

Notes on inflation technique adapted to quantum and GPT causal models

TF, EW, RWS

What we said in the inflation paper

I. CAUSAL INFERENCE IN QUANTUM THEORY AND IN GENERALIZED PROBABILISTIC THEORIES

Recent work has sought to explore quantum generalizations of the notion of a causal model, termed *quantum causal models* [citations]. The causal structures are still represented by DAGs, but (i) whereas classically each latent node of the DAG represents a random variable, quantumly, each latent node represents a quantum system associated to a complex Hilbert space, and (ii) whereas classically the manner in which a node depends causally on its parents is represented by a conditional probability distribution, quantumly, it is represented by a completely-positive trace-preserving linear map. Note, however, that a quantum causal model is still ultimately in the service of explaining joint distributions over classical variables. These variables represent the settings of preparation procedures and the outcomes of measurements that are used in an experiment on quantum systems, and the statistical distribution over such variables is the only experimental data with which one can confront a given quantum causal model. The basic problem of causal inference for quantum causal models, therefore, concerns the compatibility of a joint distribution over observed classical variables with a given DAG when the model supplementing the DAG is quantum (that is, includes latent nodes that are quantum). If a joint distribution over observed variables is compatible with a given DAG within the framework of quantum causal models, we say that it is *quantumly compatible* with that DAG.

One motivation for studying quantum causal models is that they offer a new perspective on an old problem in the foundations of quantum theory: that of establishing precisely which of the principles of classical physics must be abandoned in quantum physics. It was shown in [?] that Bell’s theorem [?] can be leveraged to prove that there exist distributions over observed variables (predicted by quantum theory and observed experimentally [citations]) that cannot be accounted for in the Bell scenario under any classical causal model without doing violence to fundamental principles of causal inference such as Reichenbach’s principle and the principle that conditional independences should not be fine-tuned (the principle of faithfulness). Quantum causal models, on the other hand, *can* account for these distributions in the Bell scenario. In other words, although these distributions are not *classically* compatible with the Bell scenario, they are *quantumly* compatible with it. Furthermore, such quantum causal explanations suggest [cite:LeiferSpekkens, ParadigmKinematicsAndDynamics, SomethingByHenson?] that quantum theory is perhaps best understood as revising our notions of the nature of unobserved entities, how one represents causal dependences thereon and incomplete knowledge thereof, while nonetheless *preserving* the spirit of Reichenbach’s principle and the principle of no fine-tuning.

Another motivation for studying quantum causal models is a practical one. Violations of Bell inequalities have been shown to constitute resources for information processing [citations]. To achieve such an advantage, however, it is necessary that a given information processing task can be recast in such a way that the causal structure of the protocol mirrors that of the Bell scenario. Many tasks, however, may not be amenable to being cast in this form. And yet the Bell scenario is not the only DAG for which there is a quantum-classical separation, that is, for which there exist distributions that are quantumly but not classically compatible with the DAG. It has been shown that such a separation also exists in the bilocality scenario [Branciardetal][?] and the triangle scenario [?], and it is likely that there will be many more DAGs for which a quantum-classical separation can be found. The hope, therefore, is that for any DAG where one can identify a quantum-classical separation, the separating distributions may constitute a resource for information processing.

For both foundational and practical reasons, therefore, there is a strong motivation to find examples of DAGs that exhibit a quantum-classical separation. However, this is by no means an easy task. The set of distributions that are quantumly compatible with a given DAG is actually very similar to the set of distributions that are classically compatible with that DAG [? ?]. For example, both the classical and quantum sets respect all the conditional independence relations among observed nodes that are implied by the d -separation relations of the DAG [?]. Furthermore, recent work has found that set of distributions that are quantumly compatible with a given DAG satisfy many of the entropic inequalities that hold classically [? ? ?]. To date, no quantum-classical separation has been identified where the separation is achieved by a Shannon-type entropic inequality on observed variables that is derived from the Markov conditions on all nodes [? ?]. Fine-graining the scenario by conditioning on root variables (“settings”) leads to a different kind of entropic inequality, and these have proven somewhat quantum-sensitive [? ? ?]. Such inequalities are still limited, however, in that they only apply ~~in the presence of observable root nodes to DAGs that include root nodes that are observed~~¹, and they still fail to witness cases of classical incompatibility of certain distributions with a given DAG (where the incompatibility is witnessed by stronger techniques) [? ?].

¹ Rafael Chaves and E.W. are exploring the potential of entropic analysis based on ~~distributions that are~~ conditioned on non-root observed nodes. This generalizes the method of entropic inequalities, and might be capable of providing much stronger entropic witnesses.

In addition to quantum generalizations of causal models, one can define generalizations for other operational theories that are neither classical nor quantum, as was done in Henson, Lal and Pusey [?]. Such generalizations are formalized using the framework of *generalized probabilistic theories* (GPTs) [citations], a framework that is sufficiently general to describe any operational theory that makes statistical predictions about the outcomes of experiments and respects some weak constraints. Some constraints on compatibility can be proven to apply not only to classical and quantum causal models, but to any theory expressible in this framework. In the terminology of Ref. [?], such constraints on compatibility constitute *theory-independent* limits on correlations. These are of interest because they clarify what any conceivable theory of physics can achieve in a given causal scenario.

Furthermore, by knowing which constraints hold in the GPT framework, one can seek to identify DAGs that exhibit a GPT-classical separation or a GPT-quantum separation, that is, to identify DAGs for which there exist distributions that are GPT compatible but not classically compatible (respectively not quantumly compatible) with the DAG.

Consider GPT-classical separation first. Many examples of DAGs exhibiting such a separation were provided in Ref. [?]. Of the classes of DAGs with six or fewer nodes identified therein as having the potential to be interesting (by not satisfying a sufficient condition for being uninteresting), all but three classes were demonstrated to indeed be interesting. Pienaar recently demonstrated that the three remaining classes are interesting as well [citation]. In Appendix ??, we show how the inflation technique provides an alternative means of demonstrating the interestingness of one of these three classes.

GPT-quantum separations are harder to find. The Bell scenario is known to manifest such a separation, through the work of Tsirelson and Popescu and Rohrlich [citations]. The identification of such stronger-than-quantum correlations in the Bell scenario has been a focus of much foundational research in recent years. Traditionally the foundational question has always been: why does quantum theory predict correlations that are *stronger* than one would expect classically? But now there is a new question being asked: why does quantum theory predict correlations that are *weaker* than those predicted by other GPTs? There has been some interesting progress in identifying physical principles that can pick out the precise degree of correlations that are exhibited by quantum theory [cite:informationcausality, macroscopiclocality, etcetera]. Further opportunities for identifying such principles would be useful. This motivates a classification of DAGs into those which have a quantum-classical separation, those which have a GPT-quantum separation and those which have both.

The difference between classical, quantum and GPT causal models is the manner in which latent nodes are represented as well as the manner in which the causal influence of a latent node on other nodes is represented. (It shares, however, the feature that the entitites that represent such causal influences mirror the causal structure—for instance, in all cases, the state of a pair of root nodes in the DAG is presumed to factorize.)

Henson, Lal and Pusey[?] were able to identify constraints on compatibility that hold for *any* GPT by avoiding making reference to the latent nodes of the DAG (which are treated differently in different GPTs) and only making reference to the observed variables in the DAG.

The inflation technique also only makes reference to the observed nodes in a DAG. In some cases, the inequalities one thereby derives hold for the GPT notion of compatibility. Indeed, the inequality for the triangle scenario that we presented in Eq. ?? is the probabilistic version of the one that Henson Lal and Pusey derive in their article, and—as we already noted—in this case, their proof technique can be understood as an instance of the inflation technique.

Nonetheless, the inflation technique also yields inequalities that hold only for the *classical* notion of compatibility. It is for this reason that it can be used to derive inequalities that are quantumly violated. For instance, in Sec. ?? and Appendix ??, we have shown how the inflation technique can be used to derive Bell inequalities, which clearly admit of quantum violations. The inflation technique can also be used to derive inequalities for the Triangle scenario that are quantumly violated. Specifically, our technique is able to reproduce the result shown by Fritz [citation], that a particular distribution, inspired by the Bell example, is quantumly but not classically compatible with the triangle scenario. [R: Elie: should we say this here or leave it for some putative future paper with TC?]

It is possible that one could generalize the inflation technique to derive inequalities that can achieve a GPT-quantum separation. This, however, requires a careful definition of a quantum causal model and is beyond the scope of this article. Nonetheless, we will note below in which inflation in the quantum context will differ from its classical counterpart.

Our main focus here, however, will be to explain what distinguishes applications of the inflation technique that yield inequalities for GPT compatibility from those that yield inequalities for classical compatibility. The distinction rests on a structural feature of the inflated DAG, which we now define.

Definition 1. A DAG $G' \in \text{Inflations}[G]$, is said to contain an *inflationary fan-out* if it contains a latent node that has two or more children that are copy-index equivalent.

The inflations of the triangle DAG that are depicted in ?? and ?? contain one or more inflationary fan-outs, as does the inflation of the Bell DAG that is depicted in ??. On the other hand, the simplest inflation of the triangle DAG that we consider in this article, that of ??, does not contain any inflationary fan-outs.

In the classical context, for every inflationary fan-out, all of the copy-index-equivalent children of the latent node are required to causally depend on the latent node in precisely the same way as their counterparts in the original DAG did. However, the only maps that can achieve such a duplication of dependences, known as *broadcasting maps*, are not physically realizable, as was famously shown for quantum theory in Ref. [citation Barnum et al.] and generalized to GPTs in Ref. [Lefier, Barnum, Barrett, Wilce]. These no-broadcasting results are related to the monogamy of entanglement in quantum theory [cite:monogamypapers] (see also the discussion of the quantum conditional problem in Ref. [LeiferSpekkens]). Mathematically, there *are* linear maps that can achieve broadcasting, as shown in ??, but if one used these to define the inflation of a quantum causal model, then distributions over certain sets of variables in the inflated DAG (those containing the children of an inflationary fan-out) might fail to be nonnegative. Such inflations might still be useful as a mathematical tool for ultimately deriving causal compatibility inequalities on the original DAG, but one would need to proceed differently from the way we have proceeded in this article: whereas we have here assumed that all joint valuations of the set of observed variables on the inflated DAG are nonnegative, one could not impose such a constraint for the type of quantum inflation just described. Rather than demanding nonnegativity of the full joint distribution, one could only demand nonnegativity for distributions on sets of variables that did not include multiple children of an inflationary fan-out. Note that the inflated DAG in such cases could not be interpreted as a causal structure. Rather, it would be interpreted as describing multiple different *counterfactual* scenarios within which the causal dependences are the same.

An analysis along these lines has already been carried out successfully by ?] in the derivation of entropic inequalities that capture quantum compatibility. Although ?] do not invoke the inflation technique, they do seem to employ a similar type of structure to model the conditioning of an observed variable on a “setting” variable, a structure that we would describe as an inflated DAG that has no inflationary fan-outs. ?] take pains to avoid including full joint probability distributions in any of the entropic inequalities they apply to this structure R: *why must they take pain to do so when the DAG has no inflationary fan-out?*, precisely as we would want to do in constructing inequalities on our inflated DAG, and they successfully derive entropic inequalities for quantum compatibility. But so far, no inequalities polynomial in the probabilities have been derived using this method.

II. FLOTSAM FROM THE INFLATION PAPER

The reason that the inflation technique can see the difference is that it does not confine its attention to the observed nodes of the original DAG, but considers also the observed nodes of the inflated DAG. Regardless of the theory one considers, ancestral independence always yields factorization of probabilities.

The assumption that two sets of variables in the inflated DAG that are copy-index equivalent can both causally depend on their parents in the same way that they could have in the original DAG *does not hold* for broadcasting DAGs, because of the quantum and GPT no-broadcasting theorems.

Suppose that quantum systems B and C both depend causally on quantum system A . The two dependences are represented by quantum channels from A to B and from A to C respectively. However, not every pair of channels is permitted. A well-known result known as the *quantum no-broadcasting theorem* establishes that the two channels cannot both be the identity channel. In fact, if one of the channels is the identity channel, then the other is a so-called deletion channel wherein the output does not depend at all on the input. (The latter result is equivalent, through the Choi-isomorphism, to the fact that pure entanglement is monogamous.) Even if neither channel is the identity channel, the amount of quantum information transmitted down one channel necessarily trades off with the amount transmitted down the other. [Are there good references for this?]

[My stuff w Jon Barrett may be considered a solution of the quantum conditionals problem when the conditioning system is a complete common cause. In this case, the condition is simply commutativity of the quantum conditionals.]

In a quantum causal model, a specification of the causal structure includes a specification of the dimension of Hilbert space of the quantum system associated to each node, and a specification of the set of causal parameters specifies, for each node, the quantum channel from the tensor product of the operator spaces associated to the node’s parents to the operator space of the node. In the case of root nodes, the parents are the null set so that the quantum channel is simply a preparation of a quantum state. We will denote a quantum channel from a system X to a system Y as $\mathcal{E}_{Y|X} : \mathcal{L}(\mathcal{H}_X) \rightarrow \mathcal{L}(\mathcal{H}_Y)$. Therefore, a given set of causal parameters has the form

$$\{\mathcal{E}_{A|\text{Pa}_G(A)} : A \in \text{Nodes}[G]\}. \quad (1)$$

A quantum causal model M constitutes a causal structure together with a set of causal parameters, $M := (G, \{\mathcal{E}_{A|\text{Pa}_G(A)} : A \in \text{Nodes}[G]\})$.

A quantum causal model specifies a joint distribution over all observed variables in the DAG by tracing over the

latent systems,

$$P_{\text{ObservedNodes}[G]} = \text{tr}_{\text{LatentNodes}[G]} \bigotimes_{A \in \text{Nodes}[G]} \mathcal{E}_A | \text{Pa}_G(A), \quad (2)$$

where the tensor product represents Hardy’s causaloid product. Actually, there is a constraint on what quantum channels can appear here. For any two nodes, if we consider the complete common cause of the two, then the quantum channels from this common cause to each of the nodes is such that their Choi-isomorphic states commute.

A given distribution over observed variables is said to be **quantum compatible** with a given causal structure if there is some choice of the causal parameters that yields the given distribution via Eq. (2). Note that a given set of marginal distributions over various subsets of observed variables is said to be quantum compatible with a given causal structure if and only if the joint distribution over observed variables that yields these marginals is quantum compatible with the causal structure.

Why should *any* of the inequalities we derive continue to hold for GPT causal models when the definition of an inflation of a classical causal model is different from that of a GPT causal model? I think the answer is this: as long as it is unambiguous how the causal dependences in the inflated GPT model are determined by those in the original GPT model (i.e., one does not need to specify how to broadcast a state in the model), then the analogues of lemma ?? and corollary ?? for quantum compatibility can be proven to hold. Therefore, for certain types of inflations of a DAG, it is possible to convert causal compatibility inequalities on the inflated DAG into causal compatibility inequalities on the original DAG.

But this introduces a new question: which techniques for deriving causal compatibility inequalities on the inflated DAG are true not only of classical causal models, but of GPT causal models as well? I believe that if one uses a solution of the marginal problem on the inflated DAG, together with ancestral independences to factorize some of the distributions on the pre-injectable sets, then the causal compatibility inequalities one obtains are true for GPT compatibility.

Evidently, if one derived a causal compatibility inequality using d -separation criteria that made reference to latent variables, then this would not apply to GPT compatibility. Similarly if one uses copy-index equivalence relations to derive the causal compatibility inequality.

Because quantum and GPT causal models also presume that a DAG describes the causal structure, the notion of the inflation of a DAG continues to apply to these cases. However, for a DAG G and an inflation G' thereof, our definition of the $G \rightarrow G'$ inflation of a causal model *has* presumed that the latent nodes are random variables. Consequently, one must generalize this definition to the quantum case as follows.

We hope that causal compatibility inequalities derived from broadcasting inflated DAGs will provide an additional tool for witnessing certain quantum distributions as non-classical. For example due to the results of ??, it seems conceivable that these inequalities will be much stronger and provide much tighter constraints than entropic inequalities.

It is worth pondering how it is possible that some of the inequalities that can be derived via inflation—such as Bell inequalities—have quantum violations, i.e., why one cannot expect them to be valid for all quantum distributions as well. The reason for this is that duplicating an outgoing edge in a DAG during inflation amounts to **broadcasting** the value of the random variable. For example, while in the triangle scenario, ??, the information about X in ?? is “sent” to A and C , in the inflated DAG of ??, the information about X_1 is sent to A_1 and A_2 and C . Since quantum theory satisfies a no-broadcasting theorem [? ?], one cannot expect such broadcasting to be possible quantumly. More generally, there is an analogous no-broadcasting theorem in various alternatives to quantum theory; indeed it is a generic feature of so-called generalized probabilistic theories (GPTs) [? ? ? ?]. As a consequence, a quantum or GPT causal model on the original DAG does not generally inflate to a “quantum inflated model” or “GPT inflated model” on the inflated DAG.

Some inflations, such as the one of ??, do not require such broadcasting. By removing A_1 from the broadcasting inflation of ?? we obtain the non-broadcasting inflation of ??. In ?? the channel from X to A is merely *redirected* from its original configuration in ??; there is no broadcasting of information required.

Definition 2. A DAG $G' \in \text{Inflations}[G]$, is said to be **broadcasting** if there exists a latent node in G' that has two or more children that are copy-index equivalent. Otherwise, it is said to be **nonbroadcasting**.

The broadcasting terminology stems from the fact that copy-index-equivalent variables in an inflated DAG must have the same statistical dependence on their parents, such that the map from a distribution on the parents to a distribution on the copy-index-equivalent variables is one that yields the same marginal distribution for each such variable, together with the fact that such maps have been termed *broadcasting maps* in the quantum context [cite: nobroadcastingtheorem].

It follows that every quantum causal model can be inflated to a non-broadcasting DAG, so that one obtains a quantum and general probabilistic analogue of Lemma ?? in the non-broadcasting case. Constraints derived from non-broadcasting inflations are therefore valid also for quantum and even general probabilistic distributions. In the specific case of the entropic monogamy inequality for the Triangle scenario, i.e. ?? here, this was originally noticed in Ref. [?]. Another example is ??, which was derived from the non-broadcasting inflation of ?. ?? too, therefore, is a necessary criterion for compatibility with the Triangle scenario even when the latent nodes are allowed to carry quantum or general probabilistic systems. Since the perfect-correlation distribution considered in ?? violates both of these inequalities, it evidently cannot be generated within the Triangle scenario even with quantum or general probabilistic states on the hidden nodes. This was also pointed out in Ref. [?].

On the other hand, by intentionally using broadcasting in an inflated DAG, we can specifically try to witness certain quantum or general probabilistic distributions as non-classical. This is exactly what happens in Bell’s theorem.

Even when using broadcasting inflated DAGs, it may still be possible to derive inequalities valid for quantum distributions if one appropriately modifies the nonnegativity inequalities in the marginal problem, e.g. such as ?. The modification would replace demanding nonnegativity of the full joint distribution with instead demanding the nonnegativity of only quantum-physically-meaningful marginal probability distributions.

Even when using broadcasting inflated DAGs, it may still be possible to derive inequalities valid for quantum distributions if one appropriately modifies ?? to generate a different initial set of nonnegativity inequalities. This new set should capture the nonnegativity of only quantum-physically-meaningful marginal probability distributions. Indeed, a quantum causal model on the original DAG can potentially be inflated to a quantum inflated model on the inflated DAG in terms of the logical broadcasting maps of [?]. From this perspective, a broadcasting inflated DAG is an abstract logical concept, as opposed to a feasible physical construct. However, this would result in a joint distribution over all observable variables that may have some negative probabilities, and one cannot expect ?? to hold in general. But one can still try to reformulate the marginal problem so as to refer only to the existence of joint distributions on non-broadcasting sets rather than the existence of a full joint distribution from which the marginal distributions might be recovered. Here, a set U of observable nodes is non-broadcasting if $An(U)$ does not contain two distinct copies of a node both sharing a common latent parent.

An analysis along these lines has already been carried out successfully by [?] in the derivation of entropic inequalities that are valid for all quantum distributions. Although [?] do not invoke inflated DAGs, they do seem to employ a similar type of structure to model the conditioning of a variable on a “setting” variable, and this also gives rise to non-broadcasting sets. [?] take pains to avoid including full joint probability distributions in any of their initial entropic inequalities, precisely as we would want to do in constructing our initial probability inequalities, and they successfully derive quantumly valid entropic inequalities. But so far, no inequalities polynomial in the probabilities have been derived using this method.

A tight set of inequalities characterizing quantum distributions would provide the ultimate constraints on what quantum theory allows. Deriving additional inequalities that hold for quantum distributions is therefore a priority for future research.

Mention relation to the quantum conditionals problem. Say something about how we anticipate deriving new inequalities that might be violated by quantum theory—RWS.

III. CAUSAL COMPATIBILITY INEQUALITIES FOR MARGINAL POSSIBILISTIC ASSIGNMENTS

In ??, we showed that for any given marginal scenario, one can derive constraints on the marginal probabilistic assignments from logical relations that hold among the marginal deterministic assignments. We also noted that the idea was inspired by Hardy’s version of Bell’s theorem, which makes reference only to which valuations of the observed variables are possible and which are impossible. That is, Hardy-type arguments can be understood as deriving constraints on the marginal *possibilistic* assignments from logical relations that hold among the marginal deterministic assignments. To get *all* constraints that follow from logical tautologies in a given marginal scenario, it suffices to follow the prescription of ??, with one alteration: replace the application of the union bound to a possibilistic analogue thereof.

First, we formalize the notion of a marginal possibilistic assignment. Let $S_U(\mathbf{u})$ denote the possibilistic assignment to the joint valuation \mathbf{u} of the subset of observed variables U . Specifically, we set $S_U(\mathbf{u}) = 0$ if $P_U(\mathbf{u}) = 0$ and we set $S_U(\mathbf{u}) = 1$ if $P_U(\mathbf{u}) \neq 0$.

Next, recall that if a logical tautology can be expressed as

$$[E_0] \implies [E_1] \vee \dots \vee [E_n], \tag{3}$$

then the union bound implies that

$$P(E_0) \leq \sum_{j=1}^n P(E_j). \quad (4)$$

The possibilistic analogue of the union bound is the inference from Eq. (??) to

$$S(E_0) \leq \max_{j \in \{1, \dots, n\}} S(E_j). \quad (5)$$

For instance, in the marginal problem where the variables are $\{A, B, C\}$, with each being binary, and the contexts are the pairs $\{AB\}$, $\{AC\}$, and $\{BC\}$, we considered in ?? the tautology

$$[\mathbf{A=1}, \mathbf{C=1}] \implies [\mathbf{A=1}, B=1] \vee [B=0, \mathbf{C=1}], \quad (6)$$

and derived therefrom the following constraint on marginal probability assignments

$$P_{AC}(\mathbf{11}) \leq P_{AB}(\mathbf{11}) + P_{BC}(\mathbf{01}). \quad (7)$$

The corresponding constraint on marginal possibilistic assignments is simply

$$S_{AC}(\mathbf{11}) \leq \max\{S_{AB}(\mathbf{11}), S_{BC}(\mathbf{01})\}. \quad (8)$$

Ancestral independences also imply factorization for possibilistic assignments. For example, recall that the DAG of Fig. ?? has $\{A_2, B_1\}$ as a pre-injectable set, where A_2 and B_1 are ancestrally independent. This implies that if a given valuation $[A_2 = a]$ is possible *and* a given valuation $[B_1 = b]$ is possible, then the joint valuation $[A_2 = a, B_1 = b]$ is also possible,

$$A_2 \not\lhd_{\emptyset^*} B_1 \implies S_{A_2 B_1}(ab) = S_{A_2}(a)S_{B_1}(b). \quad (9)$$

If one wishes, one can derive causal compatibility inequalities for possibilistic assignments rather than probabilistic ones. For instance, for the DAG of Fig. ??, the marginal possibilistic assignments to $\{A_2 C_1\}$, $\{B_1 C_1\}$ and $\{A_2 B_1\}$ satisfy the analogue of Eq. (??), namely,

$$S_{A_2 C_1}(\mathbf{11}) \leq \max\{S_{A_2 B_1}(\mathbf{11}), S_{B_1 C_1}(\mathbf{01})\}. \quad (10)$$

Because Fig. ?? is an inflation of the triangle scenario, and $\{A_2 C_1\}$, $\{B_1 C_1\}$ and $\{A_2 B_1\}$ are pre-injectable sets, we can substitute Eq. (??) into Eq. (??) to obtain

$$S_{A_2 C_1}(\mathbf{11}) \leq \max\{S_{A_2}(\mathbf{1})S_{B_1}(1), S_{B_1 C_1}(\mathbf{01})\}. \quad (11)$$

which is a possibilistic causal compatibility inequality for the triangle scenario. It can be interpreted as asserting that if $[A_2=1, C_1=1]$ is possible, then *either* $[B_1=0, C_1=1]$ is possible or *both* $[A_2=1]$ and $[B_1=1]$ are possible. Equivalently, it asserts that if $[B_1=0, C_1=1]$ is impossible and *either* $[A_2=1]$ or $[B_1=1]$ is impossible, then necessarily $[A_2=1, C_1=1]$ is impossible as well.