A causal Bayes net analysis of dispositions

Alexander Gebharter[1] · Florian Fischer[2]

Abstract
In this paper we develop an analysis of dispositions in terms of causal Bayes nets. In particular, we analyze dispositions as generic cause–effect structures that increase the probability of the manifestation when the stimulus is brought about by intervention in certain circumstances. We then highlight several advantages of our analysis and how it can handle problems arising for classical analyses of dispositions such as masks, mimickers, and finks.

Keywords Dispositions · Causal Bayes nets · Masks · Mimickers · Finks

1 Introduction

In this paper we develop an analysis of dispositions on the basis of causal Bayes nets.[1] Causal modeling techniques such as causal Bayes nets have already been applied to various philosophical problems (see, e.g., Gebharter 2017a, c; Hitchcock 2016; Meek and Glymour 1994; Schaffer 2016). Using the causal Bayes net formalism as a framework for analyzing philosophical concepts and issues intimately connected to

---

[1] Note that we do not claim that causal Bayes nets are necessary for analyzing dispositions. There might be other theories of causation, for example Woodward's (2003) interventionist theory, which might also do a good job. However, we think that causal Bayes nets are particularly interesting because they come with several advantages that might be especially compelling for empirically minded philosophers. We point to some of them below. Also note that we use the word "analysis" in a quite moderate sense. We do not assume that by providing an analysis of a concept B in terms of another concept A one ontologically reduces B to A. We rather understand analyzing B in terms of A as doing conceptual geography (cf. Carroll 1994): It shows how B and A are conceptually related; by learning something about one of the concepts one can, at the same time, learn something about the other.

✉ Alexander Gebharter
alexander.gebharter@gmail.com

Florian Fischer
florian.fischer@uni-siegen.de

[1] Department of Theoretical Philosophy, University of Groningen, Oude Boteringestraat 52, 9712, GL Groningen, The Netherlands

[2] University of Siegen, Adolf-Reichwein-Straße 2, 57076 Siegen, Germany

causation seems promising for several reasons. One advantage of causal Bayes nets is that they make causation empirically tangible. The framework provides powerful tools for formulating and testing causal hypotheses, for making predictions about what would happen under hypothetical interventions, and for the discovery of causal structure (see, e.g., Spirtes et al. 2000). In addition, it can be shown that the theory of causal Bayes nets satisfies standards successful empirical theories satisfy as well: It provides the best explanation of certain empirical phenomena and can, as a whole theory, be tested on empirical grounds (Gebharter 2017b; Schurz and Gebharter 2016). The causal Bayes net framework treats causation like theoretical concepts (such as force, charge, etc.) are treated in scientific theories: Instead of providing an explicit definition of causation, causation is rather implicitly characterized by several axioms which relate it to empirical data (in the form of probability distributions).[2] The theory is, thus, non-reductive and does not come with deep metaphysical commitments, which we take to be an advantage over metaphysically more laden reductive theories of causation. This makes the framework especially attractive for empirically minded philosophers.

In the following we use causal Bayes nets to analyse dispositions as generic cause–effect structures that increase the probability of the manifestation when the stimulus is brought about by intervention in certain circumstances. Such an analysis of dispositions comes with several advantages: It allows one to apply powerful causal discovery methods to find and specify dispositions. But the analysis' main upshot is that it is flexible enough to account for the fact that dispositions might change their behavior in different circumstances. In other words, one and the same disposition may give rise to different counterfactual conditionals if its causal environment is changed. The causal Bayes net framework can be used to study such behavior of dispositions in different causal environments on empirical grounds. Because of this flexibility, our analysis can also provide novel solutions to philosophical problems posed by masks, mimickers, and finks which, one way or another, plague all other accounts of dispositions currently on the market.[3] We agree with Cross (2012) that the "recent literature on dispositions can be characterized helpfully, if imperfectly, as a continuing reaction to this family of counter-examples" (p. 116). Another advantage of our analysis is that it allows for a uniform representation of probabilistic and non-probabilistic dispositions.[4] Other analyses of dispositions often either have trouble switching from non-probabilistic dispositions to probabilistic dispositions, or exclude probabilistic dispositions altogether.

The paper is structured as follows: In Sect. 2 we introduce dispositions and the problems arising for classical dispositional theories due to masks, mimickers, and finks. Then, in Sect. 3, we present the basics of the causal Bayes net framework and our proposal for an analysis of dispositions within this particular framework. We also highlight several advantages of our analysis. In Sect. 4 we finally show how

---

[2] In Glymour's (2004) words, it is a Euclidean rather than a Socratic approach.

[3] Of course, some accounts are able to deal with some of the counterexamples; but there is still no conditional account of dispositions that is able to handle all counterexamples in a satisfying way. Our diagnosis in Sect. 3 will be that conditional accounts have to face these problems due to systematic shortcomings.

[4] Probabilistic dispositional theories have, for example, been put forward by Ellis and Lierse (1994), Popper (1990) and Prior et al. (1982).

our analysis of dispositions can handle problems with masks, mimickers, and finks classical accounts have to face. We illustrate how these problems can be solved by means of prominent exemplary scenarios which shall stand proxy for all kinds of masking, mimicking, and finking cases. We conclude in Sect. 5.

2 Dispositions, classical analyses, and troubles with masks, mimickers, and finks

Disposition ascriptions are frequently made in science as well as in everyday life. We take glass to have the disposition of fragility, sugar to be soluble, and masses to attract each other. Dispositions are intimately connected to their manifestations and stimulus conditions. A lump of sugar, for example, has the disposition of solubility. This means that it will dissolve under certain conditions such as being put into water. The dissolving is the manifestation of the disposition, while the submergence into water is one of the disposition's stimulus conditions.

According to classical approaches, a disposition[5] D corresponds to a stimulus-manifestation pair $\langle \Sigma, M \rangle$ or a class of stimulus-manifestation pairs. For instance, the manifestation of the fragility of a glass can be triggered by striking it with a hammer, by throwing it to the ground, and so on. This disposition's class of stimulus-manifestation pairs would be $\{\langle striking, breaking \rangle, \langle throwing\text{-}to\text{-}the\text{-}ground, breaking \rangle, \dots \}$. The manifestation is roughly the same for all of these stimulus-manifestation pairs, yet it can be triggered by different stimuli. We can thus call dispositions like fragility multi-stimulus dispositions. If a disposition shows different manifestations under different stimulus conditions, it is called a multi-track disposition (Ryle 1949, pp. 43–45; Vetter 2013). If one and the same stimulus can lead to different mutually exclusive manifestations, i.e., if the class of stimulus-manifestation pairs of a disposition contains two pairs $\langle \Sigma, M_i \rangle$ and $\langle \Sigma, M_j \rangle$ with $i \neq j$ which cannot occur together, we are dealing with a probabilistic disposition (Prior et al. 1982, p. 251). Note that this makes the outcome of the disposition (M_i vs. M_j) chancy once the stimulus (Σ) is present.

Due to the pre-theoretical closeness of dispositions to conditionals (cf. Prior 1985, p. 5) a multitude of conditional analyses of dispositions have been proposed over the years (see, e.g., Choi 2006; Lewis 1997; Malzkorn 2000). Though the details of the different conditional accounts vary greatly, all of them share the basic idea that a disposition D should be analyzed as a conditional or a class of conditionals connecting the stimulus conditions Σ_i with the manifestations M_j of the disposition D's stimulus-manifestation pairs $\langle \Sigma_i, M_j \rangle$. If these conditionals are satisfied, the corresponding disposition is present, if not, it is not present, either. Although the early conditional accounts were ontologically reductionistic, analyses of dispositions in terms of conditionals do not have to be (cf. Fischer 2018). Malzkorn (2000, p. 454), for example, claims that his conditional analysis is consistent with a non-reductionistic theory of

---

[5] Lewis (1997) distinguishes between conventional and canonical dispositions. Examples of conventional dispositions are fragility and solubility, whereas canonical dispositions explicitly mention the stimulus and the manifestation (like dissolving when submerged into water). Following this distinction our approach concerns canonical dispositions. However, to keep things simple we will loosely talk about dispositions throughout the paper.

dispositions. Our own analysis will be neutral; it can be interpreted as ontologically reductionistic or non-reductionistic.[6]

Three kinds of problematic cases are notorious in the dispositions debate. Martin (1994) has argued that dispositions might be gained or lost. If a bouncy ball, for example, is deep frozen, it seems to lose its elasticity and to gain the disposition of fragility instead. If this idea is correct, it poses a problem to conditional analyses as the conditions for gaining or losing a disposition can coincide with the stimulus conditions. In order to illustrate this problem, Martin came up with the idea of an electro-fink, a device which can be connected to a dead wire. A dead wire does not seem to have the disposition to be conductive. The electro-fink, however, reliably detects whether a conductor is about to touch the wire. Whenever the wire is touched by a conductor, the electro-fink instantaneously renders the wire live for the duration of contact. Hence, conditional analyses of dispositions will lead to the result that the dead wire has the disposition of being conductive, simply because the electro-fink device ensures that the conditional corresponding to this disposition is true. Note that the fink can also work in a reverse-cycle. In this mode it renders an otherwise live wire dead whilst it is touched by a conductor. A wire in such a device has, according to Martin, the disposition of liveness although it will not conduct electricity if touched by a conductor. Once again the truth of the conditional and the disposition ascription fall apart. As a consequence, the conditional analysis provides false results if dispositions can in fact be lost or gained.

The second threat for conditional analyses of dispositions is posed by masking cases. Masking cases exploit the fact that a disposition does not manifest under all circumstances. As Goodman (1983, p. 36) wrote, "matches do not always light when scratched. They light only if attendant circumstances are propitious". So if circumstances are not propitious, the match does not light. In case of lack of oxygen, for example, the match does not light up when scratched though it is inflammable. Here is another prominent example: Johnston (1992, p. 233) discusses a fragile glass cup strengthened by internal packing material. This packing material masks the fragility of the cup, as the amended cup would not break when struck. Thus, once again, the disposition and the corresponding conditional fall apart. The cup has the disposition of fragility but does not show the manifestation (breaking) under the stimulus (striking the glass). Hence, conditional analyses wrongly result in the glass not having the disposition of fragility.

Although finking and masking cases have originally been put forward in order to attack reductionist analyses of dispositions, they also threaten ontologically non-reductive accounts (cf. Schrenk 2010) such as the ones put forward by Bird (2007) and Ellis (2001). Both in the case of finking and of masking, the manifestation of a disposition is prevented. They only differ in the way the manifestation is prevented. Every theory of dispositions, be it reductionist or non-reductionist, has to account for this problem somehow.

The third kind of scenarios problematic for classical analyses of dispositions are so-called mimicking cases. A prominent example of a mimicker is Lewis' (1997, p.

---

[6] Note that the fact that one concept can be analyzed in terms of another does not imply that the latter is ontologically more fundamental than the former. It might well be ontologically non-reducible or both concepts might be reducible to another ontologically more fundamental concept.

153) hater of styrofoam. A plate made of styrofoam is not fragile yet it produces a distinctive sound when struck. Whenever the hater of styrofoam hears that sound, he rips the plate apart. Thus, even though the plate is not disposed to break when struck, it does break when struck due to the presence of the hater of styrofoam. Examples like this are intended to show that conditional analyses can fail because a disposition's manifestation can be mimicked: The presence of the mimicker makes the conditional corresponding to a disposition true though the object lacks the disposition.

Mimickers are clearly problematic for reductive conditional accounts of dispositions. Like masks and finks they make the truth-value of a disposition ascription and the corresponding conditional fall apart. However, "more is at stake than the fate of conditional analyses" (Cross 2012, p. 120). Cross observes that even if one does not deem them a fit reduction basis, conditionals provide us with one standard way of epistemic access to dispositions. Going one step further, he stresses that the debate about dispositions "will be put to rest only if some conditional-friendly theory is widely acknowledged to be free from counter-example" (p. 121). So, non-reductive accounts seem also to be threatened by mimickers, at least unless another method for epistemic assessment of dispositions is found. Our analysis of dispositions in terms of causal Bayes nets to be developed in Sect. 3 will provide such a method. It will also turn out to be conditional-friendly.

3 Dispositions and causal Bayes nets

3.1 Causal Bayes nets

The causal interpretation of Bayesian networks was developed by Spirtes et al. (1993) and later by Pearl (2000). A causal Bayes net is a triple $\langle \mathbf{V}, \mathbf{E}, P \rangle$. $\mathbf{V}$ is a set of random variables $X_1, \ldots, X_n$ and $\mathbf{E}$ is binary relation on $\mathbf{V}$ (i.e., $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$). Whenever a variable X_i stands in relation $\mathbf{E}$ to another variable X_j, we graphically represent this by $X_i \longrightarrow X_j$. The graph $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ represents a specific causal structure. If $X_i \longrightarrow X_j$ is part of the graph, X_i is interpreted as a direct cause of X_j w.r.t. $\mathbf{V}$, and X_j is interpreted as a direct effect of X_i w.r.t. $\mathbf{V}$. If two variables X_i and X_j are connected by a path of the form $X_i \longrightarrow \cdots \longrightarrow X_j$, then X_i is called a (direct or indirect) cause of X_j, and X_j is called a (direct or indirect) effect of X_i. The set of all direct causes (or causal parents) of a variable X_j is referred to by $\mathbf{Par}(X_j)$. The variables X_j that are connected to a variable X_i via a path of the form $X_i \longrightarrow \cdots \longrightarrow X_j$ are called X_i's descendants. For technical reasons, every variable X_i is assumed to be a descendant of itself. The set of a variable X_i's descendants is referred to by $\mathbf{Des}(X_i)$. While the graph $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ of a causal Bayes net $\langle \mathbf{V}, \mathbf{E}, P \rangle$ captures a particular causal structure, P is a probability distribution over $\mathbf{V}$ that represents the strengths of the causal influences propagated along the direct cause–effect relations of this causal structure.

Causal Bayes nets are assumed to satisfy the causal Markov condition (Spirtes et al. 2000, p. 29):

Definition 3.1 (*Causal Markov condition*) $\langle \mathbf{V}, \mathbf{E}, P \rangle$ satisfies the causal Markov condition if and only if every $X \in \mathbf{V}$ is probabilistically independent of its non-descendants $\mathbf{V} \backslash \mathbf{Des}(X)$ conditional on its causal parents $\mathbf{Par}(X)$.[7]

The causal Markov condition implies several probabilistic independencies for a given causal structure $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$. It is basically a generalization of two ideas that can already be found in Reichenbach's (1956) book *The direction of time*: Firstly, that conditionalizing on all common causes of two correlated variables screens these two variables off each other (if there are no other causal connections between these two variables) and secondly, that conditionalizing on all of a variable's direct causes screens this variable off from its indirect causes.

Whenever a model $\langle \mathbf{V}, \mathbf{E}, P \rangle$ satisfies the causal Markov condition, its probability distribution P over $\mathbf{V}$ factors according to the following Markov factorization (Spirtes et al. 2000, pp. 29f):

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | \mathbf{Par}(X_i)) \tag{1}$$

The conditional probabilities $P(X_i | \mathbf{Par}(X_i))$ appearing on the right hand side of the equation are called X_i's parameters. They represent the strengths of the influences of X_i's direct causes on X_i.

Another important condition is the causal faithfulness condition. If satisfied, it guarantees that causal connections produce probabilistic dependencies. The causal faithfulness condition can be formulated as follows (Spirtes et al. 2000, p. 31):

Definition 3.2 (*Causal faithfulness condition*) $\langle \mathbf{V}, \mathbf{E}, P \rangle$ satisfies the causal faithfulness condition if and only if the probabilistic independencies implied by $\mathbf{G} = \langle \mathbf{V}, \mathbf{E} \rangle$ and the causal Markov condition are all the probabilistic independencies featured by P.

While the causal Markov condition is intended as a general requirement for causal Bayes nets, there are many possibilities how the causal faithfulness condition can be violated, for example by fine-tuning the parameters of a model in such a way that several causal paths cancel each other out. While the causal Markov condition will be crucial for our analysis of dispositions, the causal faithfulness condition (which allows for a drastic reduction of underdetermination in causal search) only becomes important when it comes to finding dispositions.

Before we will develop our analysis of dispositions, we need to introduce two additional causal concepts. The first one is the notion of an intervention. Causal Bayes nets can be used to make two different kinds of predictions: One can predict X_j's probability distribution either after observing that another variable X_i has taken value x_i or after X_i has been forced to take value x_i by intervention.[8] According to the causal

---

[7] X is probabilistically independent of Y conditional on Z if and only if $P(x|y, z) = P(x|z)$ or $P(y, z) = 0$ holds for every combination of X-, Y-, and Z-values x, y, and z, respectively. X, Y, and Z can be variables os sets of variables and probabilistic dependence is defined as the negation of independence.

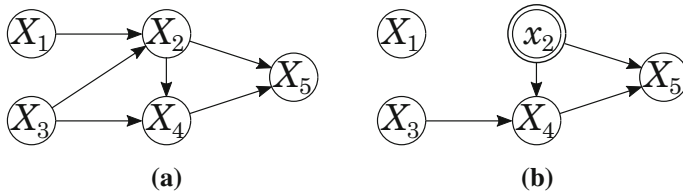[8] X_i and X_j can also be sets of variables.

Fig. 1 Original graph (**a**) used for making predictions under observation and truncated graph (**b**) after bringing $X_2 = x_2$ about by intervention (indicated by a double-circle). While observing $X_2 = x_2$ might have a probabilistic influence on X_2's causes and on X_2's effects, setting X_2 to x_2 by intervention breaks the arrows into X_2 and, thus, can only lead to probability changes of the effects of X_2 (i.e., X_4 and X_5)

Markov condition and Eq. 1, the causal structure $\langle \mathbf{V}, \mathbf{E} \rangle$ of the system of interest implies certain independencies for the associated probability distribution P. If one forces X_i to take value x_i by intervention, one basically generates additional independencies. In particular, the variable X_i on which one intervenes is assumed to be under full control of the intervener and, hence, becomes independent of its direct causes in $\langle \mathbf{V}, \mathbf{E} \rangle$. In other words, the causal arrows pointing at X_i are deleted. One can then apply the causal Markov condition and Eq. 1 to the resulting truncated structure $\langle \mathbf{V}, \mathbf{E}' \rangle$ in order to compute X_j's post intervention probability distribution $P(X_j|do(x_i))$; we use Pearl's (2000, Sect. 7.1) *do*-operator to mark X_i's taking value x_i due to an intervention as "$do(x_i)$". The corresponding observational distribution $P(X_j|x_i)$ can be computed by applying the causal Markov condition and Eq. 1 to the original structure $\langle \mathbf{V}, \mathbf{E} \rangle$. For a graphical illustration, see Fig. 1.

The last causal notion relevant for our analysis of dispositions is that of a causal context. A causal context $\mathbf{C} = \mathbf{c}$ typically consists of several variables $C_1, \ldots, C_n$ taking certain values $c_1, \ldots, c_n$.[9] When speaking of a causal context $\mathbf{C} = \mathbf{c}$, it is often assumed in the literature that the variables $C_1, \ldots, C_n$ making up that causal context are not included in the set of variables $\mathbf{V}$ to be modeled. However, since our analysis (see Definition 3.3) will require that contexts can be fixed by intervention and because interventions and their effects are only defined for specific causal models (satisfying the Markov condition), we will assume throughout the paper that the variables making up a context are included in $\mathbf{V}$. If not explicitly saying otherwise, we will also assume that the values variables take in a context are brought about by intervention. One central feature of causal contexts that will become important later on is that changing between contexts will typically change the probability distribution over other variables in $\mathbf{V}$ and, thus, lead to different predictions about what would happen under different interventions. Note, however, that our analysis will not require that the variables $C_1, \ldots, C_n$ making up a causal context stand in specific causal relationships to other variables in $\mathbf{V}$.

3.2 Analyzing dispositions

We now have all the tools required for our analysis of dispositions within the causal Bayes net framework. We start with the most simple case: a disposition with exactly one

---

[9] Note that the smallest possible causal context is the empty set (i.e., $\mathbf{C} = \emptyset$).

stimulus condition. One of the prime examples of such a disposition is supposably water-solubility. Clearly, the stimulus event (putting a sugar cube u into water) is causally relevant for the manifestation (u dissolves) of the disposition water-solubility. Consequently, we suggest analyzing this disposition as a generic cause–effect relation. So there will be some causal model $\langle \mathbf{V}, \mathbf{E}, P \rangle$ featuring an arrow $W \longrightarrow D$ (or a path $W \longrightarrow \cdots \longrightarrow D$), where W is a binary variable with the values 1 for u is put into water and 0 for u is not put into water, and D is a binary variable with the values 1 for u dissolves and 0 for u does not dissolve. Our analysis of dispositions as generic cause–effect relations nicely corresponds to one of the adequacy conditions for dispositions formulated by Malzkorn (2000). According to Malzkorn, "dispositions are causal properties; the *analysans* of a disposition D must state some kind of causal relation between the corresponding test and the corresponding manifestation" (p. 462). This constraint is based on the idea that just to bring about the manifestation does not suffice. Instead, the object in question has to display the manifestation *because* it is under the stimulus conditions.

A successful analysis of the disposition water-solubility must meet two more requirements: Firstly, the causal model $\langle \mathbf{V}, \mathbf{E}, P \rangle$ mentioned above must be adequate, i.e., $\langle \mathbf{V}, \mathbf{E} \rangle$ must correctly represent a part of the true causal structure of the world and P must correctly represent the regularities among variables in $\mathbf{V}$ to be found in the world.[10] Secondly, objects like u must dissolve at least sometimes if put into water. This means that putting objects like u into water (this is an intervention) must increase the probability of dissolving if the variables in some causal context $\mathbf{C} = \mathbf{c}$ are fixed to their values $\mathbf{c}$ by intervention.[11] More formally: $P(D = 1|do(W = 1, \mathbf{C} = \mathbf{c}))$ must be greater than $P(D = 1|do(\mathbf{C} = \mathbf{c}))$. We will explain why we added the phrase "in some causal context $\mathbf{C} = \mathbf{c}$" and why $\mathbf{c}$ must be brought about by intervention below. But first, in order to get the full picture, let us state the proposed analysis of dispositions more explicitly:

Definition 3.3 (*Disposition*) Objects u of type U have the disposition $[Y = y$ if $X_1 = x_1, \ldots, X_n = x_n]$ if and only if there is a model $\langle \mathbf{V}, \mathbf{E}, P \rangle$ (with $X_1, \ldots, X_n, Y \in \mathbf{V}$) satisfying the causal Markov condition and a context $\mathbf{C} = \mathbf{c}$ (with $\mathbf{C} \subseteq \mathbf{V} \setminus \{X_1, \ldots, X_n, Y\}$) such that

1. $X_1, \ldots, X_n, Y$ describe possible events involving objects u of type U, and
2. $\langle \mathbf{V}, \mathbf{E}, P \rangle$ correctly represents a part of the true causal structure of and the true regularities to be found in the world, and
3. $P(y|do(x_1, \ldots, x_n, \mathbf{c})) > P(y|do(\mathbf{c}))$, and
4. $P(y|do(x_1, \ldots, x_n, \mathbf{c})) > P(y|do(\mathbf{x}, \mathbf{c}))$ holds for every subsequence $\mathbf{x}$ of $x_1, \ldots, x_n$.

The expression between the square brackets […] represents a disposition. The variables $X_1, \ldots, X_n$ describe whether the disposition's stimulus conditions occur and Y

---

[10] What it means for a model to correctly represent and under which conditions a model correctly represents is an open field of research. For the endeavor of this paper we simply assume that this problem can in principle be solved.

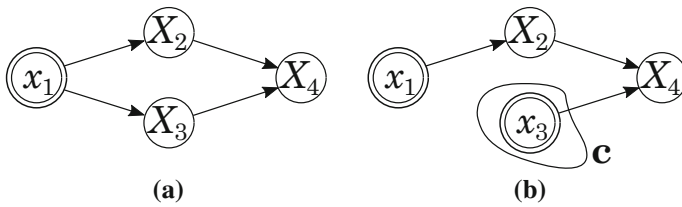[11] Recall that the causal context might be empty (i.e., $\mathbf{C} = \emptyset$).

(a)   **(b)**

Fig. 2 Assume X_1 models the presence of a disposition's stimulus condition and X_4 the presence of its manifestation. If X_1 and X_4 are connected via two causal paths (**a**) that cancel each other out, then intervening on X_1 will not make any difference for X_4. However, in such cases there is a context $\mathbf{C} = \mathbf{c}$ (e.g., $\mathbf{C} = \{x_3\}$) such that producing $\mathbf{c}$ by intervention will block one of these paths (**b**). As a consequence, X_1's causal influence on X_4 will show up when intervening on X_1 if $\mathbf{C}$ is fixed to $\mathbf{c}$

describes whether the disposition is manifested.[12] This disposition is present if there is at least one model $\langle \mathbf{V}, \mathbf{E}, P \rangle$ and one context $\mathbf{C} = \mathbf{c}$ (with $\mathbf{C} \subseteq \mathbf{V} \setminus \{X_1, \ldots, X_n, Y\}$) that satisfy conditions 1–4. Condition 1 guarantees that not just any selection of causal variables can describe stimulus or manifestation conditions of a disposition ascribed to objects u of type U. The variables must describe events involving such objects u such as striking u with a hammer or u breaking, dissolving, etc. Condition 2 guarantees that the causal Bayes net is a correct representation of a part of the world. To this end it must correctly represent a part of the true causal structure of the world and the probabilities featured by P must fit the true regularities to be found in the world. Condition 3 ensures that bringing about the stimulus conditions by intervention increases the probability for the disposition's manifestation if also the variables in context $\mathbf{C} = \mathbf{c}$ take their values due to an intervention. This also guarantees that the variable describing the manifestation is an effect of the variable(s) describing the stimulus condition(s).[13] Adding the context $\mathbf{C} = \mathbf{c}$ is important because dispositions often do not manifest in all, but only in specific circumstances. Because of this they can be tricky to identify (see Fig. 2 for such a more tricky situation). The context is intended to identify such situations in which the disposition manifests. $\mathbf{C} = \mathbf{c}$ must be produced by intervention in order to avoid spurious correlations, i.e., correlations that are not due to the stimulus conditions' causal influence on the manifestation (see Fig. 3 for an illustration). Finally, condition 4 guarantees that each $X_i = x_i$ of the stimulus conditions $X_1 = x_1, \ldots, X_n = x_n$ is actually relevant for the manifestation in context $\mathbf{C} = \mathbf{c}$. Omitting one or several of the stimulus conditions would lower the probability of the manifestation.

Before we highlight some features and advantages of the analysis of dispositions proposed, several comments on our analysis seem to be in order. Firstly, note that while our analysis of dispositions requires the existence of a set of variables satisfying

---

[12] Note that our causal Bayes net analysis is flexible enough to also allow for a representation of multi-track dispositions. One way to do this would be to have more stimulus and more manifestation variables for a single disposition in a causal Bayes net. $X_1 = x_1$ and $X_2 = x_2$ could, for example, stand for two different stimulus conditions, and $Y_1 = y_1$ and $Y_2 = y_2$ for two different manifestations of the disposition.

[13] Recall from Sect. 3.1 that a change in Y's probability distribution after an intervention on X can only occur if Y is a descendant of X in the model.
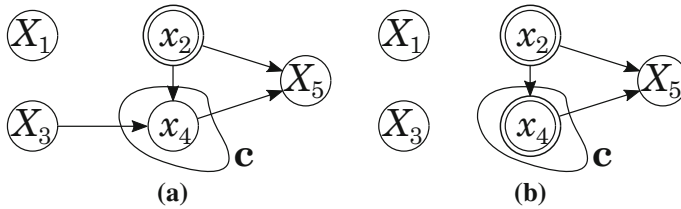
Fig. 3 Observing the context $\mathbf{C} = \mathbf{c}$ might produce a probability change for non-effects of the variable intervened on. This can, for example, happen if the context features common effects: Learning that X_4 has taken value x_4 by observation (**a**) might lead to a change in X_3's probability distribution even if x_2 is brought about by intervention. Not requiring that $\mathbf{C} = \mathbf{c}$ was brought about by intervention might, thus, lead to the false result that a disposition with stimulus condition $X_2 = x_2$ and manifestation $X_3 = x_3$ is present. Fixing the context by intervention (**b**), however, breaks the arrows into variables in the context and, thus, excludes such problematic consequences

the causal Markov condition,[14] it does not rely on the causal faithfulness condition. As mentioned before, the causal faithfulness condition only becomes relevant when employing search procedures. Secondly, though Definition 3.3 uses causal models for analyzing dispositions, dispositions are not identified with specific causal paths represented in particular causal models; the notion of a disposition is not model-relative.[15] Thirdly, our analysis does not directly apply to single objects u (without further specification), but rather to objects u of a certain domain U sharing relevant characteristic marks. U can, for example, contain all objects u made of porcelain. According to our analysis, objects of type U would then be fragile (if struck) because in any subclass U' of U made up of objects being struck (by intervention), the relative frequency of objects breaking would be higher than in U.[16] This restriction is related

---

[14] We assume that for any disposition there will be a suitable model satisfying the causal Markov condition. That any set of variables can be expanded in such a way that the causal Markov condition is satisfied is a common assumption made in the causal modeling literature (see, e.g., Spirtes et al. 2000, ch. 6). Anyway, one might still be worried about very specific cases which seem to violate the causal Markov condition because of common causes that do not screen off their effects. Prominent examples include EPR/B experiments, decay processes, but also macro pheomena such as the ones produced by Cartwright's (1999a,b) chemical factory. Whether such causal dependencies really exist is still controversial and several proposals have been made how the causal Markov condition could be saved (for contributions to this debate see, e.g., Glymour 1999; Näger 2016; Retzlaff 2017; Wood and Spekkens 2015; Hausman and Woodward 1999). But even if one is not convinced by these proposals and takes violations of the causal Markov condition seriously, there are ways to get the causal Bayes net machinery working (see, e.g., Gebharter and Retzlaff 2018; Näger 2013; Schurz 2017).

[15] Our analysis of dispositions is in the same sense not model-relative as, for example, Woodward's (2008, Sect. 7) analysis of causation simpliciter.

[16] The detour over such a domain U is required because the causal Bayes net machinery can (without further assumptions) not capture causal relations between token-level events involving single objects u. It can, for example, not capture the token-level causal claim that striking this particular thing u caused its breaking. But it can establish the type-level claim that striking objects u made of porcelain (i.e., objects u in the domain U) is causally relevant for breaking. So, strictly speaking, what our analysis can tell us is that objects u with a certain property U (such as being made of porcelain) are fragile (if struck). This seems not too far away from how we actually ascribe dispositions. If we say things like "glass is fragile" or "this particular glass is fragile", for example, we implicitly specify U as the class of things made of glass. Since the relevant class U is at least implicitly specified in all the examples we discuss, we will most of the time just speak of u or objects u without explicitly mentioning the relevant domain U.

to the fourth point we would like to mention here: Our analysis can be supplemented by an account of actual causation (see, e.g., Halpern and Pearl 2005; Woodward 2003, Sect. 2.7).[17] Accounts of actual causation are intended to specify the actual causes of a specific event given all other possibly causally relevant variables of interest took the values they actually took. But the move from type-level causal claims expressed within a causal model to actual causation comes with specific problems and which account gets things right or is the best one is still somewhat controversial (see, e.g., Glymour et al. 2010). Because of this and since our account is simpler, we stick with it in this paper. However, readers who would like to avoid the detour over U and who are more sympathetic to accounts of actual causation might replace conditions 3 and 4 above by requiring that $X_1 = x_1, \ldots, X_n = x_n$ would be actual causes of $Y = y$ in context $\mathbf{C} = \mathbf{c}$ according to their favourite account of actual causation. But note that most of these accounts are designed for deterministic settings only; an exception is Fenton-Glynn (2017).

3.3 Dispositions and conditionals

We have not defined dispositions by reference to conditionals about the disposition's stimulus-manifestation pairs $\langle \Sigma, M \rangle$ the way classical approaches would. Rather, dispositions are identified with cause-effect relations that, in some circumstances (i.e., contexts), give rise to the counterfactual conditionals used in classical analyses. Counterfactual conditionals are typically connected to causal structures via the notion of an intervention (cf. Pearl 2000, Sect. 7). To put it in a nutshell, counterfactuals such as "if X_i had taken value x_i, then X_j had taken value x_j with probability r" can be evaluated by checking whether $P(x_j|do(x_i))$ equals r.[18] As we will see below, changing a disposition's context might change the probability distribution over the set of variables of interest and, hence, also the counterfactuals a disposition gives rise to. We consider this the main advantage our approach has over conditional analyses: It allows one and the same disposition to support different counterfactual conditionals in different causal environments. In case of the structure in Fig. 2, for example, setting X_1 to its value x_1 by intervention relative to the background of the empty context $\mathbf{C} = \emptyset$ (this is (a)) will not have any probabilistic effect on X_4, i.e., $P(x_4|do(x_1)) = P(x_4)$ will hold for all X_4-values x_4. But if $X_3 = x_3$ is added to the context, then bringing about x_1 by intervention might have an effect on X_4, i.e., $P(x_4|do(x_1, x_3)) > P(x_4|do(x_3))$ might hold for some X_4-values x_4.

3.4 Dispositions and causal contexts

Let us come back to our simple water-solubility example in order to illustrate how a dispositions' behavior can change with change in contexts. According to our analysis, u has the disposition of water-solubility if whether u is put into water (W) and whether u dissolves (D) stand in the right causal relationship and putting u into water (by

---

[17] We are indebted to an anonymous reviewer for this point.

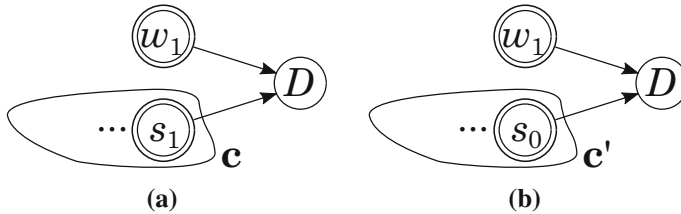[18] We can cover the strict or non-probabilistic case by setting r to 1 or 0.

Fig. 4 Changing contexts from (**a**) $\mathbf{C} = \mathbf{c}$ to (**b**) $\mathbf{C} = \mathbf{c}'$ will change the impact of bringing about $W = 1$ by intervention has on D. w_1 is shorthand for $W = 1$; likewise for s_i. The dots stand for additional variables which are assumed to be fixed to the same values in both contexts

intervention) actually increases the probability of u dissolving in some context $\mathbf{C} = \mathbf{c}$. Assume that we already know that W is a cause of D and that this causal relation can be represented by $W \longrightarrow D$. Now, the context $\mathbf{C} = \mathbf{c}$ might, for example, include the specific gravitational force of the earth, the air pressure, etc. However, it might also include the fact that u has been shrink-wrapped ($S = 1$). In this case, intervening on W will have no probabilistic effect on D. But if we are looking at the alternative context $\mathbf{C} = \mathbf{c}'$, which only differs from $\mathbf{C} = \mathbf{c}$ insofar as $S = 1$ is replaced by $S = 0$, then putting u into water by intervention will increase u's probability to dissolve. Now, the handy thing about our analysis is that the disposition is present in both contexts, $\mathbf{C} = \mathbf{c}$ as well as $\mathbf{C} = \mathbf{c}'$. The difference is that it can manifest only in the latter context, i.e., when u is not shrink-wrapped (see Fig. 4 for a graphical illustration).

Note that one and the same disposition gives rise to different counterfactual conditionals in the two contexts. In context $\mathbf{C} = \mathbf{c}$, it gives, for example, rise to the conditional "if u were put into water, it would not dissolve". In context $\mathbf{C} = \mathbf{c}'$, on the other hand, it gives rise to the counterfactual "if u were put into water, it would dissolve". This observation shows what we think is the crucial advantage our approach has over classical conditional analyses. Such analyses have no other option than to identify the disposition with conditionals about its corresponding stimulus-manifestation pairs. But whether the stimulus leads to the manifestation of the disposition often depends on additional disposition-external causal factors such as whether u is shrink-wrapped.

At this point, one might object that classical analyses could circumvent the problem in the same way as our approach does, i.e., by requiring that the conditionals used to analyze a disposition have to hold only in some contexts. Unfortunately, such a move would be deemed to fail simply because there are contexts in which certain conditionals come out as true even though the objects of interest do not possess the corresponding disposition. These are the contexts which give rise to the typical mimicking cases (see, for example, the concrete block and sledgehammer scenario discussed in Sect. 4). Another possible way to go for classical analyses seems to consist in adding details about all additionally relevant factors to the antecedents of the conditionals used for the analysis. One could then compare all of these conditionals and the factors mentioned in their antecedents and exclude those factors that turn out to be irrelevant for the manifestation in all conditionals in which they appear. This would exclude problematic factors in mimicking cases (such as dropping in the concrete block and sledgehammer example discussed in Sect. 4) and finally allow for identifying the right stimulus

conditions. In the end, the procedure would give rise to a full specification of the disposition. However, it would also give rise to quite complex dispositions—every causally relevant factor would have to be one of the disposition's stimulus conditions—while, at the same time, one would not be able to account for non-complex dispositions such as solubility in water in a simple way. In addition, one would have to possess knowledge of all the causal factors that may possibly be relevant for a disposition's manifestation in order to define or specify this disposition.

In contrast, our approach can handle mimicking cases quite easily (see Sect. 4 for details). The key difference in how such cases are handled by the two approaches lies in the distinction between ordinary and intervention (or causal) conditionals. Conditional analyses will identify dispositions with ordinary conditionals and, thus, as we saw above, either postulate dispositions where there are none in the presence of mimickers or require a full specification of all the possibly relevant factors in the antecedents of the conditionals. The causal Bayes net approach, on the other hand, goes one step further: Instead of identifying dispositions with conditionals, it traces conditionals back to generic cause–effect relations. Since there is no disposition present in mimicking cases, the apparent stimulus conditions do not stand in such cause–effect relations to the manifestation and, thus, do not support intervention counterfactuals. Note that our approach does also not require full causal knowledge for identifying dispositions. We just need to find one causal model and causal context to whose background a disposition's manifestation can be brought about by bringing about its stimulus conditions by intervention (in many cases the empty context $\mathbf{C} = \emptyset$ will already allow us to identify a disposition). We can then investigate the disposition of interest's behavior in other causal environments on empirical grounds. Once a disposition is identified, we can add possible causal factors of interest to our model and set the corresponding variables to specific values. Having done that, we are able to measure the probability distribution over our expanded set of variables in the new context. This might show that the behavior of the disposition we are interested in changes with change of context. We consequently can discover on empirical grounds which counterfactuals the disposition in question may give rise to when changing contexts.

3.5 Finding dispositions

Here is another advantage of our analysis of dispositions: In principle, it allows for standard procedures of causal search to be applied to finding dispositions. To identify the dispositions of a certain system that can be described by the variables in some set of variables $\mathbf{V}$, we can just measure the probability distribution over $\mathbf{V}$ and run one of the established algorithms for causal search. If we have reasons to assume that the causal Markov condition and the causal faithfulness condition are satisfied, for example, we can run the prominent PC algorithm (Spirtes et al. 2000, Sect. 5.4.2).[19] Assume PC

---

[19] Of course, search procedures such as PC have their limitations. They do, for example, typically not output a unique causal structure (underdeterminiation) and additional experiments or background knowledge is required to further thin out the set of possible causal structures. There are also well-known possibilities how the causal Markov condition and the causal faithfulness condition can be violated. The causal Markov condition can, for example, be violated if the set $\mathbf{V}$ of variables to be analyzed is not causally sufficient, meaning that common causes of variables in $\mathbf{V}$ are not included in $\mathbf{V}$. And faithfulness can, for example, be

returns the structure $W \longrightarrow D \longleftarrow M$, where W and D describe, as before, whether u is put into water and whether u dissolves, respectively, while M stands for whether u is put into milk. On basis of our model we can determine whether putting u into water by intervention and whether putting u into milk by intervention increases the probability that u dissolves in some fixed context $\mathbf{C} = \mathbf{c}$. If so, we were able to identify two different dispositions: water-solubility and milk-solubility. One can further abstract from the stimulus conditions. The fact that u is water-soluble and milk-soluble implies that it is also soluble: There exist stimulus conditions $X_1 = x_1, \ldots, X_n = x_n$ such that bringing them about by intervention to the background of some fixed context would increase the probability for u to dissolve.

Note that which dispositions one will find on the basis of the causal Bayes net approach developed in this paper crucially depends on where one looks for them. If one is interested in everyday life entities such as tables, balls, pieces of sugar, etc., one will find dispositions such as solubility in water, solubility in milk, etc. If one restricts the domain of search to economics, biology, chemistry, physics, and so on, one will find quite different dispositions. We consider it as an advantage of our approach that it can be used in such a flexible way without making any realist commitments about the dispositions one finds; one can use the approach and be an ontological reductionist or a non-reductionist about the dispositions one finds within a certain domain.

3.6 Probabilistic and non-probabilistic dispositions

A major benefit of our suggestion to represent dispositions as cause–effect relations supporting intervention conditititals is that probabilistic and non-probabilistic dispositions can be handled in a unified way. We can—in principle—represent non-probabilistic dispositions as cause–effect structures in which the effect's parameters are 1 or 0. However, putting u into water (without any additional information about the background conditions) will typically not determine u to dissolve with probability 1 when u possesses the disposition of being water-soluble, for example. That is because there may be other (additional) influences on D, such as shrink-wrapping u, that may decouple D from the causal influence of W or, at least, make it less probable that $W = 1$ leads to $D = 1$.

A related advantage of a probabilistic analysis of dispositions is that it allows us to also capture the fact that many dispositions come in degrees (cf. Manley and Wasserman 2007, 2008). For instance, objects u may be more fragile when struck than other objects v. This can neatly be represented in a causal Bayes net via a causal arrow exiting a variable S (modeling whether an object is struck) and pointing to a variable B (modeling whether an object breaks). If the objects u in our domain are

---

Footnote 19 continued

violated in the presence of deterministic causal dependencies or in cases where different causal paths cancel each other out. For problems with deterministic structures and a more limited proposal how to handle them anyway see, for example, Glymour (2007). For a proposal how to detect violations of faithfulness and a procedure for causal search if faithfulness is violated see, for example, Zhang and Spirtes (2008). There are also ways to handle variable sets that are not causally sufficient. One prominent algorithm for causal search in such cases is, for example, FCI (Spirtes et al. 2000, pp. 144f).

made of very thin glass, $P(B = 1|do(S = 1))$ will be higher than it would be for objects v in a domain of objects made of thick glass, etc.

4 Handling masks, mimickers, and finks

In this section we show by means of prominent examples how our proposal to analyze dispositions as generic cause–effect relations that increase the probability of the manifestation if the stimulus conditions are brought about by intervention in some fixed context can handle masks, mimickers, and finks. Some of these examples are more realistic, while others are purely fictional. We use these examples as proxies for all kinds of possible scenarios involving masks, mimickers, and finks, as one can reasonably assume that the problems coming with classical analyses can be avoided in a similar way in other scenarios as well. We presuppose that all the relevant causal and probabilistic information required for constructing causal models is already available. This information could have been obtained in the way described in Sect. 3.5, for example.

Our diagnosis of why classical analyses of dispositions fail in the presence of masks, mimickers, and finks, is basically the following: Though the involved dispositions are the same in all contexts, their behaviors can change drastically when changing contexts. This characteristic of dispositions is not captured by standard analyses of dispositions that define dispositions in terms of their behavior. Our causal Bayes net analysis, however, goes one step further: We do not identify dispositions with conditionals describing their behavior, but with generic cause–effect relations that give rise to counterfactual conditionals. While the dispositions (i.e., the generic cause–effect relations) themselves remain the same, they give rise to different post intervention probabilities in different contexts, just as dispositions give rise to different counterfactual conditionals in different contexts. Some of the results of the analysis below will nicely fit our intuitions, while others will be more revisionary. In any case we think that our analysis in terms of causal Bayes nets makes interesting contributions to the present debate and allows for handling masks, mimickers, and finks in a fruitful way.

4.1 Fragility and packing material

One prominent example of a disposition that can be masked is fragility (if dropped). Glass, for example, possesses this disposition. The standard analysis of dispositions postulates that glass has the disposition of being fragile (if dropped) if and only if it holds that glass breaks if dropped. One major problem with this analysis becomes obvious when masks come into play. A mask, roughly speaking, prevents the manifestation of a disposition though the stimulus conditions of this disposition are satisfied. If some dispositions can be masked, the classical analysis fails. One possible mask for fragility of glass (if dropped) consists in adding an internal support structure.[20] If a

---

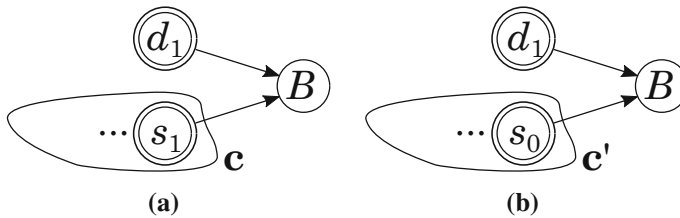[20] This is the original masking example from Johnston (1992).

Fig. 5 Context $\mathbf{C} = \mathbf{c}$ masks the disposition being fragile if dropped (**a**). However, when changing the context to $\mathbf{C} = \mathbf{c}'$ by intervention, dropping the glass increases the probability for breaking (**b**). d_1 is shorthand for $D = 1$; likewise for s_i

glass contains such a supporting structure, it—though still possessing the disposition of being fragile (if dropped)—will not break if dropped.

Let us now demonstrate how the analysis of dispositions we developed in Sect. 3 can avoid problems raised by masks. Let us analyze the disposition of being fragile (if dropped) by means of a causal arrow $D \longrightarrow B$, where D is a binary variable standing for whether u is dropped and B is a binary variable standing for whether u breaks. Since we are interested in how this disposition behaves in the presence of the mask, let us represent this mask by a binary variable S standing for whether internal supporting structure is placed within u or not. We add S as a direct cause of B and arrive at a causal Bayes net with the graph $D \longrightarrow B \longleftarrow S$.

As to the probability distribution, let us assume that dropping a glass increases the probability for it to break, i.e., that $P(B = 1|D = 1) > P(B = 1)$ holds. Yet in the case of an added support structure, let us assume that dropping a glass ceases to influence whether or not it breaks, i.e., that $P(b|d, S = 1) = P(b|S = 1)$ holds for all B- and D- values b and d, respectively. Together with $D \longrightarrow B \longleftarrow S$, these assumptions represent masking the disposition of fragility (if dropped) of glass via an added internal support structure.

Now our analysis avoids the problems the classical conditional analysis of dispositions faces at this point (see Fig. 5 for an illustration). In the presence of the mask dropping ($do(D = 1, S = 1)$) will make no difference for the probability of breaking ($B = 1$), i.e., $P(B = 1|do(D = 1, S = 1))$ will equal $P(B = 1|do(S = 1))$. According to our analysis, this means that the disposition breaking when dropped is not manifested. Though the disposition does not manifest in the presence of an internal supporting structure (i.e., when fixing S to 1 by intervention), u still possesses the disposition of being fragile (if dropped). Recall that we identified dispositions as cause–effect relations which propagate probabilistic influences brought about by interventions in some fixed context. Actually, there is such a context. If we change $S = 1$ to $S = 0$ in the original context $\mathbf{C} = \mathbf{c}$, we arrive at the new context $\mathbf{C} = \mathbf{c}'$. Bringing about $D = 1$ by intervention in context $\mathbf{C} = \mathbf{c}'$ will have a positive probabilistic impact on $B = 1$. Hence, our analysis yields the correct diagnosis of the situation: Glass actually has the disposition of being fragile, but it can manifest only if the glass is not filled with an internal supporting structure. Summarizing, we can identify masks as causal factors that cancel the causal influence of the stimulus conditions on the disposition's manifestation in some causal contexts.
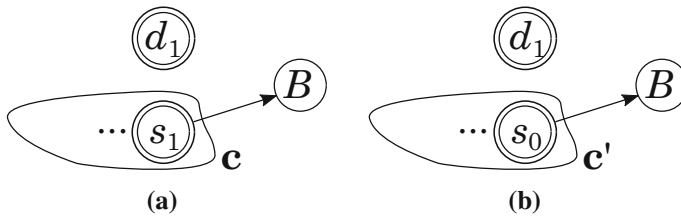
Fig. 6 Context $\mathbf{C} = \mathbf{c}$ mimicks the disposition being fragile if struck with a sledgehammer (**a**). Neither changing $\mathbf{C} = \mathbf{c}$ to a context $\mathbf{C} = \mathbf{c}'$ in which the concrete block is not struck with a sledgehammer (**b**), nor to any other context will make a difference: Dropping a concrete block will never increase the probability for breaking; concrete blocks do not possess the disposition fragile (if dropped). Note that, contrary to masking cases, the variable D representing the apparent stimulus condition is not a cause of B. d_1 is, again, shorthand for $D = 1$; likewise for s_i

4.2 Concrete blocks and sledgehammers

Suppose a concrete block is dropped and breaks because it is struck with a sledge-hammer just as it hits the floor.[21] If the concrete block had not been struck with the sledgehammer, it would still be unharmed after it was dropped. This is because concrete blocks do not possess the disposition of fragility (if dropped). Striking a concrete block with a sledgehammer, however, mimicks this disposition in this specific situation: Striking a concrete block right when it hits the floor after being dropped makes the concrete block behave as if it possessed the disposition of fragility (if dropped). This is fatal for the standard analysis of dispositions which would falsely tell us that concrete blocks are fragile (if dropped).

Let us now see how our suggestion of analyzing dispositions can handle concrete blocks and sledgehammers. We describe the relevant possible behaviors of the sledge-hammer (i. e., whether a concrete block u is struck with it) with a binary variable S. Whether the concrete block is dropped is modeled by a binary variable D, and whether the concrete block breaks by a binary variable B. We assume S to be directly causally relevant for B. But what about D's relevance for B? Since intervening on D will not make a difference for B in any circumstances (simply because D is not a cause of B at all), Definition 3.3 gives us the correct result that the concrete block does not possess the disposition of fragility (if dropped), no matter if the breaking of a concrete block might sometimes coincide with it being dropped (see Fig. 6 for an illustration). Summarizing, we can identify mimickers with causal factors that, if set to certain values, produce the same effect as the apparent stimulus conditions if u would actually have the disposition under consideration.

4.3 Hater of styrofoam

Let us come back to the case of the hater of styrofoam introduced in Sect. 2. This case has classically been discussed as a mimicker. Our analysis, however, can shed new light on the case. Whenever a styrofoam plate is struck, this produces an annoying

---

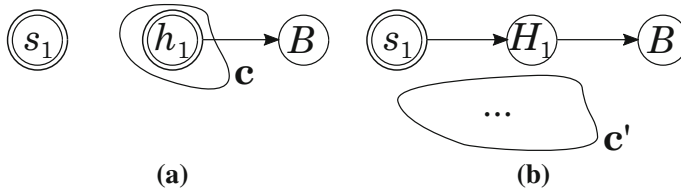[21] This example is taken from Manley and Wasserman (2008, p. 67).

Fig. 7 If the behavior of the hater of styrofoam is included in one's context (**a**), then no intervention on S will lead to a change in B. However, once H is removed from the context (**b**), setting S to 1 by intervention will increase the probability for $B = 1$. This supports the view that styrofoam is fragile (if struck)

sound. If a hater of styrofoam is around and hears that sound, he immediately rips the styrofoam plate in pieces. The basic idea underlying the case is that styrofoam does not have the disposition of being fragile (if struck), but that haters of styrofoam mimic this disposition and, thus, that conditional analyses lead to the wrong result that styrofoam is fragile (if struck).

Applying our causal Bayes net analysis to this case has an interesting consequence. Since the hater's behavior is a reaction to whether the plate of styrofoam is struck, the example's underlying causal structure can be analyzed as $S \longrightarrow H \longrightarrow B$ with $S = 1/0$ for the plate being struck/not struck, $H = 1/0$ for the hater ripping the plate apart/not ripping the plate apart, and $B = 1/0$ for breaking/not breaking. If the context includes the hater of styrofoam, then intervening on S will not have any effect on B (Fig. 7a). However, if the context is changed in such a way that it does not contain $H = 1/0$ anymore, striking the styrofoam plate will most likely cause the hater to rip it apart and, thus, bringing $S = 1$ about by intervention will increase the probability of $B = 1$ (Fig. 7b). As a consequence, our analysis tells us that styrofoam plates actually are fragile (if struck). However, this disposition only manifests in very specific circumstances such as in the presence of a hater of styrofoam.

At first glance, this result seems to go against our intuitions. However, the whole example is somewhat exotic. In our view the result shows two things: Firstly, that purely fictional examples and playing around with possibilities can be dangerous. Our analysis can help to ground the debate about dispositions in empirical facts about the structure of the actual world. Secondly, it shows that the hater of styrofoam case is not really a mimicking case. The hater of styrofoam does not mimic the disposition fragile (if struck), but rather mediates between the disposition's stimulus condition (being struck) and its manifestation (breaking). The hater is rather a kind of enabler.

4.4 Electro-finks

Yet another problem for standard analyses of dispositions is the possible existence of finks. A fink, roughly speaking, takes away a disposition of an object that it had before or vice versa. The most prominent examples of finks are probably Martin's (1994) electro-fink devices already mentioned in Sect. 2. The object of interest in Martin's example is a dead wire, i.e., a wire that does not have the disposition of being conductive. The wire is, however, connected to an electro-fink device. This device renders the wire conductive while the wire is touched (and only while the wire
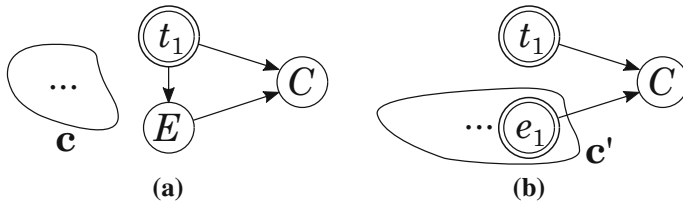
Fig. 8 Electro-finks work in the same way as masks. If E is not included in context $\mathbf{C} = \mathbf{c}$ (**a**), it cancels T's causal influence on C. Once $\mathbf{C} = \mathbf{c}$ is changed to $\mathbf{C}' = \mathbf{c}'$ by adding $E = 1$ (**b**), bringing $T = 1$ about by intervention increases the probability of $C = 1$. t_1 is shorthand for $T = 1$; likewise for e_1

is touched). The problem for classical analyses of dispositions is that the wire would conduct electricity if it were touched though it does not seem to have the disposition of being conductive at the moment it is touched. Recall from Sect. 2 that the electro-fink can also work in a reverse-cycle. In this version, the wire is originally conductive, but when it is touched, the electro-fink device is assumed to take away the wire's disposition of being conductive. Hence, the wire will not conduct electricity when touched though, according to Martin (1994), it possesses the disposition of being conductive.

Let us now reconstruct the first electro-fink scenario as a causal Bayes net. Assume E to be a binary variable standing for whether the electro-fink device is on, T a binary variable standing for whether a wire u is touched, and C a binary variable for whether u conducts electricity. The causal structure underlying this scenario is the concatenation of $T \longrightarrow C \longleftarrow E$ and $T \longrightarrow E$. The probability distribution of our causal Bayes net has to satisfy the following constraints in order to adequately represent Martin's (1994) electro-fink scenario: $P(C = 1 | E = 1, T = 1) > P(C = 1 | E = 1)$ holds and $P(c | E = 0, t) = P(c | E = 0)$ holds for all C- and T-values c and t, respectively.

Our analysis sheds new light on dispositions in the presence of finks (see Fig. 8 for an illustration): It leads to the result that the wire actually possesses the disposition of being conductive if touched. This contradicts Martin's (1994) assumption that the wire does not possess the disposition of being conductive. The way we analyze it, the wire does indeed have this disposition, yet it manifests only under very special conditions such as the presence of an electro-fink device. Note that electro-finks could also be described according to the general characterization of masks we provided earlier: Masks are causal factors that cancel the causal influence of the stimulus conditions on the disposition's manifestation in some causal contexts. In the electro-fink case this context is the empty context or a context that does not contain E. What E does, if not controlled for, is to detect whether the wire is touched and to cancel T's causal influence on C accordingly. However, once E is controlled for and $E = 1$ is added to one's context, then touching the wire will increase the probability of $C = 1$. This finding supports the view that finks are actually certain kinds of masks. It also fits Bird's (2007) observation that "the dividing line between finkishness and antidotes is not clearly perceptible, or that there is an overlap" (p. 32).[22]

---

[22] Bird (2007) uses the term "antidotes" in the same sense as we use the term "masks".

Let us finally see what can be said about the reverse-cycle electro-fink case. The graph of a causal Bayes net representing this scenario is also the concatenation of $T \longrightarrow C \longleftarrow E$ and $T \longrightarrow E$. In the reverse-cycle electro-fink case, of course, the constraints on our probability distribution must be different: We have now to expect that $P(C = 1 | E = 0, T = 1) > P(C = 1 | E = 0)$ holds and that $P(c | E = 1, t) = P(c | E = 1)$ holds for all C- and T-values c and t, respectively. The first means that touching the wire increases the probability for the wire to conduct electricity if the electro-fink is off, and the second that touching the wire does not have any probabilistic influence on whether or not it conducts electricity if the electro-fink device is active. Because of the first, forcing T to take its value 1 by intervention will increase the probability of $C = 1$ when fixing E to 0. Hence, our Definition 3.3 yields the same result as in the non-reverse case: The wire actually possesses the disposition of being conductive (if touched), but this disposition manifests only under very special conditions such as the presence of an active reverse-cycle electro-fink device. Note that the reverse-cycle electro-fink case is structurally identical to the non-reverse case. Thus, according to our analysis, the reverse-cycle electro-fink device strictly speaking is also some kind of mask.

5 Conclusion

In this paper we developed an analysis of dispositions in terms of causal Bayes nets (Sect. 3). In particular, we proposed to analyze dispositions as generic cause–effect relations such that bringing about the stimulus conditions by intervention increases the probability of the manifestation to the background of some fixed context. Our causal Bayes net analysis succeeds in the difficult task of capturing the stability of dispositions in different contexts, while simultaneously allowing for the required variability of their behavior in different contexts. It provides empirically informed epistemic access to dispositions and is conditional-friendly in so far as it can be used to compute probabilities for counterfactual conditionals. We have then applied our causal Bayes net analysis of dispositions to long-standing problems within the dispositions debate (Sect. 4), namely problems with masks, mimickers, and finks. In masking and mimicking cases, the conditionals used to analyze a disposition by classical accounts do not coincide with the actual presence of the disposition in question. Hence, classical analyses might lead to false results about whether dispositions are present. Our account, on the other hand, can handle masks and mimickers and, at the same time, explain why conditional analyses fall victim to masks and mimickers: In certain contexts dispositions do simply not give rise to the right conditionals. Depending on the specific causal context, one and the same disposition might support quite different conditionals. It is because of this that dispositions cannot be adequately analyzed on the basis of conditionals describing their behavior.

Another interesting result of the paper is that the hater of styrofoam is, contrary to how it is typically discussed in the literature, not a mimicker. It rather causally mediates between the stimulus being scratched and the manifestation breaking. It also turned out that finks are—in accordance with Bird's view (2007)—some kind of masks. It seems that finks do not actually remove or produce an object u's disposition D as Martin

([1994](#)) supposes. The causal Bayes net analysis tells us that the wire in the electro-fink scenario had the disposition of conductivity all along. If our analysis is correct, the presence of some device can neither take away nor generate the wire's disposition of being conductive. What such a device can do is rather to causally interfere with the wire in such a way that the disposition's manifestation does or does not come about. The bouncy ball briefly mentioned in Sect. [2](#), for example, can be seen in a similar light: Deep freezing the ball does not remove the disposition of elasticity and replace it by the disposition fragility. The bouncy ball possesses both dispositions at the same time, but each of them shows itself only in the right circumstances: The ball's elasticity manifests only if the temperature is high enough, and its fragility manifests only if the temperature is low enough. Having both dispositions is essential to the bouncy ball—it seems reasonable to assume that it has these dispositions because of the very nature of the material it is made of; this material does not change with change in temperatures.[23]

We recognize that some of the conclusions our analysis leads to (e.g., that finks are nothing over and above masks or that the hater of styrofoam does not really mimic the disposition fragile if struck) can still be seen as controversial. However, we hope to have shown successfully that causal Bayes nets provide an interesting and promising new way of approaching the debate about dispositions. Even if there may be disagreement about the specific way how to capture certain cases, we are optimistic that the framework of causal Bayes nets can help to frame the debate more sharply and can help to ground it in empirical facts. It would also be interesting to bring our approach in contact with other theories of dispositions which we have not discussed in this paper. But this is something to be done in future work.

References

Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Oxford: Oxford University Press.
Carroll, J. W. (1994). *Laws of nature*. Cambridge: Cambridge University Press.
Cartwright, N. (1999a). Causal diversity and the Markov condition. *Synthese*, *121*(1/2), 3–27.
Cartwright, N. (1999b). *The dappled world*. Cambridge: Cambridge University Press.
Choi, S. (2006). The simple vs. reformed conditional analysis of dispositions. *Synthese*, *148*(2), 369–379.
Cross, T. (2012). Recent work on dispositions. *Analysis*, *72*(1), 115–124.

---

[23] Note that this line of thoughts is intended as an independent motivation for why it makes sense that an object can have seemingly incompatible dispositions at the same time. It is not intended as part of the analysis provided in this paper.

Ellis, B. (2001). *Scientific essentialism*. Cambridge: Cambridge University Press.

Ellis, B., & Lierse, C. (1994). Dispositional essentialism. *Australasian Journal of Philosophy*, *72*(1), 27–45.

Fenton-Glynn, L. (2017). A proposed probabilistic extension of the Halpern and Pearl definition of 'actual cause'. *British Journal for the Philosophy of Science*, *68*(4), 1061–1124.

Fischer, F. (2018). *Natural laws as dispositions*. Berlin: De Gruyter.

Gebharter, A. (2017a). Causal exclusion and causal Bayes nets. *Philosophy and Phenomenological Research*, *95*(2), 153–375.

Gebharter, A. (2017b). *Causal nets, interventionism, and mechanisms: Philosophical foundations and applications*. Cham: Springer.

Gebharter, A. (2017c). Uncovering constitutive relevance relations in mechanisms. *Philosophical Studies*, *174*(11), 2645–2666.

Gebharter, A., & Retzlaff, N. (2018). A new proposal how to handle counterexamples to Markov causation à la Cartwright, or: fixing the chemical factory. *Synthese*. https://doi.org/10.1007/s11229-018-02014-7.

Glymour, C. (1999). Rabbit hunting. *Synthese*, *121*(1/2), 55–78.

Glymour, C. (2004). Critical notice. *British Journal for the Philosophy of Science*, *55*(4), 779–790.

Glymour, C. (2007). Learning the structure of deterministic systems. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 231–240). New York: Oxford University Press.

Glymour, C., Danks, D., Glymour, B., Eberhardt, F., Ramsey, J., Scheines, R., et al. (2010). Actual causation: A stone soup essay. *Synthese*, *175*(2), 169–192.

Goodman, N. (1983). *Fact, fiction, and forecast* (4th ed.). Cambridge: Harvard University Press.

Halpern, J. Y., & Pearl, J. (2005). Causes and explanations: A structural-model approach. Part I: Causes. *British Journal for the Philosophy of Science, 56*(4), 843–887.

Hausman, D., & Woodward, J. (1999). Independence, invariance and the causal Markov condition. *British Journal for the Philosophy of Science*, *50*(4), 521–583.

Hitchcock, C. (2016). Conditioning, intervening, and decision. *Synthese*, *193*(4), 1157–1176.

Johnston, M. (1992). How to speak of the colors. *Philosophical Studies*, *68*(3), 221–263.

Lewis, D. (1997). Finkish dispositions. *Philosophical Quarterly*, *47*(187), 143–158.

Malzkorn, W. (2000). Realism, functionalism and the conditional analysis of dispositions. *Philosophical Quarterly*, *50*(201), 452–469.

Manley, D., & Wasserman, R. (2007). A gradable approach to dispositions. *Philosophical Quarterly*, *57*(226), 68–75.

Manley, D., & Wasserman, R. (2008). On linking dispositions and conditionals. *Mind*, *117*(465), 59–84.

Martin, C. B. (1994). Dispositions and conditionals. *Philosophical Quarterly*, *44*(174), 1–8.

Meek, C., & Glymour, C. (1994). Conditioning and intervening. *British Journal for the Philosophy of Science*, *45*(4), 1001–1021.

Näger, P. M. (2013). Causal graphs for EPR experiments. http://philsci-archive.pitt.edu/id/eprint/9915. Retrieved on January 9, 2019

Näger, P. M. (2016). The causal problem of entanglement. *Synthese, 193*(4), 1127–1155.

Pearl, J. (2000). *Causality* (1st ed.). Cambridge: Cambridge University Press.

Popper, K. (1990). *A world of propensities*. Bristol: Thoemmes.

Prior, E. (1985). *Dispositions*. Aberdeen: Aberdeen University Press.

Prior, E., Pargetter, R., & Jackson, F. (1982). Three theses about dispositions. *American Philosophical Quarterly*, *19*(3), 251–257.

Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.

Retzlaff, N. (2017). Another counterexample to Markov causation from quantum mechanics: Single photon experiments and the Mach–Zehnder interferometer. *Kriterion*, *31*(2), 17–42.

Ryle, G. (1949). *The concept of mind*. London: Hutchinson and Co.

Schaffer, J. (2016). Grounding in the image of causation. *Philosophical Studies*, *173*(1), 49–100.

Schrenk, M. (2010). The powerlessness of necessity. *Noûs*, *44*(4), 725–739.

Schurz, G. (2017). Interactive causes: Revising the Markov condition. *Philosophy of Science*, *84*(3), 456–479.

Schurz, G., & Gebharter, A. (2016). Causality as a theoretical concept: Explanatory warrant and empirical content of the theory of causal nets. *Synthese*, *193*(4), 1073–1103.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search* (1st ed.). Dordrecht: Springer.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge: MIT Press.

Vetter, B. (2013). Multitrack dispositions. *Philosophical Quarterly*, *63*(251), 330–352.

Wood, C. J., & Spekkens, R. W. (2015). The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning. *New Journal of Physics, 17*, 1–29.

Woodward, J. (2003). *Making things happen*. Oxford: Oxford University Press.

Woodward, J. (2008). Response to Strevens. *Philosophy and Phenomenological Research*, *77*(1), 193–212.

Zhang, J., & Spirtes, P. (2008). Detection of unfaithfulness and robust causal inference. *Minds and Machines*, *18*(2), 239–271.