Linking Cause and Effect

The role of mechanisms in causation

Bernd Lindemann 08.03.2016 phblin@uks.eu

1. World model

In classical physics the world is modelled as a large set of real physical objects. These are ordered in space and time, separate and of independent existence.[1] In finite-state models the objects have properties characterizing one of several states. On transient encounter of two objects these may interact, changing their properties, i.e. assuming different states. Not abstracta but only real objects, which can physically interact, have causal power.[2]

While the classical world model was set aside by quantum physics, it remains close to our experience and acceptable to our common sense. The classical model is still the playground of philosophical authors and of astronomers, chemists, biologists, engineers etc. concerned with causation.[3] The classical view is also the starting point of the following treaties.

2. Causation

Causality is the relation between cause and effect. With it we all have extensive, if sometimes doubtful experience. If the second of two subsequent events occurs reliably after the first, we tend to feel that the first is its direct cause. Is this attribution of cause and effect justified? David Hume famously concluded that generally the cause-effect explanation is based on insufficient data.[4] It seems a habit of the mind, is interpretation rather than a fact of the world. Natural laws (like causality) are merely regularities interpreted.

Philosophical physicalism,[5] to name a mainstream school of thought, accepts this view in part, but goes beyond. Causation may be a first hypothesis, a tentative or a rushed causal explanation of an experienced regularity. Yet in many cases investigation proves the hypothetical explanation to be justified. Then physical cause may predictably result in physical effect *because* of a linking mechanism allowing predictable interaction, which closer investigation will reveal. If it is physical and reliable, it may be a mechanism. This expectation is supported by the general experience that any

1 e.g. Einstein, A., *Quanten-Mechanik und Wirklichkeit.* Dialectica, 1948. **2**: p. 320-324. Note that *inseparable* entities of dependent existence (like a constituted whole being inseparable from its parts) cannot be causally related. There is no *bottom-up* or *top-down* causation between system levels since an entity cannot transiently encounter and interact with what it contains.

2 Alexander's Dictum. Alexander, S., *Space, Time, and Deity.* 1927, London: Macmillan.

3 e.g. Andersen, P.B., et al., eds. *Downward Causation. Mind, Bodies and Matter.* 2000, Aarhus University Press: Aarhus. 354 pages., Woodward, J., *Making Things Happen. A theory of causal explanation*. 2003, Oxford: Oxford University Press. 412 pages., Pearl, J., *Causality: Models, Reasoning and Inference*. 2 ed. 2009, New York: Cambridge University Press. 464 pages., Kim, J., *Physicalism, or something near enough*. 2005, Princeton: Princeton University Press. 186 pages., Craver, C.F., *Explaining the brain. Mechanisms and the mosaic unity of neuroscience* 2007, New York: Oxford University Press. 308 pages.

4 Hume, D., *An enquiry concerning human understanding*. 1741, Oxford.

5 Ontologically everything is physical. Everything supervenes over the physical. All physical changes can be explained physically. (Introduced by Rudolf Carnap, 1932)

physical change has a physical cause.[6] Then, if all explanations are physical, so will be the link of cause and effect. Let us probe the consequences.

In detail, the presumable mechanism linking cause and effect may often be modelled as a 2-step process: First the physical event of encounter of two objects, then their interaction, causing a change of their properties, denoted as a change of state.

Below, the monocausal relation **C** stands for such a process of encounter and interaction of the unspecified objects x and y. The change of state by interaction is denoted by a prime (Figure 1). The fall of a domino tile causing the fall of its nearest neighbour may serve as an example:[7]

Encounter:	x', y **E** (x' y)	The falling x' topples towards a standing neighbour y.
Interaction:	(x' y) **I** x'', y'	Impact of falling x' on y, which falls too.
amounts to	x', y **C** x'', y'	Fall of x **causes** fall of y.

With '⟾' meaning ' is the direct cause of ', an alternative notation is:

x', y ⟾ (x' y) **I** ⟾ x'', y'

i.e. properties of x and y cause encounter, then interaction of y with x' causes y to change to y' and x' to x''. Thus the causal relation **C** is substituted by two causal relations, **E** and **I**. Generally the mechanistic interpretation moves the multiple causal steps of the mechanism into focus.

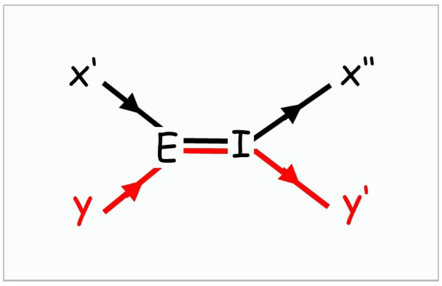

Figure 1. State-transition-diagram (STD) of causation. Direct-cause events (beginning encounter) and effect events (change of state) are linked by an encounter-interaction mechanism.

The brief notation based on **C** is

x', y ⟾ x'', y'.

6 Action on matter, also in causation, is interaction with other matter, a consequence of Newton's third law.
7 a **R** b consists of relata a and b (objects) and the relator **R** (a predicate). With **R** = "is larger than" follows a > b.
a' **R** a'' means that **R** relates (links) 2 states of an object, here a' (a in state prime) to a'' (a in state double-prime).
a', b **R** a, b' means that **R** relates the pair a' and b to the pair a and b'.
a', b **C** a, b' means that the pairing (event) of a', b is the direct cause of a, b', or a', b ⟾ a, b'.

However, given knowledge about the mechanism this is a truncation which hides the 2 mechanistic steps, making '⟾' the explanandum. Even shorter is the common phrasing: "The change in x ⟾ the change in y", which further hides the equally relevant "The change in y ⟾ the change in x".

Finally, the causal relation can be expressed in terms of conventional mathematics. The stationary transfer-function of an idealized model mechanism may be:

$$y = f(x, a, b...).$$

If the function is continuous differentiable, it determines where a change in x will reliably result in a change in y. Then dx may be said to be the **cause** of dy, which responds with a delay due to the time requirement of the mechanism.[8]

3. Causal chains

A causal chain is a sequence of direct cause-effect occurrences, drawing a cyclic path in the state-transition-diagram (STD). With objects a, b, c the chain reads (Figure 2):

a', b ⟾ (a' b) **I** ⟾ a, b' followed by b', c ⟾ (b' c) **I** ⟾ b, c'

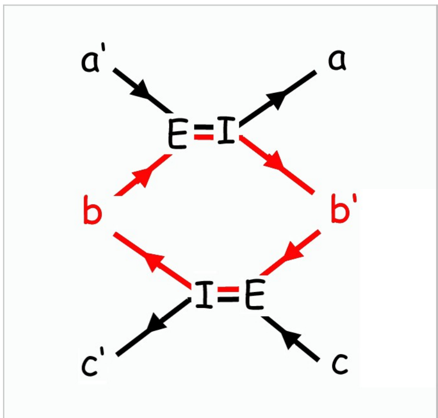

Figure 2. STD of a monocausal chain.

A common phrasing is: "The change in a caused the change in b, which caused the change in c." Then the statement "The change in a caused the change in c" is also true.

It is a general finding that all causes are parts of causal chains, no effect is without cause, any physical change has a physical cause, there is no 'Erstverursachung'. Thus true chance-events, i.e. events without cause, are not possible.[9]

8 However, some authors (e.g. Ernst Mach, Max Verworn) have pointed out that there is no mathematical necessity nor an advantage in applying a cause ⟾ effect scheme to the function.

9 Schnakenberg, J., *Öffnet der Zufall unsere Welt?*, in *Symposium Zufall*. 2015: Bernkastel-Kues, 28. bis 31. August 2014. p. 1-12.

4. Probabilistic choice

Suppose object a may encounter one of two neighbours, f or g, with transition-probabilities P_f and P_g. As long as the two probabilities are < 1, each selection of f or g will be unpredictable. With $P_f = P_g$ the path in the STD will be an unbiased random walk.

Such a choice among 2 possible causal chains may be denoted as:

$$a', f, g \quad \begin{matrix} P_f \Longmapsto (a'\ f)\ \mathbf{I} \Longmapsto a, f' \\ P_g \Longmapsto (a'\ g)\ \mathbf{I} \Longmapsto a, g' \end{matrix}$$

i.e. "The change in a ⟾ the change in f **or** in g according to P."

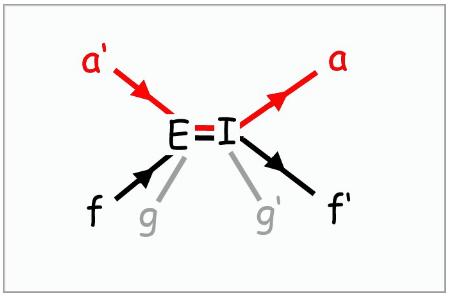

Figure 3. Encounter chooses f among 2 alternatives, f and g.

The causal path is shown in Figure 3. Random choice goes with less transition-probability, i.e. less overall reliability (predictability). The more demanding case of one choice among 4 possible targets is considered below (Figures 4, 5).

Of the two relations composing **C**, it is the encounter relation **E**, of course, to which varying probabilities are attached, depending on the number of competing encounter-targets and other issues. These transition-probabilities of **E** are not due to objective chance, but are resolvable into mechanisms on closer scrutiny. Another probability, attached to **I**, may be high or unity, as it would be in the domino example above. Of **E** and **I**, the latter is the causation-specific step: If causation occurs at all, it will occur according to **I,** while **E** merely varies the predictability of the causation. Therefore, though it sounds paradox, causation may be unpredictable (subjectively probabilistic) and random choice may be followed by causation resolvable into mechanisms. There is a continuous change-over through increasingly biased choices to predictability.

5. Quasi-random paths

In this sense random walk exemplifies unpredictable causation. Still, it remains causation rather than chaos, for with **E** the branching rules of the STD are observed (Figure 4) while only what is specified in **I** can happen. Predictability improves when one of the encounters with neighbours is statistically favoured. If the bias is slight, then *prima facie* the path in the STD may still appear random, only long observation reveals a pattern. If the bias is stronger, the pattern becomes more distinct. Thus deviation from randomness may be observed by comparing the transition probabilities for a long time.

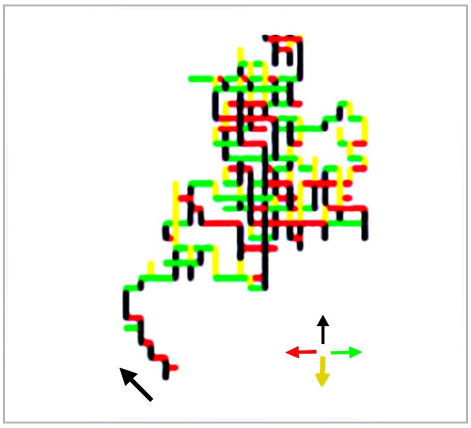

Figure 4. STD showing an unbiased random walk. The causal chains contained one choice among 4 at every stage, with $P_{up} = P_{right} = P_{down} = P_{left} = 0{,}25$. Colour coding was upward = black, to right = green, downward = yellow, to left = red (inset).

Encounter with one out of 4 competing neighbours (rather than with one out of 2 neighbours) at every stage is shown in Figures 4 and 5. For each stage **E** specifies the local transition-probabilities of encounter, expected to be resolvable mechanistically, and **I** the mode of interaction, here indicated by changing state and moving to the next stage.

Figure 5 illustrates the gliding change from random walk to biased walk to predictable mechanism. In locations where transition-probabilities were equal, an unbiased random walk is seen. Where one probability was slightly favoured, the STD begins to show the rectangular pattern of an underlying mechanism, still punctured by frequent escapes into areas of unbiased randomness. The rectangular cycle consisted of 80 steps to the right (green), followed by 40 steps downward (yellow), etc. In locations where one probability dominated, the rectangular pattern is more distinct, escapes into areas supporting unbiased random walk become seldom. Where one probability approached unity, the cyclic STD of a fully predictable mechanism appeared.

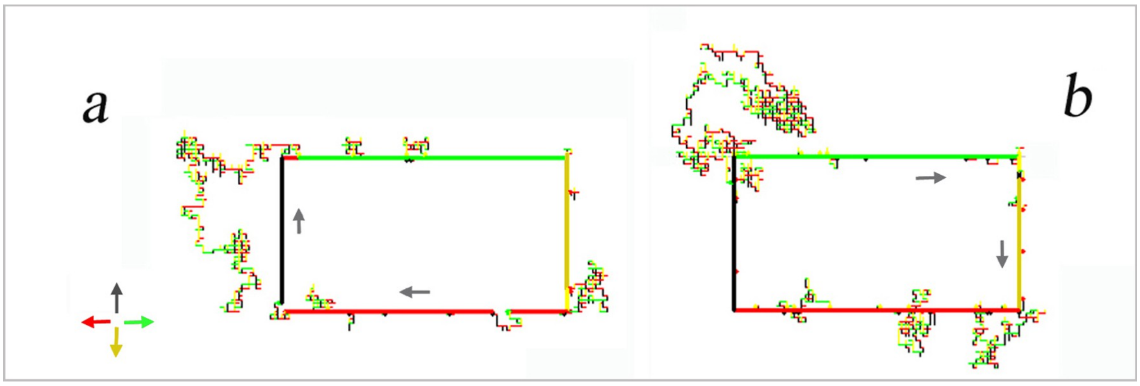

Figure 5. State-transition-diagrams (STD) illustrating the gliding change from random walk to increasingly predictable mechanism. ***a, b:*** Transition probabilities varied with location. Those of areas surrounding the mechanism were 0,25. The probabilities of the mechanism (rectangular pattern) are given in Table 1. The horizontal paths had 80, the whole mechanism 240 steps of encounter and interaction.

Transition Probabilities					
	surrounding	rectang. pattern up: black	rectang. pattern to right: green	rectang. pattern down: yellow	rectang. pattern to left: red
P_{up}	0,25	0,7	0,01	0,28	0,01
P_{right}	0,25	0,01	0,7	0,01	0,28
P_{down}	0,25	0,28	0,01	0,7	0,01
P_{left}	0,25	0,01	0,28	0,01	0,7

Table 1. Transition probabilities of Figure 5.

6. Predictability of mechanisms

Mechanisms were optimized for predictable performance by biological evolution or technical design. In simple cases they follow just one particular causal chain.[10] A mechanism without predictable performance would be useless.

Per example, let us consider the STD of a cyclic causal chain of a mechanism performing in the steady state.[11] For simplicity encounters involve only one neighbour, the path is unbranching. Using again the notation which separates encounters from interactions, we find:

$$a', b \Longmapsto (a'\, b)\, \mathbf{I} \Longmapsto a, b'$$
$$b', c \Longmapsto (b'\, c)\, \mathbf{I} \Longmapsto b, c'$$
$$c', d \Longmapsto (c'\, d)\, \mathbf{I} \Longmapsto c, d'$$
$$d', a \Longmapsto (d'\, a)\, \mathbf{I} \Longmapsto d, a'$$

Thus the cyclic overall mechanism is constituted of several steps of encounter and interaction which are arranged in sub-cycles, as shown in Figure 6. Reliability is ideal since encounter-probabilities are unity. Note that this causal chain cannot exist in isolation. It will be linked to external events, for instance by multicausal merging with an encounter-activating path (green arrow on top).

10 like the causal chain in an old-fashioned clock. A more complex example with a population of parallel causal chains is an evolved neuronal feedback circuit regulating blood pressure. Such mechanisms are multicausal but sufficiently reliable to serve their function. See Lindemann, B., *Mechanisms in World and Mind.*. 2014, Exeter, UK: imprint academic. 152 pages. Lindemann, B., *A whole affects its parts?* 2015, Homburg, Germany: invoco-verlag. 64 pages.

11 The chain has to close a cycle to allow steady state action, see: Hill, T.L., *Free Energy Transduction in Biology. The steady-state kinetic and thermodynamic formalism*. 1977, New York: Academic Press. 229 pages.

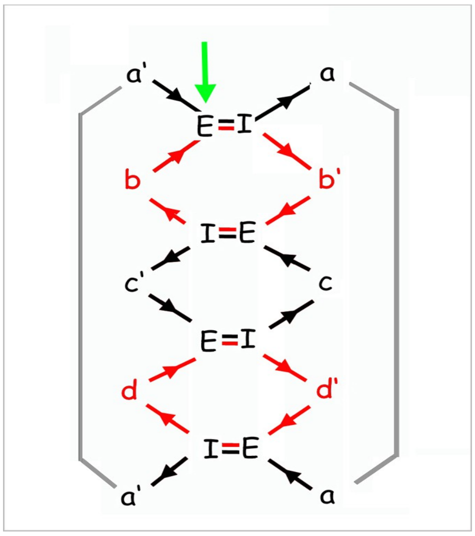

Figure 6. STD of a cyclic causal chain capable of steady state performance. The overall mechanism is a loop consisting of four cyclic **E-I** submechanisms. The triggering input (green arrow) shares its target with a', thus participating in multicausal action on one encounter element **E**.

7. Supporting causes

While monocausality may be claimed for causal chains of many designed mechanisms, causation is generally not mono- but multicausal. It assumes the structure of causal networks rather than chains. For instance, the careful arranging of a set of upright domino tiles certainly is essential, though unmentioned in the “direct cause $\Longmapsto$ effect” formalism. Without such supporting processes the 'monocausal' domino events, triggered by direct causes, were not possible.[12] Unfortunately the world's multicausal network is so complex that it cannot be represented, of course, in full detail. Using a global conditional instead, all of the above relations should be preceded by something like

> “If supporting processes were completed and... “

While supporting events must happen prior to the causation in question, their exact time is often not critical. In contrast, direct causes must happen just before the effect appears, cause and effect are separated by the fixed timespan needed by the causal mechanism. Since supporting events are always required, monocausation is really fictive, an idealisation.

8. Non-physical causes

Above we reasoned: If all explanations are physical, so will be the link of cause with effect. But what about non-physical causes, how are they linked to their effect? Take for example an error as a non-physical item:

> An error in the software caused the rocket to explode.

This seemingly harmless statement is misleading because an error (a fault in a script) is abstract (not

12 Fred Dretske calls them 'structuring' and 'triggering' causes (e.g. Dretske, F., *Mental events as structuring causes of behavior*, in *Mental Causation*, J. Heil and A. Mele, Editors. 1995, Clarendon Press: Oxford. p. 121-136.).

ordered in space and time). An error cannot physically interact, cannot cause. Rather, it was a spark in the fuel line which caused the rocket to explode.

Generally non-physical objects are mental objects. Most of them, like thoughts, do not exist independently but are “about”.[13] Being abstracta, they cannot interact in space and time, have no causal power. Thus non-physical *causes* are not linked to effects because such causes do not exist.

9. Conclusion

Causality carries the expectation of something reliable linking two subsequent events. I suggest that the reliable link is a physical mechanism. Yet causation commonly comes with a truncated formalism “direct cause $\Longmapsto$ effect” which tends to hide the possibility of reduction to physical interaction. Aware of this, we may explain the causal relation **C** as the linkage of two subsequent events by an (at first hypothetical) particular mechanism of interaction.

In detail, the linking relation may itself be composite, for instance composed of the encounter-relation **E** specifying the probabilistic branching-rules of the STD and the interaction-relation **I**, which specifies the characteristics of the mechanism. There is a gliding change-over from probabilistic causation to predictable causation; from unbiased choice through increasingly biased choice to reliability. The abundance of successful mechanistic explanations suggests that reliable mechanisms, optimized by evolution or by design, exist in the world. This justifies the concept of reliable causation, with predictable mechanisms linking direct causes with their effects.

10. Outlook

At the same time, the reductive explanation of causation as a mechanistic linkage of two events raises further questions. They address explanations at lower levels of the system hierarchy,[14] like quantum physics, where the concept of 'causality as linkage', like other concepts of classical physics and even the concept of objects composed of parts, must be abandoned. Thus reductive physicalism explains causality in terms of physical relations of objects, but goes beyond towards the ultimate level. Here objects are neither composite nor distinct any more, and concepts like causation and interaction break down.

---

13 They have Brentano-intentionality: Brentano, F., *Psychologie vom empirischen Standpunkt*. 1874, Leipzig: Meiner Verlag 1924.

14 See 'downward drainage' in e.g. Block, N., *Do causal powers drain away?* Philosophy and Phenomenological Research, 2003. **67**: p. 133-150.