# Probabilistic Causation

Margherita Benzi
Università del Piemonte Orientale "Amedeo Avogadro"
margherita.benzi@lett.unipmn.it

## 1. Introduction

A well known 1988 survey of probabilistic philosophical theories of causation begun with the remark that «[p]robabilistic theories of causation have received relatively little attention» (Davis 1988, 133). Conversely, a more recent important survey, Williamson (2009, 185), begins in the following way: «Causal relationships are *typically* accompanied by probabilistic dependencies [*italics added*]». As the two different *incipit* show clearly, in about twenty years the perception of the relationship between causes and probabilities has strengthened, and a vast majority of the most promising current philosophical theories of causation deal with probabilities. However, a necessary *caveat* must be introduced: in saying that (almost) all current approaches to causation "deal with probabilities" we are not saying that every philosophical theory of causation must analyze causation in terms of probabilities, or "explain causation away" by means of probabilities: in the last few years, the philosophical program consisting in the *reduction* of causes to probabilities has not been the prevailing one, and, in its original form has ultimately proven unsuccessful (we will return to this point later). Rather, what is currently acknowledged in stressing the relationship linking causes to probabilities is the legitimacy of causal statements also when – due to lack of information or to genuine randomness – causes do not *determine* their effects. The combination of causality with probability should be of no surprise, if we consider that contemporary science is highly probabilistic; but of course, allowing indeterminacy within causation raises the question of *where* the indeterminacy lies (should we place it at the ontic or

Margherita Benzi, "Probabilistic Causation", in Stefano Bordoni & Sara Matera,
*Forecasting the Future. Scientific, Philosophical, and Historical Perspectives*, pp. 31-55

University of Urbino Carlo Bo
http://isonomia.uniurb.it/epistemologica

at the epistemic level?) and what notion of probability should be adopted. A further, but not less important question, is whether there are some theories of probabilistic causation (in this weak sense) that seem to be commonly accepted, i.e. that are accepted independently of the specific analysis of causation – and of probability – endorsed, in the same way in which we say that probability calculus can be accepted regardless of the specific interpretation of probability adopted.

In this paper we will proceed in the following way: firstly we will give a brief reminder of the main tenets of the philosophical program connected to strong probabilistic causation; then we will recall some of the principal problems that the mathematization of probability-based causal theories have tried to solve in the last two decades (with a certain degree of success, which explains the two different *incipit* quoted above). Eventually we will discuss the connection between kinds of probabilities and kinds of probabilistic causations, an issue that has a strong impact on the problem of determinism.

## 2. Causes as Probability Modifiers

According to one of the leading philosophical approaches to causation, the essence of causation is linked to constant regularities. From a historical point of view, the source of the so called *regularity view of causatio*n lies in Hume's famous definition: «We may define a cause to be *an object, followed by another, and where all the objects similar to the first, are followed by objects similar to the second*» (Hume 1748, VII, II). This definition states two key elements of the regularity view, namely temporal succession and constant conjunction. As supporters of regularity view, Hume included, know well, perfectly constant conjunctions are not so frequent. Sometimes conjunctions between types of events which prove to be constant – and are unanimously considered causal – in some contexts, cease to be valid in other contexts, just because some contributing factors are missing, or because some conflicting factors are present. In other cases the "erosion" of constancy does not depend strictly on the context, but on the very fact that the conjunctions in question do not exhibit an inescapable character, but only a certain frequency.

The need of revising the Humean definition in order to account for conjunctions that manifest themselves with not strict, but rather 'gappy' regularities can be seen as a valid motivation for a probabilistic approach to causation: in fact the first step toward a probabilistic account of causation

consists in recognizing the existence of associations which are not invariable, but nevertheless exhibit regular frequencies. A possible paraphrase of the Humean definition could run as follows: «We may define a cause to be an object, followed by another, and where *most* objects similar to the first, are followed by objects similar to the second».

As it is well known, Hume's first definition was accompanied by a second definition whose counterfactual approach seems, at least to contemporary readers, rather at odds with the regularity view: «Or in other words *where, if the first object had not been, the second never had existed.* [It is often remarked that the phrase 'in other words' here is totally misleading]» (*ibid.*). Again, if we wanted to transpose this last definition into a 'gappy' context, the new definition could run as follows: «Or […], if the first object had not been, the probability of the second would have been lower». If we temporarily put aside the philosophical difficulties which can easily be associated to a *counterfactual* lowering of probabilities, it should be natural to think of causes not only as those factors which determine their effects, but also as those factors which raise the probability of their effects. In fact, the first and more basic trait of probabilistic causation consists in seeing positive causes as *probability raisers*; if we define *negative* causes as those preventative or impeding factors which *lower* the probability of their effect, then we can, more generally, see (positive or negative) causes as *probability modifiers*.

The basic idea of probability raising can be stated by saying that, if *C* and *E* are both events[1], *C* raises the probability of *E* when the probability of *E* conditional to *C* is greater than the probability of *E* alone;

$$P(E \mid C) > P(E) \qquad (2.1)$$

or, alternatively, by saying that the probability of *E*, conditional on *C*, is greater than the probability of *E* conditional on not-*C* :

$$P(E \mid C) > P(E \mid \sim C). \qquad (2.2)$$

The two formulations (2.1) and (2.2) are almost equivalent[2]; as (2.1) holds just when (2.2) holds, here we will use the second. Now the basic idea of probabilistic causation can be expressed by the following formula:

[1] In absence of further specifications, we will use the term "events" both for particular events, which are the subjects of particular causal claims, as «Elizabeth's smoking caused her bronchitis», and event types, or *generic events,* which are referred to by general causal claims as «Smoking causes bronchitis».

*C* is a probabilistic cause of E only if P*(E | C)* > P*(E | ~C*). (2.3)

If in (2.3) we substituted the 'only if' with 'if and only if', we would obtain a possible definition of probabilistic causation

*C* is a probabilistic cause of *E* if and only if P(*E* | *C*) > P(*E* | ~*C*). (2.3*)

Given that on the right hand side of the biconditional in (2.3*) we find only probabilities, (2.3*) is a definition which reduces causes to probabilities. However, probability raising, as expressed by (2.3*), is neither a sufficient nor a necessary condition to obtain a satisfying reductive definition of causality, as there are probability raisers which are not causes, and causes that do not raise the probability of their effects. In what follows we will present a short summary of the most relevant problems deriving from insufficiency, namely *Simmetry* and *Spurious causation.*

## 3. Probability Raising is an Insufficient Condition for Defining Causation

### 3.1. Symmetry

A definition of causation in terms of probabilities like (2.3*) does not satisfy the generally accepted requisite that a relation should satisfy in order to be causal, that of being an *asymmetric* relation; in probability theory, if the occurrence of *C* raises the probability of *E*, then also the presence of *E* raises the probability of *C*; in symbols:

If P(*E* | *C*) > P(*E* | ~*C*), then P(*C* | *E*) > P(*C* | ~*E*). (3.1)

Should we define causation by mere probability raising, like in (2.3*), we would be forced to accept that each effect causes its causes, contrary to our intuition that effects are non-causal probability raisers of their causes.[3]

[2] See Hitchcock (2010, 2.1).

[3] It is perfectly possible to think of bidirectional or cyclical causation, but these cases are generally seen as different events of the same kind of the effects causing new events of the same kind of the original causes; for example, the intended meaning of the statement that «poverty causes ignorance and ignorance causes poverty» is that ignorance causes *further* impoverishment.

### 3.2. Spurious correlation

A definition of causation which is based exclusively on *probability raising* does not allow us to distinguish between genuine causation and spurious correlation. A typical problem of spurious correlation occurs whenever two events *A* and *B* are statistically associated, but the link is due to a third factor *C*, which is the cause of both *A* and *B*. Given two events *A* and *B* of which we know only that are statistically correlated we cannot say if

*i*) it is *B* causing *A*, or the other way around, or

*ii*) there is a third causal factor, or complex of factors, causing both *A* and *B*[4], or

*iii*) it is just a mere coincidence, as in the famous case of the parallel increasing of sea levels in Venice and bread prices in London (Sober 2001).

Incidentally, it can be observed that difficulties in distinguishing causal from non-causal associations plague all the regularity theories of causation, probabilistic and not probabilistic ones, since their origins, as attested by Thomas Reid's well known claim that Hume's theory could not rule out the unwanted conclusion that the night causes the day and the day causes the night because they regularly follow each other.[5]

As we will see in a while, the search for suitable solutions for the problems deriving by the so-called "naive probabilistic analysis" of causation (Glynn 2011) inspired a great deal of philosophical work.

## 4. The No-Screening-Off Condition

A large part of what has been built within this research area finds its origins in the contribution given by Hans Reichenbach in *The Direction of Time*, a

[4] A good example of *i*) and *ii*) is given by the controversial causal explanation of the correlation between smoking and depression: «Depression may cause people to smoke (perhaps to self-medicate their symptoms), or smoking may cause increased risk of depression (via alterations in neurotransmitter pathways following chronic exposure). The relationship may even be bidirectional (acute or infrequent tobacco use may reduce negative effect, but chronic use may exacerbate it), or be caused by shared risk factors (possibly genetic) so that the relationship is not causal at all». (Munafò & Araia 2010, 452)

[5] «It follows from [Hume's] definition of a cause, that night is the cause of day, and day the cause of night. For no two things have more constantly followed each other since the beginning of the world». (Reid 1788, 4.9)

volume published three years after his death (Reichenbach 1956). It is worth noting that in a work written in 1923 – but published ten years later - he had presented a Kantian position, in which causal judgements were given a synthetic *a priori* status (Reichenbach 1933). Reichenbach (1956) assumes a different position, claiming that probabilities are more fundamental, and that causality can be derived from probabilistic relations; moreover, he purports to show how the direction of time is derived from the direction of causation. *The Direction of Time* also introduces a concept which will prove very important for the subsequent theories of probabilistic causality, namely the concept of *screening off*. The *no-screening-off condition* formalizes the intuition that spurious causes become uninformative once the real cause is known. Before seeing how this intuition can be put at work, let us introduce two classic examples of *screening off*.

### 4.1. Common cause

Let us consider a typical case of spurious correlation, the situation in which whenever the atmospheric pressure in a certain region drops below a certain level, the height of the column of mercury in a particular barometer also drops and after a short time a storm occurs. The drop of the mercury column (*A*) is therefore associated to the occurrence of a storm (*B*): therefore it *raises* the probability of the storm and its probability is raised by the occurrence of the storm itself; in symbols:

$$P(A \mid B) > P(A \mid \sim B), \qquad (4.1)$$
$$P(B \mid A) > P(B \mid \sim A).$$

But even if *A* is a *probability raiser* of *B* and *B* a probability-raiser of *A*, neither is cause of the other: the right causal picture is restored if we consider a third factor, the atmospheric pressure (*C*). If we analyze the relations of probabilistic dependency among the factors *A*, *B*, and *C*, we notice two important things:

– *C* is a *probability raiser* of each of the other two factors as $P(A \mid C) > P(A \mid \sim C)$ and $P(B \mid C) > P(B \mid \sim C)$;

– *C* has the further property of cancelling, whenever we condition on *C*, the positive statistical association between *A* and *B*. *C* is said to *screen A off from B* if *A* and *B* are statistically correlated, but the correlation vanishes once we consider *C*. The vanishing of the statistical

association is more precisely stated in terms of *probabilistic conditional independence*:

If P(*A* & *B*) > 0, then *A* and *B* are probabilistically independent given *C* if and only if P(*A* & *B*| *C*) = P(*A* | *C*) P(*B* | *C*). (4.2)

The diagram in figure 4.1 represents the causal structure of the barometer/storm/pressure example:

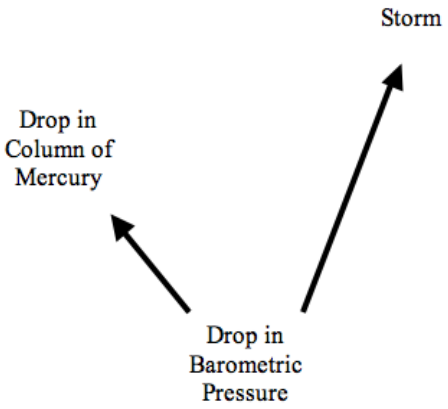


**Fig. 4.1**

In this case the role of the screening factor, *C*, is that of a common cause; screeners are often referred to as *confounders*, or *confounding factors*.

## 4.2. Causal intermediates

There is another important type of screening off, due to the presence of causal intermediates (Figure 4.2).

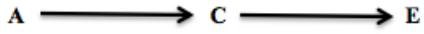


**Fig. 4.2**

In this case we have a causal chain, where the direct cause *C* screens off the effect *E* from *A*; in this case we still want to say that *A* is a cause of *E*, even if it is an *indirect* cause.

## 5. The Principle of Common Cause

The notion of common cause plays a fundamental role in Reichenbach's conception of causation. «It will be advisable – he says in introducing his argument (Reichenbach 1956, 158) to treat the principle of a common cause like a statistical problem». The *Principle of Common Cause* states that «if coincidences of two events A and B occur more frequently than would correspond to their independent occurrence», that is, in statistical terms, if the following formula is satisfied[6]

$$P(A \,\&\, B) > P(A)\, P(B), \qquad (5.1)$$

then the explanation of their association should be ascribed to a third factor $C$ [with $0 < P(C) < 1$], which causes $A$ and $B$. A *conjunctive fork* is a causal structure (such as that in figure 4.1) where $A$ and $B$ occur simultaneously, and $A$, $B$, $C$, satisfy the following conditions:

$$P(A \,\&\, B \mid C) = P(A \mid C)\, P(B \mid C) \qquad (5.2)$$

$$P(A \,\&\, B \mid \sim C) = P(A \mid \sim C)\, P(B \mid \sim C) \qquad (5.3)$$

$$P(A \mid C) > P(A \mid \sim C) \qquad (5.4)$$

$$P(B \mid C) > P(B \mid \sim C), \qquad (5.5)$$

Here (5.2) and (5.3) describe the fact that C screens off $A$ from $B$ and (5.4) and (5.5) convey the dependence, respectively, of $A$ from $C$ and of $B$ from $C$; Reichenbach (1956, 159-161) shows that (5.1) is derivable from (5.2) - (5.5). If $C$ satisfies (5.2) - (5.5), and there is no other factor that satisfies these conditions, then $ACB$ form a conjunctive fork 'open to the future': $C$ is the common cause of $A$ and $B$ and it precedes in time both $A$ and $B$. According to Reichenbach, the majority of conjunctive forks are open to the future, and this hypothesis allows us to derive temporal order from causal asymmetry, which in turn is based on a particular set of statistical relations. He concedes that there are conjunctive forks open to the past, but they are defined by a different set of statistical relations. As Hitchcock (2010) notes, «Reichenbach considered this asymmetry to be a macrostatistical analogue of the second law of thermodynamics»[7].

[6] The notation has been changed for coherence.

[7] «Reichenbach (1956) saw his fork asymmetry as a macro-statistical analog of the second law of thermodynamics. The idea is roughly along the following lines. Suppose we have a system such as a beach that is essentially isolated from the rest of its environment. Suppose

## 5.1. Suppes' Definition

An alternative way of defining probabilistic causation consists in defining probabilistic causation by "building in" the definition of the temporal order (i.e. the fact that causes precede their effects):

*Probabilistic Causation* – If $C_t$ is an event occurring at time *t*, $E_{t'}$ is an event occurring at time *t'*, and *t* precedes *t'*, then $C_t$ causes $E_{t'}$ if and only if:

*i)* $P(E_{t'} \mid C_t) > P(E_{t'} \mid \sim C_t)$;

*ii)* There is no further event $B_{t''}$, occurring at a time *t''* earlier than or simultaneously with *t*, that screens $E_{t'}$ off from $C_t$.

A definition of this kind is presented by Suppes (1970). First, Suppes gives a preliminary definition of *prima facie* causes - in short, prima facie causes are *probability raiser* with a probability greater than *0* that precede their putative effect in time. He then defines *spurious causes* as those prima facie causes which are screened off from their effects by members of a partition of events which precede the effect. Finally, he defines *genuine causes* as those prime facie causes that are not spurious. While Reichenbach focused on structures as the one in figure 4.1, Suppes' definition is inspired by cases like the causal chain depicted in figure 4. 2.

Our brief sketch of the first period of probabilistic approaches to causation should include a presentation of the contribution given by Good (1959, 1961a, 1962). We will not provide it, referring instead, for a first presentation, to Williamson (2009). In the next section we will sketch some of the main problems linked to the definition of causation *à la* Suppes-Reichenbach, before introducing the new formal approach to the philosophy of probability and causation.

---

moreover that we find this system in a state of low entropy; e.g. there are footprints on the beach. The second law of thermodynamics tells us that the system did not spontaneously evolve into this state; rather, the low entropy state must be the result of an earlier interaction with some other system (a human walking on the beach). This interaction 'prepares' the system in a low entropy state, but once the system is isolated, its entropy will increase. Now suppose that we have two events *A* and *B*. If we hold fixed the probability of each event individually, a probability distribution over the partition *A*&*B*, *A*&~*B*, ~*A*&*B*, ~*A*&~*B* will have more information, in the sense of Shannon (1948), when *A* and *B* are correlated. The formal definition of entropy is closely related to that of information, the two being inversely proportional. So when *A* and *B* are correlated, we have the analog of a system in a state of low entropy. This state is then to be explained in terms of some earlier event that prepares the system» **(**Hitchcock 2010).

## 6. Probability Raising is an Unnecessary Condition for Defining Causation

If we wanted to summarize the preceding sections, we could say that a definition of probabilistic causation in terms of probability raising (or probability modifying) should be accompanied by something like *No-Screening off Condition*, stating that *C* causes *E* just in case it raises the probability of *E* and there is no third factor (or conjunction of factors) *C* such that *C* screens off *C* from *E*. However, even this definition must face serious counter-examples, notably those deriving from causes that do not raise (modify) the probability of their effects. One of the most well known examples is given by Skyrms (1980). In general, smoking is a positive cause of lung cancer. But suppose that, due to air pollution (which is also assumed to be another positive cause of lung cancer), city-dwellers tend to stop smoking in order to protect their lungs, whereas in the country people feel freer to smoke. Then, smoking is a positive cause of lung cancer, living in the country is negative cause of lung cancer and a positive cause of smoking; however, if the city air quality is bad enough to cause lung cancer to a high number of non smokers, then the frequency of lung cancer in the entire population can be lower among smokers than among not smokers; in this case smoking will be negatively correlated with lung cancer even if it is a positive cause of it.

In this example, living in the country is a common cause of lung cancer (negative) and smoking (positive), but does not screens off lung cancer from smoking because smoking causally affects lung cancer in an independent way. Therefore, *No-Screening-Off Condition* is *insufficien*t to restore the right causal picture. Other well known examples of contrast between probabilistic influence and causal influence are given below, together with attempts to show how to overcome this problem.

## 7. Hesslow Problem and Simpson's Paradox

A well known example of causation which is not reflected by probability raising is given by Hesslow (1976, 192):

> It is possible however that examples could be found of causes that lower the probability of their effects. Such a situation could come about if a cause could lower the probability of other more efficient causes. It has been claimed, e.g., that contraceptive pills *(C)* can cause thrombosis *(T)*, and that consequently there are cases where $C_t$ caused $T_t$. But pregnancy can also cause thrombosis, and *C* lowers the probability of pregnancy. I do not know the values of *P(T)* and *P(T/C)* but it seems possible that *P(T/C)* < *P(T)*, and in a population which lacked other

> contraceptives this would appear a likely situation. Be that as it may, the point remains: *it is entirely possible that a cause should lower the probability of its effect.*

In Hesslow's example, we have causation unaccompanied by probability-raising. In general, those cases in which two or more properties, which are positively or negatively correlated or independent from each other, exhibit a "sign reversal" – or become uncorrelated when we consider the subpopulations separately – are called "Simpson's paradoxes". Here we present a typical example of Simpson's paradox, following Pearl (2000).

Imagine a drug trial (where some subjects are given the drug and the others are given a placebo) where the drug appears positively associated to recovery in the overall population, but negatively associated to recovery in the two subpopulations of males and females.

Here the 'factors' are: *Recovery*, *Drug*, *Female*. The absence of each factor is represented, as usual, by the negation symbol '~'.The sign-reversal is represented by the following inequalities:

*i)* P(*Recovery* | *Drug*) > P(*Recovery* | ~*Drug*)

*ii)* P(*Recovery* | *Drug* & *Female*) < P(*Recovery* | ~ *Drug* & *Female*)

*iii)* P(*Recovery* | *Drug* & ~*Female*) < P(*Recovery* | ~ *Drug* & ~*Female*).

The data in following three tables show, respectively, the different frequencies of recoveries of patients who received the drug with respect to those who received the placebo in the whole group (*a*), in the subpopulation of female patients (*b*) and in the subpopulation of male patients (*c*).

| **Combined** | Recovered | Not recovered | Total | Recovery rate |
|---|---|---|---|---|
| Drug | 20 | 20 | 40 | 50% |
| No drug | 16 | 24 | 40 | 40% |
| | 36 | 44 | 80 | |

*(a)*

| **Females** | Recovered | Not recovered | Total | Recovery rate |
|---|---|---|---|---|
| Drug | 2 | 8 | 10 | 20% |
| No drug | 9 | 21 | 30 | 30% |
| | 11 | 29 | 40 | |

*(b)*

| **Males** | Recovered | Not recovered | Total | Recovery rate |
|---|---|---|---|---|
| Drug | 18 | 12 | 30 | 60% |
| No drug | 7 | 3 | 10 | 70% |
| | 25 | 15 | 40 | |

*(c)*

If we look at the three tables, we see that *(a)* seems to point toward drug efficacy, as the rate of recovery is 10% higher among the treated patients. Therefore, it seems that we are justified in asserting:

*iv)* P(*Recovery* | *Drug*) > P(*Recovery* | *~Drug*).

However, tables *(b)* and *(c)* show a recovery rate which is 10% *lower* both among female patients and among male patients, displaying a value reversal with respect to the mixed group:

*v)* P(*Recovery* | *Drug* & *Female*) < P(*Recovery* | *~Drug* & *Female*),

*vi)* P(*Recovery* | *Drug* & *~Female*) < P(*Recovery* | *~Drug* & *~Female*).

In our example, the drug appears to be a positive cause of recovery in the overall population because male patients, who recover more often than female patients *independently*, also undergo treatment with higher frequency. The larger proportion of treated patients and of recovering patients within the male population, therefore, masks, in the whole population, the true causal picture. However, if we investigate the effect of the drug in the two subpopulations separately, holding fixed the factor "sex", that in this case results to be a confounding factor, the right causal picture emerges.

The philosophical moral of these examples is that *probability raising* plus *Screening Off* are neither sufficient for analyzing causation nor for licensing correct causal judgements: it is also important that causal links are evaluated with respect to the 'right' populations. But how should we identify the right populations? Are there general rules to do so? Cartwright (1979, 423) suggests incorporating the requirement that probability raising is ascertained in causally homogeneous populations into the definition of probabilistic causation in the following way:

> *Definition of probability:* C *causes* E if and only if C increases the probability of E in every situation which is otherwise causally homogeneous with respect to E.

In more formal terms:

> *Contextual Unanimity*: C causes E if and only if P(E| C & B) > P(E| ~C & B) in every background context.

Skyrms (1980) offers a weaker formulation of this principle, requiring that the alleged cause raises the probability of the presumed effect in at least a background context and lowers it in none. *Background contexts* are defined as conjunctions of factors which are causally relevant with respect to the causal relation under enquiry. Due to this appeal to (other) causally relevant factors, neither Cartwright's nor Skyrms' theories are to be considered as reductive theories of probabilistic causation.

## 8. Causal modeling

Starting from late '80, the research on the relationship between causal and statistical dependencies has become interdisciplinary, and it has undergone a strong mathematization, due also to the contribution from research in artificial intelligence, at least since the works by Pearl (1988, 2000) and Spirtes, Glymour & Scheines (1993). The result of this area of research is called *causal modeling* and, due to its ability to offer a systematized account of formerly sparse statistical, philosophical and mathematical results, has become the state of the art concerning the approach to causal inference from statistical data in many scientific areas, with particular respect to epidemiology and the social sciences. From a philosophical point of view, the research of causal modeling has yield an important clarification of many principles both of *causal reasoning* and on *reasoning about* causation, but at the same time has raised a lively debate on the validity of its methods and principles. In this section we will present a brief sketch of the theory.

Our introduction starts with the definition of *Bayesian networks.* A Bayesian network is a mathematical object consisting of:

> *i*) a directed acyclic graph **G**, i.e., a set of nodes and a set of arrows that connect pairs of nodes, and in which there are not cycles (like $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_1$). Each node is associated to a *variable*, whose values represent the occurrence/non occurrence of events; for brevity, the terms "node" and "variable" are often used interchangeably. Variables can be binary (e.g. the variable FEVER having value 1 or value 0 represents the presence or the absence of fever), or range over richer sets of values

(e.g. BODY TEMPERATURE = ⟨32, …, 41⟩. The set of variables is denoted by "**V**". The relations between subsets of variables are described as family relations: an arrow from $X_1$ pointing directly to $X_2$ is said to be a *parent* of $X_2$, where $X_2$ is said to be a *child* of $X_1$. The set formed by the children of $X_1$, the children of children of $X_1$, …, is called *the set of the descendants* of $X_1$;

*ii*) a probability distribution on the variables of **G**, such that for each node $Y$ with parents $X_1$, …, $X_n$ is specified a conditional probability distribution $P(Y \mid X_1, \ldots, X_n)$.

An important assumption concerning the relationship between the graph and the probability distribution is the *Markov Condition* (*MC*):

(*MC*) In a Bayesian network any node is conditionally independent of its non-descendants, given its parents.

In figure 8.1, the probability of the node $X_5$ is independent of its non-descendants $X_1$ and $X_4$ conditionally on its parents $X_3$ and $X_2$.

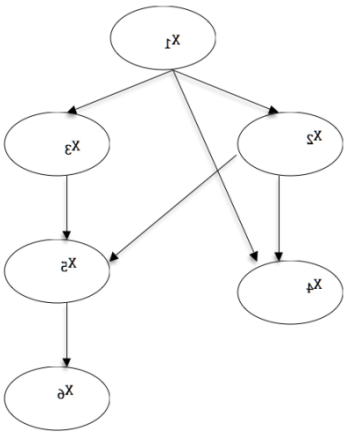

**Fig. 8.1**

The Markov Condition turns out to be equivalent to a relation, called *d-separation*, which allows a 'reading' of the independence relationships among the variables from the graph.[8] Besides Markov Condition, two other conditions can be – and often are – assumed:

[8] For an introduction to d-separation see e.g. Scheines (2005).

> *Minimality* – no edge can be removed from the graph with the resulting sub-graph violating the Markov condition,

And

> *Faithfulness* – probabilistic independencies are due only to the holding of the Markov Condition, and not to incidental mutual cancellations of probability values.

Originally Bayesian networks were devised as an useful tool for representing and inferring probabilistic dependencies. Given the close relationship between probabilistic dependencies and causal dependencies, it is not surprising that they have been used also for causal reasoning. In order to be an useful tool in this area, however, Bayesian networks must be given a causal interpretation. Under such an interpretation, the arrows of the graph represent direct causal relationships. For *causal* Bayesian networks the following *Causal Markov Condition* (*CMC*) holds

> *(CMC)* In a causal Bayesian network, any node is probabilistically independent of its *non-effects*, conditional on its direct causes.

As Williamson (2009) points out, «*CMC* implies the Principle of Common Cause in the following version: if variables *A* and *B* are probabilistically dependent then one causes the other, or there is a set *U* of common causes in *V* which screen off *A* and *B*». He also remarks that this version of the Principle of Common Cause is also a consequence of Reichenbach's own version «under a suitable mapping between events and variables».

How does causal modeling fares with respect to problems like Hesslow's problem and Simpson paradox? A general answer could be that by the constant request of making every step and assumption explicit, classical problems can be handled quite well: every step of the building of the model, and every causal assumption about both the causal scenario under enquiry and the inferential principles adopted in the process of drawing causal inferences will be expressed explicitly, helping to disambiguate what is hidden in Hesslow's problem and Simpson Paradox. However, recent researches have made clear that good causal modeling requires good background causal knowledge, and some principles are still controversial. In cases *à la* Hesslow, for example, the possibility that the tendency of birth

control pills to (directly) cause thrombosis and the tendency of birth control pills to prevent thrombosis by preventing pregnancy (which in turns is a cause of thrombosis) is excluded by the Faithfulness Condition that, as we have seen, implies that the causal influences of a variable on another along different routes do not cancel each other. As many critics have pointed out, however, it is difficult to find metaphysical compelling reasons to accept Faithfulness, which seems therefore to be an object of methodological decisions.

As far as Simpson's paradox is concerned, it is clear that it can "dissolve" once we condition on the variable "Sex" (= male, female): even it is not perfectly clear in what sense the sex is "a cause" of recovery, in Simpson's scenario it causally influences both compliance to the treatment and recovery, so it should be held fixed and the "true" causal history is revealed by the disaggregated data. But it is not always easy to decide what factors should be hold fixed, and there are not methods based on purely statistical criteria for identifying confounders. The theory of causal modeling offers a powerful method to deal with confounding biases, based on the so called "back-door criterion" (see, for example, Pearl 2013). However, applying back-door criterion correctly requires the building of a "good causal model" and this, in turn, seems to require a large amount of (causal) background knowledge. It should not be surprising, therefore, that many authors working with causal models endorse a non–reductionist position. We will discuss later open questions in the metaphysics of probabilistic causation; here we just point out some cases in which the Markov Condition does not hold. The Markov Condition can fail for the following reasons:

a) the variables are linked by intrinsically non-causal probabilistic dependencies, such as logical relations, mathematical relations (e.g. are connected by a mathematical equation), or semantic relations (e.g. synonymy);

b) the variables are connected by a non-causal physical law;

c) the set of variables **V** contains variables which are linked by a common cause that is not included in **V**: in this case **V** is said to be *causally insufficient*;

d) the population is selected by a biased procedure;

e) quantum mechanical systems.

Most of these cases are ruled out by explicit prescriptions: for example, Spirtes, Glymour & Scheines (1993) explicitly rule out causally insufficient sets of variables. They conceive principles as the Faithfulness Condition, the Sufficiency Condition and Causal Markov Condition neither as a dogma nor as parts of a reductive definition of causality, but rather as working hypotheses stating that the «Markov Condition is not given by God; it can fail for various reasons [...]. The reliability of inferences based upon the Condition is only guaranteed if substantial assumptions obtains»[9]

It is therefore clear that, in this framework, principles as the Faithfulness Condition, the Sufficiency Condition and Causal Markov Condition itself appear more as methodological principles than constitutive traits of causation. Moreover, from an epistemological point of view, the opinion that causal reasoning and causal ascertainment require a great deal of knowledge, seems rather uncontroversial; however, from an ontological point of view, there seems to be space for a strong contrast: on the one hand, even most "founding fathers" of causal modeling declare themselves anti-reductionist (Pearl), or metaphysically "neutral" (Spirtes, Glymour and Scheines 1993); on the other hand, some authors (e.g. Papineau 1993, 2001; Spohn 2001; Thalos 2002) use causal modeling to argue in favour of reductionism[10]. Our last remarks will be devoted to this contrast of interpretations.

## 9. Metaphysical Questions

### 9.1. Are general causes reducible to particular causes?

Any theory addressing the metaphysics of causation should specify *what* is connected by a causal link: events, tropes, states of the world? The problem of the nature of causal *relata* is common to all philosophical theories of causation, and it does not concern only probabilistic theories of causation. In causal modeling, the prevailing view, with notable exceptions, is that causal *relata* are *events*, and that events can be adequately represented by specific values of the variables in $\boldsymbol{V}$. A specific and relevant question inside this framework is whether causal relations apply to *types of events* (as in «Drinking hemlock causes death») or to *particular events* (as in «Socrates' drinking hemlock caused Socrates' death»). Many authors agree on the opportunity of tackling type causation and token causation separately, but this

[9] Spirtes, Glymour and Scheines (1993, 9).
[10] It should be noted that Spohn (2001) qualify its reductionism as epistemological.

attitude avoids the problem of deciding whether one of these two kinds of causation depends on the other, and consequently to which kind of causation we should assign ontological priority. In defending the independence of the two theories, Eells (1991) proposes some convincing examples of general causal statements which could be assigned a truth-value even in case no particular instance ever occurred (e.g. «Drinking a pint of plutonium causes death»). As we will see, the question of general/particular probability is linked to the chosen interpretation of probability. It is worth noting, however, that the technical machinery of causal modeling can be applied either to singular or to general causation.

## 9.2. Are causes reducible to probabilities?

Within causal modeling approach, the possibility of reducing causal dependences to patterns of probabilistic (in)dependencies seems to be hampered by the so-called *statistical indistinguishability* of some patterns. Given three variables *A*, *B*, *C* and a probability distribution on the variables such that *C* screens off *A* from *C*, the three following graphs are compatible with the statistical data:

$$i)\ A \leftarrow C \rightarrow B;$$

$$ii)\ A \rightarrow C \rightarrow B;$$

$$iii)\ B \rightarrow C \rightarrow A.$$

**Fig. 9.2.1**

In other terms, the DAGs *i*)-*iii)* are all acceptable as representations of an unique set of probabilistic independence relations; however, the causal structures represented by each graph are quite different. It is therefore clear that in this case the statistical data alone underdetermine the causal picture; in such cases, we say that the graphs are statistically indistinguishable. Obviously, if we knew the temporal ordering of the variables, the underdetermination would vanish: e.g., if we knew that *C* occurred before *A* and *B*, it would be easy to recognize that *C* is the common cause of *A* and *B* and that the right graph is *i)* in figure 9.2.1 (at least if we accepted the assumption that causes precedes their effects). But if we aim to the defend the metaphysical program of reducing causes to probabilities, we may not want

to be forced to accept the assumption of temporal precedence of causes; therefore, the fact that in situations as simple as the one picture in figure 9.2.1 we are not able to elicit one causal structure seems to threaten any reductionist program. A possible way out is given by what has been sometimes called "the third arrow strategy" (see for example Papineau 1993). It is based on the assumptions that when the set of probabilistic (in)dependencies let us in doubt between different causal structures, we can search for further variables that ultimately will reveal the "right" direction of the arrows between the variables under enquiry.

Consider, again, figure 9.2.1 and suppose, as Papineau (1993, 240) suggests, that there is some further variable (*D*) which satisfies the following probabilistic relationships:

- *D* is not correlated with *A*;
- *D* is correlated with *B* and *C*;
- *D*'s correlation with *B* is screened off by *C*;
- *D*'s correlation with *C* is not screened off by anything.

If we assume that correlation is essential for causation, we can identify the correct diagram as the second chain in figure 9.2.1, otherwise) we should see *D* correlated with *A*.

As shown by figure 9.2.2 the right causal structure is obtained as a part of a larger structure.

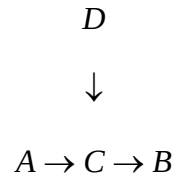


**Fig. 9.2.2**

In short, when causal judgement is made impossible from statistical indistinguishability, the third arrow strategy invites us to search new variables and new probabilistic independencies in order to give the right direction to our (causal) arrows. Papineau quotes a theorem by Spirtes, Glymour e Scheines stating that, assuming adequate screening off conditions, for any set of probabilistically related variables, there is a wider set such that the

conditions will fix the causal order of the original variables. He adds that «[i]t is of course a contingent matter whether such a *possible* wider set is *actually* available in every case, that is, whether, for any causally ambiguous probabilistic structure […] there is always a wider structure […] which disambiguates it» (Papineau 1993, 241); however, if we assume that there are such wider structures, we accept a reduction of causes to probabilities.

## 9.3. Objective and subjective probabilities

Whether one considers causes reducible to probabilities or merely linked to them, one should ask herself what kind of probabilities are at stake. While technical theories of probabilistic causality such as Suppes (1970) and Spirtes, Glymour & Scheines (1993) are pluralistic, and as such admit different interpretations of probability and causality, other theories try to be more specific and to give a metaphysical characterization of causation. Contemporary philosophy of probability is unanimous in distinguishing (at least) two kinds of probability: *physical* probabilities, seen as parts of external world and as such independent of our minds, and *mental* probabilities, which are defined with respect to epistemic agents. Frequency theories of probability (such as von Mises' and Reichenbach's), as well as propensity theories (such as Popper's) belong to the first kind of theories, while personalist (such as de Finetti's and Ramsey's) and logical interpretations of probability (such as Carnap's and Keynes') belong to the second kind.

An analogous distinction could be traced between interpretations of causation that see causes as related to our minds (such as Hume's and Kant's) and interpretations that see causes as features of the world. Combining the interpretations, we would obtain, as observed by Williamson (2005) four kinds of combinations:

a) Physical probabilities, physical causes

b) Physical probabilities, mental causes

c) Mental probabilities, physical causes

d) Mental probabilities, mental causes.

Williamson (2009) examines the various possibilities in depth; here we give just few short remarks on two points.

Firstly, if mental is intended as subjective, *c)* must be excluded, because we should impose a physical constraint (Principle of Common Cause, or Causal Markov Condition) on our probabilities, as the Causal Markov Condition, where causality is interpreted physically, implies that the agent's degrees of belief must satisfy certain independence relationships; therefore a physical interpretation of causality would conflict with a strict subjectivist notion of probability.

Secondly, Williamson rejects *a)* on the basis of the counterexamples that can be raised against the Causal Markov Condition, when physically interpreted. There are two possible objections against the assumption of the Principle of Common Cause under a physical interpretation. The first one has been highlighted by the well known remark by Sober (2001) that if we observe, say, that the price of the bread in Britain and the level of the water in Venice have both been steadily rising since records began, we should be forced, by PCC, to induce a common cause of the two phenomena, which is obviously absurd. Against this, Papineau (2001) claims we should not look for causal links between factors that are not spatio-temporally correlated. He is prepared to accept the idea that, if we would be able to reconstruct the huge Bayesian network formed by probabilistic dependencies at the right level (where probabilities are interpreted as frequencies), this would be all that there is in causation. But what is the right level? This question brings us to the second main group of objections raised against the physical interpretation of PCC, i.e. the fact that PCC (and CMC) seem to not work in quantum mechanical systems[11].

## 9.4. Macro-world and micro world, determinism and indeterminism

Despite its ambitious title, this section will be very short, aiming exclusively to point to some work in progress. As already mentioned, some of the outstanding authors in the field of causal modeling, like Spirtes, Glymour and Scheines, and Pearl, are not particularly disturbed by the (possible) failure of Causal Markov Condition in the micro-world; being non-reductionist they believe that even if some assumptions of causal modeling are not ap-

[11]CCP fails for certain quantum systems involving distant correlations. For example, if we have two particles in the singlet state, and measure the spin of each in, say, the vertical direction, we will find that the probability of spin up equals the probability of spin down equals .5 for both particles. The probability that particle one is spin up while particle two is spin down is not .25 but .5, so the two measurement results are correlated. However, it can be shown that there is no (local) common cause that screens off the two measurement outcomes.(Hitchcock 2010)

plicable in some contexts, this is not a good reason to believe that that assumptions are not applicable in other contexts. Causal modeling does not need to commit to indeterminism: Pearl assumes a Laplacean position, where Spirtes, Glymour and Scheines use the expression «pseudo-indeterminism» to refer to those situations that are deterministic, but such that we don't have sufficient information for knowing all the relevant facts. However, the failure of Causal Markov Condition - and of the Principle of Common Cause - with respect to quantum phenomena could reveal itself a trouble-maker for a position that

- adopts a physical interpretation of both probability and causality,
- states that causes are reducible to probabilities (and, possibly, that there is a wide underlying Bayes Network corresponding to the causal structure of the world)
- places the fundamental level of causality in the micro-world.

However, the real import of the alleged incompatibility of CMC with quantum theory is controversial. On the one Side, Hausman & Woodward (1999) have argued against the legitimacy of considering EPR as a genuine counterexample to CMC; on the other hand, the fate of CMC could be linked to the chosen interpretation of quantum mechanics (see Suárez & San Pedro 2010).

## 10. Conclusive remarks

From this very sketchy survey we might gain a general perspective on the area of recent research in probabilistic causality: on the epistemological and methodological side the mathematization linked to causal modeling had brought about a paradigm shift, and what once were revolutionary studies, now seem to be configuring as normal research, solving puzzles and applying causal modeling methods to new problems and disciplines. On the other hand, on the metaphysical side we still see many conflicting opinions. In particular, the multiplicity of links between probabilities, causes and the foundations of physics seems to sharply deny the Russellian description of causation as «a relic of a bygone age».

**Acknowledgments**

I would like to thank the organizers and the students of the XVI *Scuola Estiva di Filosofia della Fisica*, Cesena, September 20th, 2013, where the content of this paper was presented as introductory lecture on the topic. I have kept here the idea of a presentation of the field, trying to summarize the main ideas and the main points of contrast. Rather than a 'state of the art' this paper aims at being an introduction to a classical debate and a suggestion at following metaphysical problems that technical solutions cannot answer.

**References**

Cartwright, N., 1979, "Causal Laws and Effective Strategies", *Nous*, 13, pp. 419-37.

Davis, W. A., 1988, "Probabilistic theories of causation", in *Probability and Causality. Essays in Honor of Wesley C. Salmon*, J. H. Fetzer (ed.), Dordrecht, Reidel, pp. 133-160.

Eells, E., 1991, *Probabilistic Causality*, Cambridge, Cambridge University Press.

Galavotti, M.C., Suppes, P. and Costantini, D., (eds.), 2004, *Stochastic Causality*, Stanford, CSLI Publications.

Glynn, L., 2011, "A Probabilistic Analysis of Causation", *British Journal for the Philosophy of Science*, 62 (2), pp. 343-392.

Good, I. J., 1959, "A Theory of Causality", *British Journal for the Philosophy of Science*, 9, pp. 307-10.

Good, I. J., 1961, "A Causal Calculus I", *British Journal for the Philosophy of Science*, 11, pp. 305–18.

Good, I. J., 1962, "A Causal Calculus II", *British Journal for the Philosophy of Science*, 12, pp. 43–51. "Errata et corrigenda", vol.13, 88.

Hausman, D. M. and Woodward, J., 1999, "Independence, Invariance and the Causal Markov Condition", *British Journal for the Philosophy of Science*, 50, pp. 521–83.

Hesslow, G., 1976, "Two Notes on the Probabilistic Approach to Causality", *Philosophy of Science*, 43, (2), pp. 290-292.

Hitchcock, C., 2010, "Probabilistic Causation" in *The Stanford Encyclopedia of Philosophy* (2010 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/entries/causation-probabilistic/

Hume, D., 1748, *An Enquiry Concerning Human Understanding*.

Munafò, M. R. and Araia, R., 2010, "Editorial: Cigarette smoking and Depression: a Question of Causation", *The British Journal of Psichiatry*, 196, pp. 425-426.

Papineau, D., 1993, "Can We Reduce Causal Directions to Probabilities?", in D. Hull, M. Forbes, and K. Okruhlik (eds.). *PSA 1992*, East Lansing, Philosophy of Science Association, pp. 238 - 252.

Papineau, D., 2001, "Metaphysics over Methodology – or, Why Infidelity Provides No Grounds to Divorce Causes from Probabilities", in M.C. Galavotti, P. Suppes, and D. Costantini (eds.), *Stochastic Causality*, Stanford, CSLI Publications, 2004, pp. 15-38.

Pearl, J., 1988, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, Kaufmann.

Pearl, J., 2000, *Causality: models, reasoning, and inference*, Cambridge, Cambridge University Press, Second Edition 2010.

Pearl, J., 2013, "Understanding Simpson's Paradox", *Technical Report R-414*, December 2013, ftp://ftp.cs.ucla.edu/pub/stat_ser/r414.pdf , edited version forthcoming in *The American Statistician*, 2014.

Reichenbach, H., 1933, "Die Kausalbehauptung und die Möglichkeit ihrer empirischen Nachprüfung", *Erkenntnis* 3, pp. 32-64.

Reichenbach, H., 1956, *The Direction of Time*, Berkeley, University of California Press.

Reid, T., 1785/2002, *Essays on the Intellectual Powers of Man*, ed. by D. Brookers, University Park, Pennsylvania State University Press.

Scheines, R., 2005, "Causation", in *New Dictionary of the History of Ideas*, ed.by. M.C. Horowitz, New York, Charles Scribner and Sons, Vol. 1, pp. 280-289.

Shannon, C.E., 1948, "A Mathematical Theory of Communication", *Bell System Technical Journal*, 27, pp. 379-423, 623-656.

Skyrms, B., 1980, *Causal Necessity*, New Haven-London, Yale University Press.

Sober, E., 2001, "Venetian Sea Levels, British Bread Prices, and the Principle of the Common Cause", *British Journal for the Philosophy of Science*, 52 (2), pp. 311-346.

Spirtes, P., Glymour, C., and Scheines, R., 1993, *Causation, Prediction and Search*, New York, Springer, Second edition 2000.

Spohn ,W., 2001, "Bayesian Nets Are All There Is to Causal Dependence", in M.C. Galavotti, P. Suppes, D. Costantini (eds.), *Stochastic Causality*, Stanford, CSLI Publications, 2004, pp. 157-172.

Suárez, M. and San Pedro, I., 2010, "Causal Markiv, Robustness and the Quantum Correlations", in M. Suárez (Ed.), *Causes, Probabilities and Propensities in Physics*, New york, Springer, pp.173-93

Suppes, P., 1970, "A Probabilistic Theory of Causality", *Acta Philosophica Fennica*, 24, pp. 1-130.

Thalos, M., 2002, "The Reduction of Causal Processes", *Synthese*, 131 (1), pp. 99-128.

Williamson, J., 2009, "Probabilistic Theories", in *The Oxford Handbook of Causation*, ed. by H. Beebee, C. Hitchcock and P. Menzies, Oxford, Oxford University Press, pp. 185-212.