Linking Cause and Effect

The role of mechanisms in causation

Bernd Lindemann 10.08.2016 phblin@uks.eu

Summary: Causality carries the expectation of something predictable linking two subsequent events. I suggest that this reliable link is always a multi-step physical mechanism. Yet causation commonly comes with a truncated formalism "direct cause $\Longrightarrow$ effect", where any multi-step causal chain of physical interactions is hidden by the arrow. Aware of this, we may explain a causal relation with the linkage of two subsequent events by a particular mechanism of interaction. This is shown with examples of increasing complexity.

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. Only real objects, which can physically encounter and interact, have causal power. Abstracta, which cannot encounter in space and time, cannot cause.[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.

---

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 Sometimes attributed to: Alexander, S., *Space, Time, and Deity.* 1927, London: Macmillan (Alexander's Dictum).

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.

2. Causation

Causality or causation is the relation between two subsequent events, cause and effect. With it we all have extensive, if sometimes doubtful experience. If the second of two events occurs reliably after the first, we tend to feel that the first is its direct cause. Of course this naïve attribution of cause and effect needs justification. David Hume famously concluded that generally the cause-effect explanation is not based on sufficient 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 a mere experienced regularity. Yet in cases investigation will support the hypothetical explanation. Then physical cause may predictably result in physical effect due to a reliable physical link between them, as closer investigation will reveal.

For instance, it will be regularly observed that night follows day and day follows night, like effect follows cause. Few exceptions are on record. Ancient humans have explained the regularity with animistic or poetic fantasies. Today we prefer planetary rotation as a physical mechanism linking day and night.

The expectation of a physical link is supported by the general experience that every physical change has a physical cause.[6] Then, if all explanations are physical, so will be the link of cause and effect. The question remains, however, *which* causal chain of physical linkage is realized in a particular case.

To quote another example: Raining (the cause) makes the street get wet (the effect). The linking mechanism consists of 1. **encounter** of rain droplets with the surface of the street and 2. **interactive** spreading of the water on the surface of the street. Like in this case, the link between direct cause and effect may often be modelled as a 2-step process.

The fall of a domino tile causing the fall of its nearest neighbour is another well-known example. First we observe the physical event of encounter of two objects, then their interaction, causing a change of their properties, denoted as a change of state. The monocausal and deterministic 2-step model used here certainly is an idealisation of a more complex reality.

Below relation **C** stands for the process of encounter and interaction of the unspecified objects x and y. The change of state resulting from interaction is denoted by a prime (Figure 1). [7]

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

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

5 meaning that everything is physical or supervenes over the physical. Thus all physical changes can be explained physically. Introduced by Rudolf Carnap, 1932. Here treated as a hypothesis based on observation.

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 $\Longmapsto$ a, b'.

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

x', y ⟾ (x' y) ⟾ x'', y'

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

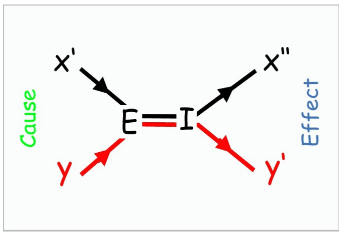

Figure 1. State-transition-diagram (STD) of causation. Direct-cause events (beginning encounter) and effect events (change of state) are linked by an encounter-interaction mechanism. A one-shot process without regeneration of x' and y is shown.

A briefer notation is x', y ⟾ x'', y'. However, given knowledge about the mechanism, this is a truncation which hides the 2 mechanistic steps, making '⟾' an 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 mathematical notation. Let y = y(x) be the stationary transfer-function of an idealized model mechanism. If the function is continuous differentiable, it determines which change in y will reliably result from a change in x. 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 path in the state-transition-diagram (STD). With three objects initially in states a', b, c the chain reads (Figure 2):

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

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 " (transitivity) is also true.

It is a general finding that all causes are parts of causal chains embedded in causal networks, no effect is without cause, any physical change has a physical cause (physicalism), there is no 'Erstverursachung'.[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

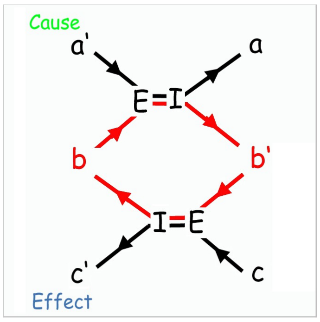

Figure 2. STD of a monocausal chain with cyclic regeneration of b. A one-shot process without regeneration of a' and c is shown.

4. Probabilistic choice

Suppose an object in state a' encounters an object in state f (transition probability P_f) or another object in state g, with transition-probability $P_g = 1 - P_f$. As long as the two probabilities are < 1, each individual selection of f or g will be unpredictable. With $P_f = P_g$ the path in the STD will be an unbiased random walk. Note that any probabilistic step in this model may stand short for a complex deterministic process not yet understood.[10]

The probabilistic choice among 2 possible causal chains may be denoted as:

with P_f : a', f $\Longmapsto$ (a' f) $\Longmapsto$ a, f'
with $1 - P_f$: a', g $\Longmapsto$ (a' g) $\Longmapsto$ a, g'

i.e. "The change in a $\Longmapsto$ the change in f **or** in g according to P_f, P_g." with $P_g = 1 - P_f$.

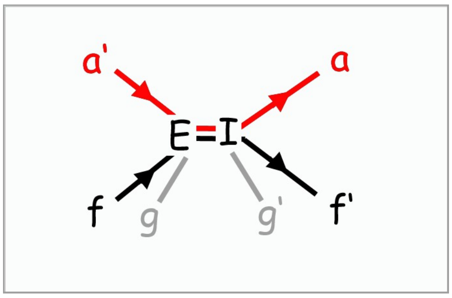

Figure 3. STD. The encounter chooses (randomly or biased) among 2 alternatives, f and g.

The causal path of the STD is shown in Figure 3. 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** 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

2014. p. 1-12.

10 e.g. see Chapter 2 in Lindemann, B., *Mechanisms in World and Mind.*. 2014, Exeter, UK: imprint academic.

occur according to **I,** while **E** merely varies the predictability of the causation. There is a continuous change-over through increasingly biased choices to predictability.

5. Quasi-random paths

In this sense a random walk exemplifies unpredictable causation. Still, it remains causation rather than chaos, for with **E** the branching rules of the STD are observed while only what is specified in **I** can happen (Figure 4). Predictability improves when one of the encounters with neighbours is statistically favoured. If the bias is slight, then *prima facie* the path 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.

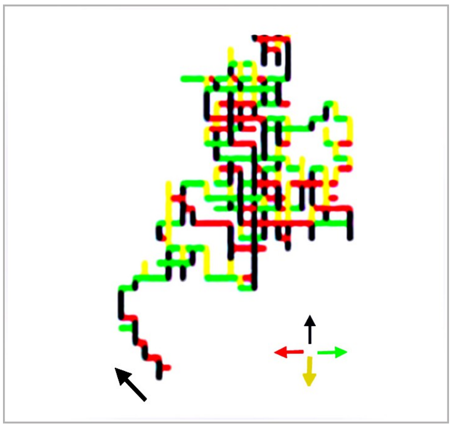

Figure 4. Unbiased random walk through a planar STD. At each location the causal chain contained one random choice among 4 neighbours, with $P_{up} = P_{right} = P_{down} = P_{left} = 0{,}25$. Colour coding was upward = black, to right = green, downward = yellow, to left = red (inset).

Encounter with one out of 4 competing neighbours at every stage of the path 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 of the STD.

Figure 5 and Table 1 illustrate the gliding change-over from random walk to biased walk to predictable mechanism, based on multiple steps of encounter and interaction. The rectangular cycle shown is the path of an ideal mechanism. It consisted of 80 **E** and **I** steps to the right (green), followed by 40 steps downward (yellow), etc. Cause and effect may be located in the lower left and upper right corner. In locations surrounding the rectangular path transition probabilities were equal, resulting in unbiased random walks. Where one probability was favoured, the rectangular pattern of an underlying mechanism began to show, still punctured by frequent escapes into areas of unbiased randomness. In locations where one probability dominated (P = 0,7), escapes into surrounding areas of unbiased random walk became more seldom and the cyclic path of a reasonably predictable mechanism appeared as the rectangular pattern.

Random Walk	Transition Probabilities P	Reliability	Figure	
unbiased	equal	none	4	
biased	unequal	poor to reasonable	5	
ideal mechanism	one P = 1	ideal		

Table 1. Gliding change from random walk to ideal mechanism.

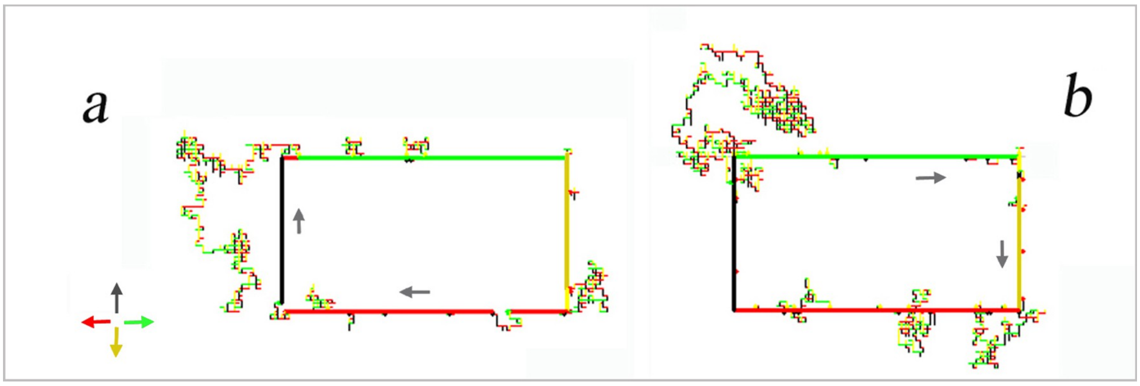

Figure 5. ***a, b:*** Two walks through a planar STD with one choice among 4 neighbours. Transition probabilities varied with location. Those of areas surrounding the mechanism were 0,25. The probabilities of the mechanism (rectangular pattern) are given in Table 2. The horizontal paths had 80, the whole 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 2. Transition probabilities of Figure 5.

6. Predictability of mechanisms

A mechanism is a physical device which reliably (or quasi-reliably [11]) causes a specific change in its environment. For instance, when its trigger is pulled, a gun fires and the bullet travels along a chosen trajectory through the environment. Physical mechanisms like planetary rotation have an inherent stability which makes them reliable. Biological and technical mechanisms were optimized for predictable performance by evolution or design. In simple cases they follow just one particular causal chain or causal cycle.[12] A mechanism always has reasonably predictable performance, even though its probability is never 1.

Per example, let us consider the STD of a cyclic causal chain of a mechanism performing in the steady state.[13] For simplicity encounters involve only one neighbour, the path is unbranching (but see below). Using again the notation which separates encounters (first arrow) from interactions (second arrow), we find:

a', b ⟾ (a' b) ⟹ a, b'
b', c ⟾ (b' c) ⟹ b, c'
c', d ⟾ (c' d) ⟹ c, d'
d', a ⟾ (d' a) ⟹ d, a'

Thus the cyclic overall mechanism is constituted of several steps of encounter and interaction. They are arranged in four sub-cycles, which serve for regeneration of states, as shown in Figure 6. Reliability approaches the ideal since probabilities of encounter are nearly unity.

Note that this causal chain, model of a mechanism, cannot perform in isolation. It will be linked to external events, for instance by multicausal merging with an encounter-activating path (green arrow on top). A change in the environment is achieved by a divergence of paths, altering the state of an environmental object (blue arrow).[14]

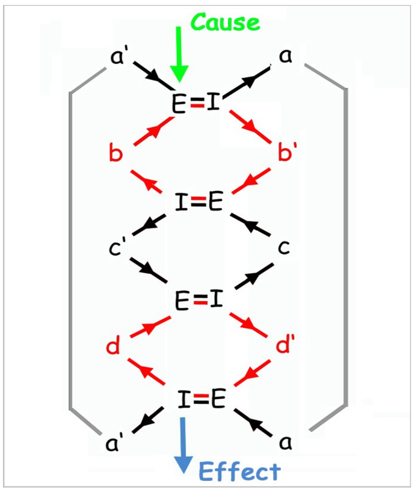

Figure 6. STD of a cyclic causal chain linking environmental cause with environmental effect. The overall mechanism, capable of steady state performance, is a loop consisting of four cyclic **E-I** sub-mechanisms. The triggering input (green arrow, 'Cause') shares its target with a', thus participating in multicausal action on one encounter event. The blue arrow stands for the caused change in the environment, the 'Effect'.

11 With sufficient though not perfect reliability.
12 Like the causal chain in a mechanical 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.
13 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.
14 Further, a difference in free energy between cause and effect may assure that the cycle turns in the indicated direction only. e.g. T.L. Hill, 1977, l.c.

7. Supporting causes

While monocausality may be claimed for causal chains of many evolved or designed mechanisms, causation is generally not mono- but multicausal. It assumes the structure of causal networks rather than chains. For instance, “Smoking increases the risk of cancer” cannot be covered with a mono-causal model. Further, while the careful *arranging* of a set of upright domino tiles certainly is essential for cause and effect, it remains unmentioned in the “direct cause ⟾ effect” formalism. Without such supporting processes (or regeneration steps, not shown in Figure 1) the 'monocausal' domino events, triggered by direct causes, were not possible.[15]

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. Yet, since supporting events are always required, monocausation is really fictive, an idealisation.

8. Non-physical causes

Above we argued that if all explanations are physical (the hypothesis of physicalism), so will be the link of cause with effect. But what about non-physical causes, how are they linked to their physical 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 quite misleading because an error (a fault in a script) is abstract (not 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, like thoughts. They do not exist independently but supervene over neuronal processes. Further, they are “about”.[16] Being abstracta, they cannot encounter and interact in space and time, they have no causal power. Thus non-physical *causes* are not linked to physical effects because such causes do not exist.

---

15 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.).

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

9. Conclusion

Causality carries the expectation of something reliable linking two subsequent events. I suggest that the reliable link is always a multi-step physical mechanism, often a complex one. Indeed, if transition probabilities always remain below unity, then all causation is probabilistic. Further, if transition probabilities may eventually be resolved into deterministic and complex causal networks, then all causation, even that approximately modelled as a monocausal chain, is really complex.

Yet causation commonly comes with a truncated formalism “direct cause ⟾ effect”, hiding any multi-step physical interaction in the arrow. Aware of this, we may explain a causal relation **C** as the linkage of two subsequent events by a particular and possibly complex chain of interactions.

In detail, the linking relation may itself be composite, in simple cases composed of the encounter-relation **E** specifying the probabilistic branching-rules of the state-transition diagram and the interaction-relation **I**, which specifies the characteristics of the mechanism.

Encounter may be probabilistic and there is a gliding change-over from probabilistic performance to deterministic, predictable causation; from unbiased random choice through increasingly biased choice to reliability. The abundance of successful mechanistic explanations suggests that reliable mechanisms exist in the world. This justifies the concept of reliable (or quasi-reliable) causation, with multi-step 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,[17] 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. Yet the classical view remains valuable as an intermediate stage of reductive explanation of intuitive plausibility.

---

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