first difference that comes to mind when considering these concepts is the difference in the the relata. The relata of determinism are states of the world, the relata of causation are usually taken to be events involving a small number (often two) objects, i.e. very small parts of the world (although perhaps under the influence of universal laws). However, there is nothing in the definition of determinism that specifies the size of the world, so presumably we can apply the concept to a world that consists of three dominos sitting on a large mass (to provide a surface for the dominos to stand on and gravity). We can say that the toppling of the third domino was determined by the state of the (shrunken) world that included the toppling of the first domino; however it would be a mistake to think we are justified in saying that the toppling of the first domino, as an isolated causal event, necessitated the toppling of the second and third dominos. For the first domino to topple the second domino, there needs to be a second domino in the right position. The second domino in the correct position is part of a state of the world that the concept of determinism applies to. An isolated causal event in itself does not necessitate anything. What happens when a mouse button is pressed is determined by the state of the system (hardware and software) that the mouse is connected to. Pressing a mouse button, when considered as an isolated event does not necessitate anything.

So while a state of the world does necessitate all future (and past) states of the world and determinism is transitive, an isolated cause does not necessitate anything; thus the causal relation is not transitive.

Nullifying support for nontransitivity from counterexamples
To avoid confusion, I repeat that I am arguing against the analysis of transitivity counterexamples from the literature; counterexamples that might be taken to support my position.

The difficulty is that I have not found any of these counterexamples convincing. All are disconcertingly contrived, some are complex, and others do not provide a complete description or accurate description of the causal chain under consideration. The examples all seem to involve some sort of trick; for example, causes that do not bring about a change to the final event in the chain or some threat that does not materialise. Here is an example (Schaffer, 2016):

> *For instance, to borrow an example from Hall (2000, see also Hitchcock 2001), suppose that the boulder begins to roll down the hill towards the hiker's head (c), which causes the hiker to duck (d), which in turn causes the hiker to survive (e). It seems that c causes d and that d causes e, yet it does not seem that c causes e …*

In other words: the boulder rolling down the hill saved the hiker's life. Taking causes to be transitive leads to absurdity, and that is taken to prove that the causal relation is not transitive.

To see that this example is defective, we can analyse it in terms of changes that are actual. A change occurred that resulted in the boulder starting to roll. Completely independently, light that was emitted from the sun eight minutes prior to the boulder beginning to roll was reflected by the boulder (another change), resulting (after a few more changes) in the hiker perceiving that there was a boulder rolling down the hill towards him. So the rolling boulder did not cause the hiker to duck. The duck was caused (taking the causal relation to be transitive as we would do in normal use of language) by the light reflected off of the boulder. If the hiker had been walking in complete darkness he would have died – the rolling boulder would not have caused him to duck.

The question that arises is: if the causal relation is not transitive, why does it need contrived examples to show that this is the case? If causes are not transitive, it should be possible to see this

A does not cause C.

lack of transitivity in uncomplicated examples, such as the three dominos example (call this the uniformity argument). It follows that the challenge is: to show that causation is not transitive for a line of three dominos. This is discussed in the next section.

Causes are not transitive
Let us return to the dominos example: three dominoes stand in a line; the first domino is caused to topple (event A); the first domino then strikes the second domino causing the second domino to topple (event B) the second domino then strikes the third domino, causing the third domino to topple (event C).

The question is: does 'A' cause 'C'?

Now it is true that most people say that 'A' causes 'C'; implying that the causal relation is transitive. But notice the first domino does not come into direct contact with the third domino. The change in the world described as 'A' causes 'C' does not exist! The first domino does not strike the third domino.

I know from experience that people have difficulty with this. To make things easier, see what happens when 'cause' is substituted with 'push': The first domino is pushed and topples (event A), it pushes against the second domino which topples (event B) which pushes the third domino, which topples (event C). We say that the first domino pushes the second domino and the second domino pushes the third, but we would not say that the first domino pushes the third domino because it is obviously not true. The required relation between event 'A' and event 'C' does not exist. Therefore causation is not transitive.

It could be objected that substituting 'cause' with 'push' is illegitimate and that the change results in a strawman argument. But remember that we do not know what a 'cause' is. We do know what a push is, and when we want to bring about the change in the position of an object we often push against it. A push is definitely a change maker and the whole purpose of the concept of a cause is to explain change. These considerations indicate that a 'push' is close to being what a 'cause' is in the world.

Now if my uniformity argument (at the end of the previous section) is sound, it also applies to the example of the boulder that threatened the life of the hiker. The correct analysis of this case also has to result in being able to show that the causal relation is not transitive:-

Light emitted from the Sun (the first event in this causal chain) interacts with the boulder some of the light is absorbed and some is reflected. Very soon after the light hits the back of the hiker's retina… If causation is transitive then the light from the Sun causes the hiker to duck (last event in the chain). But light from the Sun causes nobody to duck – this does not happen; if it did everybody would be continually ducking. Assuming (for this argument) that the causal relation is transitive leads to an absurd conclusion; therefore the causal relation is not transitive. This is exactly the result that is needed.

A does not cause C.

A does not cause C.

References:

Helen Beebee, Christopher Hitchcock and Peter Menzies (editors), (2009) The Oxford Handbook of Causation. Oxford University Press

Lewis, David (1973) 'Causation'. In Oxford Readings in Philosophy: Causation edited by Ernest Sosa and Michael Tooley. Oxford University Press

Schaffer, Jonathan, 'The Metaphysics of Causation', *The Stanford Encyclopedia of Philosophy* (Fall 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2016/entries/causation-metaphysics/>.