Changes in state *vs* persistent state in continuous time

Aurélien Fermo, ENS - EHESS

January 28, 2020

Last meeting we talked about the difference between persistence of state *versus* changes of state, and looked at a graph similar to that shown in Fig.1. Squares represent nodes which maintain their value through time after being stimulated, whereas circles represent nodes which don't maintain their value. We're considering a continuous time and we posit something like a *no coincidence principle* which means that two nodes, whatever they are, cannot be activated in the very same time. In Fig.1 let's say that this is an AND-Gate (depicted by arc of circle): node Z needs to be activated by two nodes (J and K).

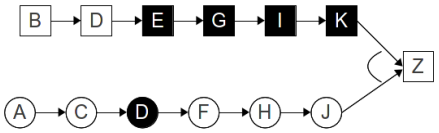

Figure 1

The idea was to evaluate whether and how this distinction challenges both the physical process (PP) and the counterfactual (CF) accounts of actual causation. But I think that we have to carefully distinguish two different − but related − types of question : on the one hand, independently of the distinction between PP and CF accounts of causation, whether the type of nodes (persistent or not) influence our causal judgements; on the other hand whether and how the type of nodes can bias our causal judgements toward one account rather the other.

Indeed, the question of the type of nodes put aside, the PP and CF accounts (in their simple and common form) actually agree on the actual causes in Fig.1 : E and D are the actual causes of Z. To see why both PP and CF accounts agree in that case, let's follow and simplify a bit an analyse of actual causation found in [1]. Let's say that the effect Z occurs at time t' and that t_0 is the time at which all the nodes are set to their initial values (so $t_0 < t'$). Moreover we define a *minimally sufficient set* at any time t as the set of all and only the nodes that are activated at t and whose activation is necessary for Z being activated. The *causation by production* (i.e. PP) and the *causation by dependence* (i.e. CF) obey different logics and, as it were, two opposite orders in identifying the actual causes of Z:

- Causation by production works by "tracing back" the causes of Z from t' to t_0 through all the intermediate causes that are connected by spatio-temporal relationship according to an *Intrinsic* structure. This structure is such that it had been the only one involved in the causal history of Z (all the other nodes *extrinsic* to that structure

being ruled out), Z would have occurred as well[1]. So at $t'-1$ a minimally sufficient set $S_{t'-1}$ of producers of Z is found and for each of its elements x is found at $t'-2$ the minimally sufficient set $S^x_{t'-2}$ of the producers of x, and so forth up to t_0. At the end the actual causes are the set S_0 which isn't but the union of all the minimally sufficient sets at t_0+1.

- Causation by dependence works in the opposite direction by identifying directly at t_0, from the outset so to speak, the minimally sufficient set S_0. Thus this set is not inferred by a kind of backward induction like above, but from the *nomological* relationship between all the intermediate causes from t_0 to t'.

Finally, when both the set S_0 obtained by tracing back the causal history of Z, and that S_0 obtained by nomological identification are identical, we say that causation by production and causation by dependence agree on the actual causes of Z. Following this analysis, we can easily show that in cases like that in Fig.1 both theories of causation give the same response about the actual causes of Z.

To the extent that both theories give the same predictions, if the type of nodes makes a difference in inferring causal judgements (saying for instance that both D and E, but D *more than* E, cause Z, or that E is a cause and J *more than* D the other one) then it will challenge both accounts in the same way. For instance, if having a sense of graded causality (e.g. J *more than D* is a cause of Z) is due (by supposition) to the non-persistence property of the nodes, then we have to say why, regardless of both theories, something like an *attenuation effect* is found in chain of non-persistent nodes. Or maybe we have to see how the specific frame of each theory can handle this effect, but in any case this cannot be used to decide *a priori* between them.

So if we want to look at the effect of the type of nodes on causal judgements regardless of the predictions of both the PP and CF accounts (when they are the same), we should then play around with the type of chain and the initial values (Fig.2&3).

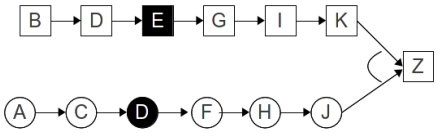

Figure 2

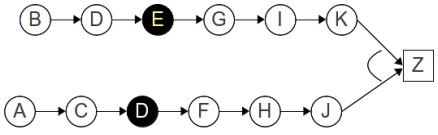

Figure 3

Now, if we want to more specifically study whether and how the (non-)persistence of states influences causal judgements toward one account of causation rather than the other, we

[1]Hall prevents us mistaking this necessary counterfactual analysis of a production process for the counterfactual theory of causation below.

need to focus on cases where PP and CF accounts make different predictions about the actual causes involved. More precisely we need cases where they disagree because of the structure itself of the graph, not because of type of nodes (circle *vs* square). Accordingly, it seems that we have to consider *pre-emption* cases, and more specifically cases of *double pre-emption* precisely because the predictions of the PP and CF accounts are different in those cases. Hence, given that they make different predictions in case of *(double) pre-emption* we could see whether the type of nodes involved (circle *vs* square) biases causal judgements in accordance with predictions given by one account rather than the other.

In Fig.4 for instance is given a case of *double pre-emption* where the already − here − activated node M prevents node N activating node Q which would have prevented node Z. Let's say that the activation of both chains of circles begins with the activation of B and C respectively. The PP account says that the actual cause of Z is C, whereas the CF one states that the actual causes are C *and* A because A hadn't been activated, Z wouldn't have been activated (and obviously C hadn't occurred, Z wouldn't have occurred).

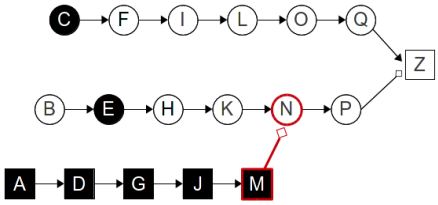

Figure 4

Again, keeping the structure of the graph the same we can study how the type of nodes, together with the initial values (for the squares), affects causal judgements by changing squares into circles and vice versa (for instance Fig.5-7).

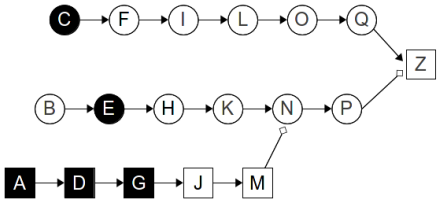

Figure 5

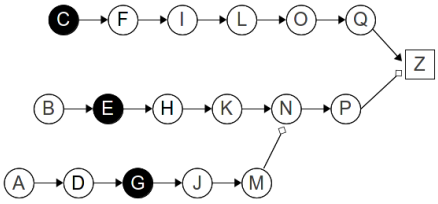

Figure 6

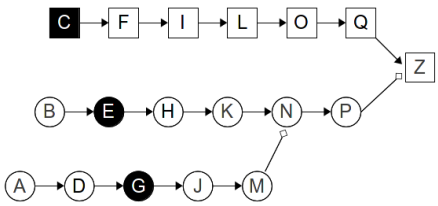

Figure 7

I think that we can follow the same plan when assessing the influence of the type of nodes in non-recursive graphs. On the one hand we should evaluate its effect on causal judgments

independently of the distinction between PP and CF accounts, and on the other hand focus on pre-emption cases to see how causal judgments could be biased toward either the PP or CF accounts in non recursive graph with different types of nodes. For the former part of the plan, Fig.8 for instance would be useful. It is similar to Fig.1 except that the chain of non-persistent nodes is now cyclic.

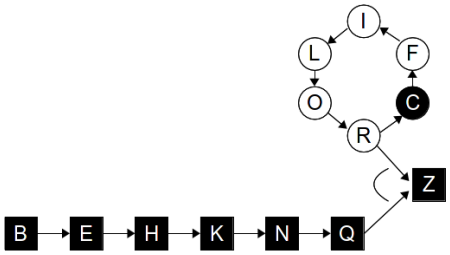

Figure 8

However, even if my intuition is that PP and CF accounts of causation would give the same answer about the actual causes in cases like that (where there is no pre-emption), maybe it is less obvious and they deal differently with cycles. Accordingly, we would need to see how each of the account of causation handles cycles of non-persistent nodes. But I find *a priori* no reason why they would treat them differently since in a cycle like in Fig.8 both the set S_0 obtained by tracing back the causal history of Z and the set S_0 obtained by identifying nomological relationship should be the same. The cyclicity might have an effect on causal judgements since it seems intuitive to think that after few rounds we don't consider as *the* main cause the node which has been activated the first. Maybe it appears after some time a sort of *globalisation effect* such that we do not pick out one node of the cycle as the cause but identify the cycle itself as a global cause. Moreover the memory has probably to do with that for we likely forget, after some rounds, which node of the cycle was activated the first . An other possibility would be choosing the proximate cause R as the main cause of Z − a kind of *proxy effect*. But in any case, if such a *globalisation effect* or *proxy effect* were to appear, I think that they would apply indifferently to both the PP et CF account of causation in those cases.

Then for the later part of the plan − assessing whether and how the type of nodes biases toward one account of causation in non-recursive graphs − we should consider cases of double pre-emption like above. Fig.9&10 show a graph similar to the ones depicted in Fig.4-7, except that the node which prevents N activating the node which prevents Z is included in a cycle. Graphs in Fig.9&10 are exactly the same, but represented at two different times. Here we assume that a preventing node, when activated, *prevents* a node being activated as well as *inhibits* a node already activated. Moreover, for the sake of simplicity, we consider that once a node is deactivated by pre-emption, all its descendants are immediately and simultaneously deactivated as well.

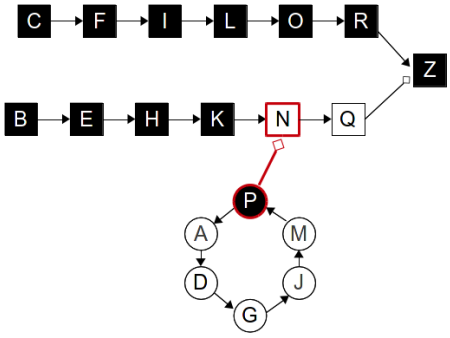

Figure 9

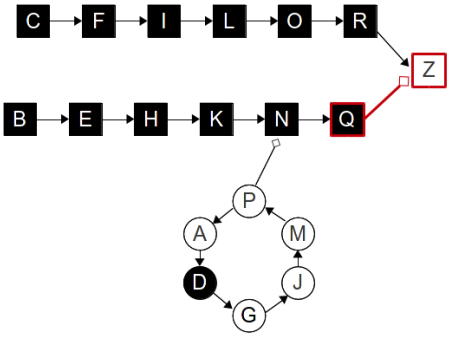

Figure 10

Finally it would be interesting I think to combine the above type of *preventing cycles*, so to speak, with *producing cycles* to see maybe more vividly the difference between PP and CF accounts, and also maybe between these two types of cycles : maybe, for instance, people are more prone to pick out P (thus breaking the cycle) as a − proximate − cause in the *preventing cycle*, while more inclined to consider the *producing cycle* in itself as as global cause.

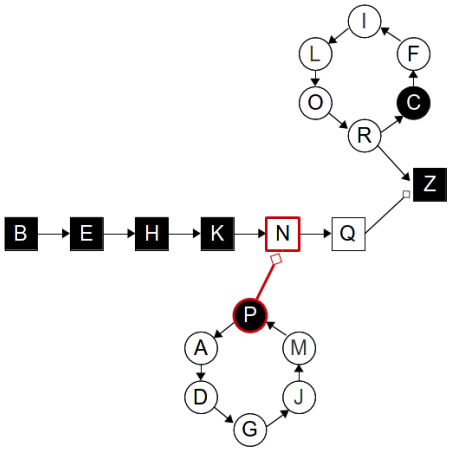

Figure 11

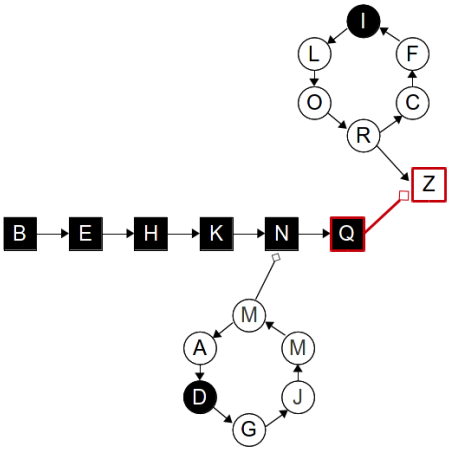

Figure 12

References

[1] Hall N. Two concepts of causation. In N. Hall & L. Paul J. Collins, editor, *Causation and counterfactuals*. MA: MIT Press, Cambridge, 2004.