Branching Space-Time, Modal Logic and the Counterfactual Conditional

Thomas Müller, Philosophisches Seminar I,
Universität Freiburg, D–79085 Freiburg, Germany;
e-mail: tmueller@fdm.uni-freiburg.de

Abstract

The paper gives a physicist's view on the framework of branching space-time (Belnap, *Synthese* **92** (1992), 385–434). Branching models are constructed from physical state assignments. The models are then employed to give a formal semantics for the modal operators "possibly" and "necessarily" and for the counterfactual conditional. The resulting formal language can be used to analyze quantum correlation experiments. As an application sketch, Stapp's premises LOC1 and LOC2 from his purported proof of non-locality (*Am. J. Phys.* **65** (1997), 300–304) are analyzed.

1 Introduction

Branching space-time (BST) was proposed by Belnap [1] as a rigorous framework for unifying relativity and indeterminism. One leading idea behind the development of that framework was that physical possibility is an important concept that should be treated rigorously. Already in the paper that introduced BST, it was shown that the framework could help elucidate the question of quantum correlations (cf. [1], sect. 11). Parallel to this development, the importance of a formal analysis of modal and also counterfactual reasoning in clarifying the premises and implications of Bell-type theorems has recently been acknowledged by many researchers (cf. the extensive list of references in [3]). It is the aim of this paper to bring together BST and formal modal and counterfactual reasoning from a physicist's perspective.

BST starts with a given partial ordering $\langle W, \leq \rangle$, from which histories are then derived as a secondary notion. From a physicist's perspective, the notion of alternative histories or scenarios appears to be primary. BST models should thus be derived from given alternative scenarios, spelled out in terms of alternative states. In this paper, models of BST are constructed by pasting states defined on Minkowski space-time (Section 2). This framework is then used to give a formal semantics for a formal language incorporating the modal operators $\Diamond$ ("possibly") and $\Box$ ("necessarily") and the counterfactual conditional $\ldots \Box\!\!\rightarrow \ldots$ ("if ... were the case, ... would be the case") (Section 3). This formal language can be employed to analyze quantum correlation experiments. As an application sketch, the framework is used to assess the premises LOC1 and LOC2 of Stapp's recent purported proof of non-locality (Section 4).

2 Pasted Minkowski space-times

Belnap's original branching space-time framework is based on a given nonempty set of point events W, called "Our World", that is equipped with a partial ordering $\leq$. From this structure $\langle W, \leq \rangle$, *histories* are carved: a history is a maximal upward directed subset of "Our World". This very general approach allows for histories to be space-times with various metrics. E.g., Our World could be a branching model of general relativity. We will here use a simpler approach that is less general, but sufficient for our purposes: Each history will be a Minkowski space-time. This means that the metric is fixed, and thus space-time points can easily be identified across histories. One may thus speak about "the same" space-time point X in history σ and in history η, even though "what happens" at X may differ in σ and η. This will now be made more precise.

States on Minkowski space-time Space-time points X, Y, etc., will be treated as elements of $\mathbb{R}^4$, i.e., $X = \langle x_0, x_1, x_2, x_3 \rangle$, $Y = \langle y_0, y_1, y_2, y_3 \rangle$, etc. The *Minkowskian space-time distance* $D^2_M(X, Y)$ between X and Y is defined as

$$D^2_M(X, Y) = -(x_0 - y_0)^2 + \sum_{i=1}^{3} (x_i - y_i)^2, \tag{1}$$

where we use units in which $c = 1$ for convenience. To introduce Minkowski space-time, the *Minkowskian ordering* $\leq_M$ is defined on $\mathbb{R}^4$ in the usual way:

$$X \leq_M Y \quad \text{iff} \quad D^2_M(X, Y) \leq 0 \quad \text{and} \quad x_0 \leq y_0. \tag{2}$$

The irreflexive ordering corresponding to $\leq_M$ is denoted $<_M$ ($X <_M Y$ iff $X \leq_M Y$ and $X \neq Y$). X and Y are *space-like related*, written $X \text{ SLR } Y$, iff $D^2_M(X, Y) > 0$, i.e., iff $X \not\leq_M Y$ and $Y \not\leq_M X$.

Let Σ be a nonempty, at most countable set of space-time labels $\sigma, \eta, \ldots$[1] We think of σ as labeling a scenario in space-time. (The term "history" will be reserved for the technical meaning that it has in BST.) Let P be a set of point properties (one may think of various field strengths — finding out what the right P is is a question of physics, not one of conceptual analysis).[2] Then let $S : \Sigma \times \mathbb{R}^4 \to P$ be a function that assigns a state S_σ (a mapping from $\mathbb{R}^4$ to P) to each space-time label σ. This global state S could be all we need to derive a full model of branching space-time: σ and η could be pasted together in the whole region before their states diverge. Starting with S alone allows for very irregular branching models, especially concerning the boundaries of regions of overlap (cf. (4) below). Thus we will adopt a different strategy.

[1] The countability assumption will only be needed in the proof of Lemma 3 below. With uncountably many scenarios, the branching construction to be given below is still possible, but the branching model could contain "non-standard" histories violating the form given in Lemma 3. Arguably, countably many scenarios are sufficient to describe any real experiment, so the assumption is not much of a constraint.

[2] The concept of a point having a property may appear questionable. However, even in the original formulation of BST, points were assumed to be "concrete particulars" [1, p. 388]. Furthermore, a more useful notion of a state of a region may supervene on the framework given here.

Given splitting points The construction of explicit models, e.g., of Bell-type experiments, is easier, and more intuitive, if one uses some structure in addition to the state S. One usually knows which space-time points are crucial for an experiment, i.e., at which space-time points scenarios split. We thus assume that together with S, for each pair of space-time labels σ, η a set $C_{\sigma,\eta} \subset \mathbb{R}^4$ of *splitting points* is given. The guiding idea is that $C_{\sigma,\eta}$ gives the locations at which a "choice" between σ and η is made. Thus, the elements of $C_{\sigma,\eta}$ must be mutually space-like related: After a choice has been made, there can be no further choice. We will also assume *historical connection*, i.e., for all σ, η, $C_{\sigma,\eta} \neq \emptyset$ iff $\sigma \neq \eta$: There is (was) a choice between any two scenarios at some point; any two scenarios share a common historical root.

The various sets of splitting points cannot be assigned completely arbitrary. In line with the guiding idea, two consistency requirements have to be imposed: First, while the notation "$C_{\sigma,\eta}$" may suggest that the order of the labels is important, this is not so: $C_{\sigma,\eta} = C_{\eta,\sigma}$. (A choice between σ and η is also a choice between η and σ.) Secondly, if $C_{\sigma,\eta}$ and $C_{\eta,\gamma}$ are given, a choice between σ and γ must respect the region that is already known to coincide for all three σ, η and γ, so that a splitting point for σ and γ must coincide with or be later than a splitting point for σ, η and for η, γ:

$$\forall X \in C_{\sigma,\gamma}\ \exists Y \in C_{\sigma,\eta} \cup C_{\eta,\gamma}\ \ Y \leq_M X. \tag{3}$$

As a further constraint, it is assumed that the set of splitting points between any two histories is "small". In the original formulation of BST, it was stated that (continuous) splitting along a simultaneity slice was conceptually possible, but "appears weird" [1, p. 414]. For the present purposes (cf. Lemma 4 below) we will need to be stricter and not just wonder, but rule out that possibility: For all $\sigma, \eta \in \Sigma$, the set $C_{\sigma,\eta}$ has to be finite.

As long as no choice has been made, scenarios overlap. The *region of overlap* $R_{\sigma,\eta}$ between σ and η is defined as

$$R_{\sigma,\eta} = \{X \in \mathbb{R}^4 \,|\, \neg\exists Y \in C_{\sigma,\eta}\ Y < X\}. \tag{4}$$

By this definition, splitting points belong to the region of overlap, while the whole future light cone above any splitting point, including the boundaries, is outside the region of overlap. Note that this definition gives the correct result in the vacuous case $\eta = \sigma$: as $C_{\sigma,\sigma} = \emptyset$, we get $R_{\sigma,\sigma} = \mathbb{R}^4$. In terms of regions of overlap, the second requirement on the assignment of splitting points (3) reads, perhaps more mnemonically:

$$R_{\sigma,\gamma} \supseteq R_{\sigma,\eta} \cap R_{\eta,\gamma}, \tag{5}$$

i.e., the region of overlap between σ and γ must include at least all points that are common to the overlap between σ and η and between η and γ.

The given state S must comply with the assignment of splitting points: In order to be consistent with $C_{\sigma,\eta}$, states S_σ, S_η must not differ inside the region of overlap, but they must differ just above. Explicitly, the check for consistency is:

DEFINITION 1 (CONSISTENT SPLITTING STRUCTURE)
A given state S is consistent *with a given set of sets of splitting points $\mathcal{C} = \{C_{\sigma,\eta} \,|\, \sigma, \eta \in \Sigma\}$ iff for all $\sigma, \eta \in \Sigma$*

1. *states coincide in the region of overlap:*

$$\forall X \in R_{\sigma,\eta} \ S_\sigma(X) = S_\eta(X),$$

2. *there is no empty splitting:*

$$\forall X \in C_{\sigma,\eta} \forall Y \in \mathbb{R}^4 \quad \big(X <_M Y \to \exists Z \in \mathbb{R}^4 \ \ (Z \leq_M Y \wedge S_\sigma(Z) \neq S_\eta(Z))\big) \, .$$

Note that we do not require that states S_σ and S_η differ everywhere outside their region of overlap $R_{\sigma,\eta}$, nor can we simply set $R_{\sigma,\eta} = \{X \in \mathbb{R}^4 \mid S_\sigma(X) = S_\eta(X)\}$. States are allowed to reconverge — only scenarios are not.

Definition 1 can be read in two ways. Minimally, one can assume S given and use the definition as a test for a guess of $\mathcal{C}$. However, one may also read the definition as a claim fixing the physically possible states: mathematical functions S that do not allow for a consistent splitting structure are "unphysical" and should not be used to construct a physical branching model.

Derivation of a branching space-time model $\langle B, \leq_S \rangle$ from S and $\mathcal{C}$ From a consistent splitting structure, a model of Belnap's branching space-time can be derived by identifying points in regions of overlap. This is done in two steps. First, space-time points are distinguished across histories: Let $\Sigma \times \mathbb{R}^4$ be the set of distinguished space-time points. Elements of $\Sigma \times \mathbb{R}^4$ are pairs $\langle \sigma, X \rangle$, conveniently written X_σ, "the space-time point X in scenario σ". The identification is then to be affected by an equivalence relation that pools together points in regions of overlap. We define the equivalence relation $\equiv_S$ on points X_σ as follows:

$$X_\sigma \equiv_S Y_\eta \quad \text{iff} \quad X = Y \quad \text{and} \quad X \in R_{\sigma,\eta}. \tag{6}$$

We have to show that $\equiv_S$ is indeed an equivalence relation:

LEMMA 1
$\equiv_S$ is an equivalence relation on the set $\Sigma \times \mathbb{R}^4 = \{X_\sigma \mid \sigma \in \Sigma, X \in \mathbb{R}^4\}$.

Proof: (1) $\equiv_S$ is reflexive, since $R_{\sigma,\sigma} = \mathbb{R}^4$. (2) $\equiv_S$ is symmetric, since $R_{\sigma,\eta} = R_{\eta,\sigma}$ (as $C_{\sigma,\eta} = C_{\eta,\sigma}$). (3) Transitivity holds in virtue of (5): if $X \in R_{\sigma,\eta}$ and $X \in R_{\sigma,\gamma}$, then also $X \in R_{\eta,\gamma}$. □

The set B that will correspond to "Our World" W in the original formulation of BST is the quotient structure of the set of distinguished space-time points relative to the equivalence relation $\equiv_S$, i.e.,

$$B = \Sigma \times \mathbb{R}^4 / \equiv_S \ = \ \{[X_\sigma] \mid \sigma \in \Sigma, X \in \mathbb{R}^4\}, \tag{7}$$

where the equivalence class $[X_\sigma]$ for the point X_σ is

$$[X_\sigma] = \{X_\gamma \mid X_\sigma \equiv_S X_\gamma\}. \tag{8}$$

So far, B is a nonempty set. The elements of B can with some plausibility be viewed as "concrete particulars", since the given state S assigns a physical property $p \in P$,

$p([X_\sigma]) = S_\sigma(X)$, to each point $[X_\sigma]$. (This p is independent of the representative X_σ of $[X_\sigma]$ in virtue of definition 1.1 and (6).) Given the base set B, the next step in a definition of a model of BST is to define a partial ordering $\leq_S$ on B. We define

$$[X_\sigma] \leq_S [Y_\gamma] \quad \text{iff} \quad X \leq_M Y \quad \text{and} \quad X_\sigma \equiv_S X_\gamma, \tag{9}$$

i.e., at the space-time point X, which has to be before Y in the usual Minkowskian ordering, σ and γ must belong to the same equivalence class. This definition is given in terms of representatives of the equivalence classes, but it does not depend on a particular representative: (1) Let $X_\eta \in [X_\sigma]$. Then $X_\eta \equiv_S X_\gamma$ by transitivity of $\equiv_S$. (2) Let $Y_\delta \in [Y_\gamma]$. Then $X_\sigma \equiv_S X_\delta$ in virtue of (5).

The strict ordering corresponding to $\leq_S$ is denoted $<_S$. We need to prove that $\leq_S$ is indeed a partial ordering:

LEMMA 2
$\leq_S$ is a partial ordering on B.

Proof: (1) $\leq_S$ is reflexive by the reflexivity of $\equiv_S$. (2) $\leq_S$ is transitive: Let $[X_\sigma] \leq_S [Y_\eta]$ and $[Y_\eta] \leq_S [Z_\gamma]$. $X \leq_M Z$ holds by transitivity of the Minkowskian ordering. As $Y_\eta \equiv_S Y_\gamma$, we have $[Y_\eta] = [Y_\gamma]$, thus $X_\sigma \equiv_S X_\gamma$, which gives $[X_\sigma] \leq_S [Z_\gamma]$. (3) $\leq_S$ is antisymmetric: Let $[X_\sigma] \leq_S [Y_\eta]$ and $[Y_\eta] \leq_S [X_\sigma]$. Then $X = Y$ by the antisymmetry of the Minkowskian ordering, and $X_\sigma \equiv_S X_\eta$ gives in fact $[X_\sigma] = [Y_\eta]$. □

Proof of the BST property The following main Theorem asserts that by now, from S and the consistent set of sets of splitting points $\mathcal{C}$ we have indeed constructed a model of BST.

THEOREM 1
$\langle B, \leq_S \rangle$ is a model of branching space-time.

Proof: B is nonempty (since Σ was required to be nonempty), and by Lemma 2 we know that $\leq_S$ is a partial ordering on B. Before we can go on to show the further required properties of $\leq_S$ and the prior choice principle, we need to prove a Lemma that states that the histories in $\langle B, \leq_S \rangle$ are the histories that were intended by our construction. To give the definition of a history explicitly, $h \subset B$ is a history in B iff it is maximal upward directed, i.e., if h is upward directed:

$$\forall e_1, e_2 \in h\, \exists e \in h\ e_1 \leq_S e \text{ and } e_2 \leq_S e, \tag{10}$$

and h is maximal w.r.t. this property, i.e.

$$\forall e_3 \in B - h\, \exists e_4 \in h\ \neg\exists e \in h\ e_3 \leq_S e \text{ and } e_4 \leq_S e. \tag{11}$$

LEMMA 3
Every history h in B is of the form $h = \{[X_\sigma] \,|\, X \in \mathbb{R}^4\}$ for some (not necessarily unique) $\sigma \in \Sigma$.

Proof: The first direction is simple. Given $h = \{[X_\sigma] \mid X \in \mathbb{R}^4\}$ for some $\sigma \in \Sigma$, we have to show that both (10) and (11) are satisfied. (10) holds since for any $e_1 = [X_\sigma]$, $e_2 = [Y_\sigma] \in h$, the Minkowskian ordering will supply Z s.t. $X \leq_M Z$ and $Y \leq_M Z$; thus, we can take $e = [Z_\sigma]$. In order to prove (11), we proceed indirectly. Thus, assume that (11) fails, i.e., that there is a $[X_\eta] \in B - h$ that could extend h and preserve (10). As $[X_\eta] \notin h$, we must have $X_\eta \not\equiv_S X_\sigma$. By (10), there has to be a point $Y_\sigma \in h$ s.t. $[X_\sigma] \leq_S [Y_\sigma]$ (so that $X \leq_M Y$), but also $[X_\eta] \leq_S [Y_\sigma]$. This, however, implies $X_\eta \equiv_S X_\sigma$, which contradicts the assumption. Thus h as given is indeed maximal. This completes the first half of the proof.

To prove the other direction, assume that $h \subset B$ is a history in B, i.e., h satisfies (10) and (11). We have to show that h is of the form $\{[X_\sigma] \mid X \in \mathbb{R}^4\}$ for some $\sigma \in \Sigma$. We proceed in three steps:

1. If for some $\sigma, \eta \in \Sigma$ both $[X_\sigma] \in h$ and $[X_\eta] \in h$, then $X_\sigma \equiv_S X_\eta$.
   Proof: By (10), there must be some $[Y_\gamma] \in h$ s.t. $[X_\sigma] \leq_S [Y_\gamma]$ and $[X_\eta] \leq_S [Y_\gamma]$. This implies $X_\sigma \equiv_S X_\gamma$ and $X_\eta \equiv_S X_\gamma$. The assertion follows by transitivity of $\equiv_S$.
   For $X \in \mathbb{R}^4$, we define $\Sigma_h(X) = \{\sigma \in \Sigma \mid [X_\sigma] \in h\}$. (So far, this set might be empty for some X, but cf. point 3 below.)

2. There is (at least) one $\sigma \in \Sigma$ s.t. for all $[X_\eta] \in h$, $X_\eta \equiv_S X_\sigma$. (I.e., all elements of h can be written as $[X_\sigma]$ for some $X \in \mathbb{R}^4$ and that one σ.)
   Proof: Assume the contrary, i.e.,
   $$\forall \sigma \in \Sigma \; \exists [Y_\gamma] \in h \;\; Y_\gamma \not\equiv_S Y_\sigma. \tag{*}$$
   Select some arbitrary $[X_\eta] \in h$. We will now give a diagonal argument leading to a contradiction. Let I be the cardinality of $\Sigma_h(X)$ (at most countably infinite by assumption), and let $\sigma : I \to \Sigma_h(X)$ be a bijection. Thus, $\sigma_i = \sigma(i) \in \Sigma_h(X)$ for $i < I$. By (*), for all σ_i there is a $[(Y_i)_{\gamma_i}] \in h$ such that
   $$(Y_i)_{\gamma_i} \not\equiv_S (Y_i)_{\sigma_i}, \quad \text{i.e.,} \quad \sigma_i \notin \Sigma_h(Y_i). \tag{**}$$
   Now we define inductively a sequence of points $[(Z_i)_{\eta_i}] \in h, i < I$. Select $[(Z_0)_{\eta_0}]$ above both $[X_\sigma]$ and $[(Y_0)_{\gamma_0}]$. (As h is a history, such a point exists by (10).) At stage $i + 1$, select $[(Z_{i+1})_{\eta_{i+1}}]$ above both $[(Z_i)_{\eta_i}]$ and $[(Y_{i+1})_{\gamma_{i+1}}]$. Thus $\Sigma_h(Z_{i+1}) \subseteq \Sigma_h(Z_i)$, $\Sigma_h(Z_{i+1}) \subseteq \Sigma_h(Y_{i+1})$ and $\Sigma_h(Z_{i+1}) \subseteq \Sigma_h(X)$. By this construction, the set
   $$\bigcap_{i<I} \Sigma_h(Z_i) \neq \emptyset,$$
   i.e., it contains some $\sigma \in \Sigma_h(X)$. But this $\sigma = \sigma_i$ for some $i < I$. Thus, in particular $\sigma_i \in \Sigma_h(Z_i)$, which implies $\sigma_i \in \Sigma_h(Y_i)$, contradicting (**) and thus our assumption (*).

3. With the σ from step 2, $h = \{[X_\sigma] \mid X \in \mathbb{R}^4\}$.
   Proof: We know that all elements of h can be written as $[X_\sigma]$ for some X. It remains to prove that for all $X \in \mathbb{R}^4$, $[X_\sigma] \in h$. This, however, follows directly

from the first half of the proof of this Lemma: As $\{[X_\sigma] \mid X \in \mathbb{R}^4\}$ is a history in B, any proper subset is not a history due to (11).

This completes the proof of Lemma 3. $\square$

We now continue the proof of the main Theorem. It remains to show the following four properties of the partial ordering $\leq_S$ and, finally, the prior choice principle:

1. The ordering $<_S$ is dense: Let $[X_\sigma] <_S [Y_\eta]$. Thus $X_\eta \equiv_S X_\sigma$. By the Minkowskian ordering, there is some $Z \in \mathbb{R}^4$ s.t. $X <_M Z <_M Y$. Then, $[X_\sigma] <_S [Z_\eta] <_S [Y_\eta]$.

2. $\leq_S$ has no maximal elements: This follows directly from the corresponding property of $\leq_M$.

3. Every lower bounded (outcome) chain O in B has an infimum in B.
Proof: As the elements of a chain are mutually comparable, O can be extended to a history, and thus (by Lemma 3), all the elements of O can be written as $[X_\sigma]$ for some specific σ. The set $\{Y \mid [Y_\sigma] \in O\}$ is a lower bounded subset of $\mathbb{R}^4$, thus it has an infimum Z w.r.t. the Minkowskian ordering $\leq_M$. The point $[Z_\sigma]$ is then the infimum of O.

4. Every upper bounded (initial) chain I in B has a supremum $\sup_h(I)$ in each history h to which is belongs.
Proof: Assume that $I \subset h$. From Lemma 3, all the elements of I have the form $[X_\sigma]$ for some $\sigma \in \Sigma$. The set $\{X \mid [X_\sigma] \in I\}$ is an upper bounded subset of $\mathbb{R}^4$, thus it has a supremum Y w.r.t. the Minkowskian ordering $\leq_M$. The point $[Y_\sigma]$ is then the supremum of I in h.

5. The prior choice principle: For any lower bounded chain $O \subset h_1 - h_2$, there is a choice point for O, i.e., a point $e \in B$ s.t. e is maximal in $h_1 \cap h_2$ and $e <_S O$ (i.e., for all $e' \in O$ $e <_S e'$).
Proof: By Lemma 3, we may call the histories $h_1 = h_\sigma$, $h_2 = h_\eta$ to indicate the generic form of their members. By the definition of $\equiv_S$ it follows that the set of space-time points at which h_σ and h_η meet is exactly their region of overlap $R_{\sigma,\eta}$: $h_\sigma \cap h_\eta = \{[X_\sigma] \mid X \in R_{\sigma,\eta}\}$. Let $[X_\sigma] = \inf(O)$. We can distinguish two cases: (1) $[X_\sigma] \notin h_\eta$. Then $X \notin R_{\sigma,\eta}$, which means that there is a $Y \in C_{\sigma,\eta}$ s.t. $Y <_M X$. The point $[Y_\sigma]$ is maximal in $h_\sigma \cap h_\eta$, since for any $[Z_\gamma]$ with $[Y_\sigma] <_S [Z_\gamma]$ we have $Y <_M Z$, i.e., $Z \notin R_{\sigma,\eta}$. (2) $[X_\sigma] \in h_\eta$. Assume that $[X_\sigma]$ is not maximal in $h_\sigma \cap h_\eta$, i.e., there is a $[Y_\gamma] \in h_\sigma \cap h_\eta$ s.t. $[X_\sigma] <_S [Y_\gamma]$. As $[X_\sigma]$ is the infimum of O, there is a point $e \in O$ s.t. $e \leq_S [Y_\gamma]$. But $O \subset h_\sigma - h_\eta$, thus $e \notin h_\eta$, which implies $[Y_\gamma] \notin h_\eta$, contradicting the assumption.

This completes the proof of the main Theorem: The structure $\langle B, \leq_S \rangle$ is in fact a model of BST, and by Lemma 3, its histories are the intended ones. $\square$

The next Lemma shows that given our construction, we recover the given set of splitting points $C_{\sigma,\eta}$ as the set of *choice points* for histories h_σ and h_η:

LEMMA 4

For histories $h_\sigma, h_\eta \subset B$, the set $C_{\sigma,\eta}$ is the set of choice points, i.e., $C_{\sigma,\eta}$ is the set of points maximal in $h_\sigma \cap h_\eta$.

Proof: The first direction has already been shown in connection with the prior choice principle: Let $Y \in C_{\sigma,\eta}$. Then $[Y_\sigma]$ is maximal in $h_\sigma \cap h_\eta$, since for any $[Z_\gamma]$ with $[Y_\sigma] <_S [Z_\gamma]$ we have $Y <_M Z$, i.e., $Z \notin R_{\sigma,\eta}$, i.e., $[Z_\gamma] \notin h_\sigma \cap h_\eta$.

To show the other direction, let $[X_\sigma]$ be maximal in $h_\sigma \cap h_\eta$. We have to show that $X \in C_{\sigma,\eta}$. Assume the contrary. We will now show that on this assumption $[X_\sigma]$ cannot be maximal in $h_\sigma \cap h_\eta$ by constructing a point $W \in \mathbb{R}^4$ s.t. (1) $X <_M W$ and (2) $W \in R_{\sigma,\eta}$, so that $[X_\sigma] <_S [W_\sigma]$ and $[W_\sigma] \in h_\sigma \cap h_\eta$. In the construction, we will treat $\mathbb{R}^4$ as a vector space with the usual Euclidean norm $|\cdot|$. For $X = \langle x_0, x_1, x_2, x_3\rangle$ and $Y = \langle y_0, y_1, y_2, y_3\rangle$, we separate the spatial distance from the temporal distance by setting[3]

$$\vec{\Delta}s(X,Y) = \langle 0, y_1 - x_1, y_2 - x_2, y_3 - x_3\rangle, \quad \vec{\Delta}t(X,Y) = \langle y_0 - x_0, 0, 0, 0\rangle. \tag{12}$$

Thus $|\vec{\Delta}s(X,Y)|^2 + |\vec{\Delta}t(X,Y)|^2$ is the Euclidean distance squared between X and Y, and X SLR Y iff $|\vec{\Delta}t(X,Y)| < |\vec{\Delta}s(X,Y)|$.

Let $Y \in C_{\sigma,\eta}$ have minimal Euclidean distance from X. (As $C_{\sigma,\eta}$ is finite, such a Y exists, and the distance is greater than zero as $X \notin C_{\sigma,\eta}$ by assumption.) Set $\vec{\Delta}s = \vec{\Delta}s(X,Y)$, $\vec{\Delta}t = \vec{\Delta}t(X,Y)$. We know that for all $Z \in C_{\sigma,\eta}$ (including Y), X SLR Z, for otherwise (equality being excluded by assumption) either $X <_M Z$, implying that $[X_\sigma]$ cannot be maximal in $h_\sigma \cap h_\eta$, or $Z <_M X$, implying that $X \notin R_{\sigma,\eta}$, both contrary to assumption. Thus, in particular, X SLR Y, i.e., $|\vec{\Delta}t| < |\vec{\Delta}s|$. Now set

$$\epsilon = \frac{1}{2} \min_{Z \in C_{\sigma,\eta}} (|\vec{\Delta}s(Z,X)| - |\vec{\Delta}t(Z,X)|), \tag{13}$$

which is greater than zero since X SLR Z for all $Z \in C_{\sigma,\eta}$. The point W to be constructed is then

$$W = X - \frac{\epsilon}{2}\frac{\vec{\Delta}s}{|\vec{\Delta}s|} + \frac{\epsilon}{2}\langle 1, 0, 0, 0\rangle. \tag{14}$$

(1) $X <_M W$: From the construction, it can be read off directly that $D^2_M(X,W) = 0$ and $x_0 < w_0$. (2) $W \in R_{\sigma,\eta}$: We need to prove that for no $Z \in C_{\sigma,\eta}$, $Z <_M W$, i.e., for all $Z \in C_{\sigma,\eta}$, either Z SLR W or $W \le_M Z$. Let $Z \in C_{\sigma,\eta}$. $W \le_M Z$ can be excluded: as $X <_M W$, we would have $X <_M Z$, contradicting the maximality of $[X_\sigma]$. Thus we need to prove Z SLR W, i.e., $|\vec{\Delta}t(Z,W)| < |\vec{\Delta}s(Z,W)|$. Using the triangle inequality for the Euclidean norm and noting that from (14), $|\vec{\Delta}s(W,X)| < \epsilon$ and $|\vec{\Delta}t(W,X)| < \epsilon$, we have

$$\begin{aligned}
&|\vec{\Delta}s(Z,X)| \le |\vec{\Delta}s(Z,W)| + |\vec{\Delta}s(W,X)| < |\vec{\Delta}s(Z,W)| + \epsilon, \quad \text{i.e.,}\\
&|\vec{\Delta}s(Z,W)| > |\vec{\Delta}s(Z,X)| - \epsilon, \quad \text{and}\\
&|\vec{\Delta}t(Z,W)| \le |\vec{\Delta}t(Z,X)| + |\vec{\Delta}t(X,W)| < |\vec{\Delta}t(Z,X)| + \epsilon.
\end{aligned} \tag{15}$$

[3]These are frame-relative notions, but they will be employed to prove frame-independent (Lorentz-invariant) assertions only.

By (13), $|\vec{\Delta}s(Z,X)| \geq |\vec{\Delta}t(Z,X)| + 2\epsilon$, and this together with (15) yields Z SLR W, finishing the proof of the Lemma. $\square$

Summing up: our construction derived a model of BST $\langle B, \leq_S \rangle$ from the given consistent splitting structure $\langle S, \mathcal{C} \rangle$ in such a way that the histories in the BST model are the given scenarios from S, and the choice points for the histories are the given splitting points from $\mathcal{C}$.

3 The formal language and its semantics

Based on models of BST like the ones constructed in the previous Section, we will now define a formal propositional language for talking about space-time, modality and counterfactuals.

3.1 Formal notions ("the language")

The statements of our language are built up from atomic formulae like $\phi(X)$, $\psi(X)$ that express a fact (or falsehood) about the state of a certain space-time point X. Even though the formula $\phi(X)$ looks like a predication, in the simple semantics given here, we will not quantify over space-time points. Technically, any $\phi(X)$ will therefore be treated as an atomic proposition. Let Φ be the set of all "predicates", i.e., for $\phi \in \Phi$, $X \in \mathbb{R}^4$, "$\phi(X)$" is an atomic formula. Complex formulae are built up recursively from the atomic ones via the usual truth-functional connectives ($\wedge$, $\vee$, $\neg$, $\rightarrow$, $\leftrightarrow$), the one-place modal operators $\Box$ and $\Diamond$, and the two-place "would" counterfactual $\Box\!\!\rightarrow$.[4]

Corresponding to the recursive definition of complex formulae, we now give a recursive formal semantics. This semantics is based on a valuation that fixes the truth value of all atomic formulae.

3.2 States and valuations

The physical property p of the space-time point X in history σ is given by the state S: $p = S_\sigma(X)$. Relative to the given set Φ, the physical state S determines a semantic *valuation* $V : \Sigma \times \mathbb{R}^4 \times \Phi \rightarrow \{T, F\}$ that assigns one of the truth values T ("true") or F ("false") to $\phi \in \Phi$ at a space-time point $X \in \mathbb{R}^4$ in a history $\sigma \in \Sigma$. We make no explicit assumption about how a state S fixes a valuation V, since we wish to separate the physics clearly from the semantics. However, it may be convenient to picture $V(\cdot, \cdot, \phi)$ as a partition of P: $V(\sigma, X, \phi) = T$ iff $S_\sigma(X) \in P_\phi$, where P_ϕ is the characteristic set of physical properties for ϕ.

A *model* $\mathfrak{M} = \langle B, \leq_S, V \rangle$ is a branching structure together with a valuation. In order to be unambiguously true or false, an atomic formula has to be evaluated in a model $\mathfrak{M}$ (derived from a state S) in a certain history. In fact, because of the modal operators,

[4] The language should be extended to include tense operators as well as the Prior-inspired operator "it appears from another Lorentz frame that" (cf. note 5 below). Probabilities should also be added; for some first steps, cf. the framework of stochastic outcomes in BST [7; 9].

formulae will be evaluated in a model and in a history at a space-time point. The intuition behind this is that a formula is evaluated from a certain "point of view".[5]

3.3 Formal semantics for the truth-functional connectives

The primitive notion for the formal semantics is "atomic formula $\phi(X)$ is satisfied by a model $\mathfrak{M}$ in a history σ and at a space-time point Y", defined as

$$\mathfrak{M}, \sigma, Y \models \phi(X) \quad \text{iff} \quad V(\sigma, X, \phi) = T, \tag{16}$$

"the given valuation V assigns the truth-value true to ϕ at X in history σ". (Note that in this atomic case, the location Y at which the atomic formula $\phi(x)$ is evaluated plays no role.) The satisfaction clauses for the truth-functional connectives are the standard ones, e.g.:

$$\mathfrak{M}, \sigma, Y \models \neg\Psi \quad \text{iff} \quad \text{not} \quad \mathfrak{M}, \sigma, Y \models \Psi, \tag{17}$$

$$\mathfrak{M}, \sigma, Y \models \Psi \wedge \Psi' \quad \text{iff} \quad \mathfrak{M}, \sigma, Y \models \Psi \quad \text{and} \quad \mathfrak{M}, \sigma, Y \models \Psi'. \tag{18}$$

3.4 Formal semantics for the modal operators

The next important step is to define satisfaction of a formula prefixed by a modal operator. Since the two modalities are inter-definable, we will only treat the weak modality $\Diamond$ (possibility; the strong modality $\Box$, necessity, is defined via $\Box\Psi \Leftrightarrow \neg\Diamond\neg\Psi$).

One can distinguish a number of different concepts of modality.[6] In the context of quantum mechanical reasoning, one crucial feature is that what was once possible can become impossible in the course of an experiment. The only notion of possibility that allows for this dynamical kind of modality is possibility based on reality (for the phrase, cf. [15]): Ψ is possible in a model $\mathfrak{M}$ in a history σ and at a space-time point Y iff either (1) the history σ up to Y already fixes the truth-value of Ψ, and Ψ holds in σ at Y, or (2) from what has happened in σ at Y, Ψ is still open to occur. This concept is called "possibility based on reality" because the real course of events (i.e., the course of events in σ up to Y) determines what is possible: either something has become actual and is thus possible as well as necessary, or it is still open to occur, given what has occurred so far. According to this notion, possibility and necessity coincide for the past, but differ for the future. This is exactly as it should be: the past is fixed, the future is open. In order

[5] For a sketch of a formal logic of points of view, inspired by Prior [11], cf. Chapter 4.5 of my dissertation *Arthur Priors Zeitlogik. Eine problemorientierte Darstellung*, which is available in electronic form (although only in German) at `http://www.fdm.uni-freiburg.de/~tmueller`, forthcoming as a book with Mentis Verlag, Paderborn, Germany, 2002.

[6] Apart from the concept of possibility based on reality (see below), one can at least distinguish the two broad notions of (1) logical and (2) physical possibility (cf. [2] for a critique of the latter notion) and two notions of possibility based on a given branching model: (3) Possibility in the given branching model: Ψ is possible in a model $\mathfrak{M}$ (in a history σ and at a point Y) iff there is a history $\eta \in \Sigma$ such that Ψ is satisfied in η at Y; (4) possibility in the given branching model based on accessibility (non-empty overlap of histories). Given our requirement on the sets of splitting points, this concept coincides with notion (3), since by historical connection, every history is accessible from (shares a common root with) every other history.

to get rid of the somewhat vague description given so far, here is the satisfaction clause explicitly:

$$\mathfrak{M}, \sigma, Y \models \Diamond\Psi \quad \text{iff there is } \eta \in \Sigma \text{ s.t. } (Y \in R_{\sigma,\eta} \quad \text{and} \quad \mathfrak{M}, \eta, Y \models \Psi)\,.^{7} \tag{19}$$

Example: At the source of a quantum correlation experiment (point Y = location of source), all outcomes $L\alpha+$ (left wing: setting α, outcome $+$), $L\beta-$ (setting β, outcome $-$) etc. are possible in this sense. After the setting α has been selected on the left (point Y above the selection event in an α-branch), $L\beta-$ is *no longer possible* (cf. Figure 2).

3.5 Formal semantics for the counterfactual connective

Like possibility, the counterfactual conditional is a modal concept. Unlike possibility based on reality, however, the counterfactual by its very name refers to a contrary-to-fact scenario. In line with the standard analysis due to David Lewis [5], a counterfactual "$\Psi \;\Box\!\!\rightarrow\; \Psi'$", to be read "if Ψ were the case, Ψ' would be the case", is true iff in the accessible history in which Ψ holds that is most similar to the actual one, Ψ' holds as well.[8] If there is no accessible history in which Ψ holds, the counterfactual is counted as vacuously true.

The given definition is based on two notions that need to be elucidated: (1) When is a history "accessible" from the actual one? (2) Which notion of similarity is applicable? The standard analysis acknowledges that the counterfactual is based on vague notions (cf. [5, p. 1]). Our formal framework, however, allows for some clarification of these two questions. (1) Two histories σ, η are accessible one from the other iff they have a non-empty region of overlap, i.e., iff they split off from some common root. Given historical connection, all histories are accessible one from another. (2) The framework of pasted Minkowski space-times allows for the definition of a variety of notions of comparative similarity. The formal semantics can be given without opting for one of these notions. We then have to decide afterwards, working back from the set of formulae that become valid under each of these notions.[9]

[7] Causal tense operators can be defined in a similar way as the modal operators, e.g., for the past-tense operator P: $\mathfrak{M}, \sigma, Y \models \mathsf{P}\Psi$ iff there is $X \in \mathbb{R}^4$ s.t. $X <_M Y$ and $\mathfrak{M}, \sigma, X \models \Psi$. (The future tense operator F can be defined entirely analogously, requiring $Y <_M X$ instead.) I do not believe that the causal tense operators capture enough of our tensed talk. More useful tense operators could be based on an absolute simultaneity relation. Arguments in favor of this approach are given in Chapter 4.4 of my dissertation (cf. note 5 above).

[8] One does not need to assume that there is a unique such history; the definition given below allows for ties in the appropriate way (cf. [5, p. 19]) and will also deal with the case of infinitely many histories without a closest one (cf. [6]). — If Ψ is in fact true, the counterfactual is inappropriate from a pragmatic point of view. Our analysis will assume so-called "centering" (cf. [5, p. 14]), i.e., the actual situation will be counted as the situation most similar to itself, so that a counterfactual "$\Psi \;\Box\!\!\rightarrow \Psi'$", where Ψ is true, is counted as true iff Ψ' is also actually true.

[9] Invariably, this step involves an appeal to intuition. I am not troubled by this, since I assume that formal languages must be built by working "from within" natural language. This doctrine was held, e.g., by Quine, but perhaps most forcefully by Prior. I have tried to argue for this doctrine in Chapter 3 of my dissertation (cf. note 5 above).

Two basic approaches to a definition of comparative similarity can be distinguished. The first approach tries to establish a global similarity ordering of histories. In Section 3.5.1 it will be shown that this approach is feasible and leads to a variety of notions of comparative similarity, but the global similarity orderings thus defined do not fulfill all the formal requirements of Lewis' standard analysis. The second approach, which leads to formally satisfactory similarity orderings, defines the orderings relative to a point of evaluation. This approach is sketched in Section 3.5.2. The formal definition of the counterfactual conditional is given in Section 3.5.3. Finally, the notions of comparative similarity are assessed in Section 3.5.4.

3.5.1 Global comparative similarity

Within the present framework, three families of partial orderings $\sqsubset$ of histories can be defined. $\eta \sqsubset_\sigma \gamma$ is to mean "η is more similar to σ than γ". A formal definition of the counterfactual can be based on this notion.

The leading intuition behind the definitions to be given is that the later histories split, the more similar they are. Before giving the full definitions, we will first discuss four typical cases.[10]

1. σ and η split only at c_1 and σ and γ split only a c_2. Here the situation is simple: η is more similar to σ than γ iff $c_1 > c_2$.

2. $C_{\sigma,\eta}$ and $C_{\sigma,\gamma}$ are such that

$$\forall x \in C_{\sigma,\eta} \forall y \in C_{\sigma,\gamma} \;\; x > y.$$

   Here the situation is again simple: η is more similar to σ than γ.

3. Problematic

$$C_{\sigma,\eta} = \{x_1, y_1\}, C_{\sigma,\gamma} = \{x_2, y_2\},$$
$$x_1 < x_2 \text{ but } y_1 = y_2.$$

4. Problematic:

$$C_{\sigma,\eta} = \{x_1, y_1\}, C_{\sigma,\gamma} = \{x_2, y_2\},$$
$$x_1 < x_2 \text{ but } y_1 > y_2$$

Depending on how one decides the question in such problematic cases, one can give three notions of strict global comparative similarity that all have the required formal property of being a partial ordering (i.e., transitive and antisymmetric). Let three histories σ, η, and γ be given, with their respective sets of splitting points $C_{\sigma,\eta}$ and $C_{\sigma,\gamma}$. (The set $C_{\eta,\gamma}$ will play no role, since we are only interested in similarity with σ.) The question of similarity can be decided as follows:

[10] The definitions in this section are based on Tomasz Placek's definitions (cf. [8], chap. 6.5 and his contribution to the conference "Quantum Structures V", Cesenatico, Italy, April 2001). Thanks to T.P. for allowing me to reproduce some material of his in this section.

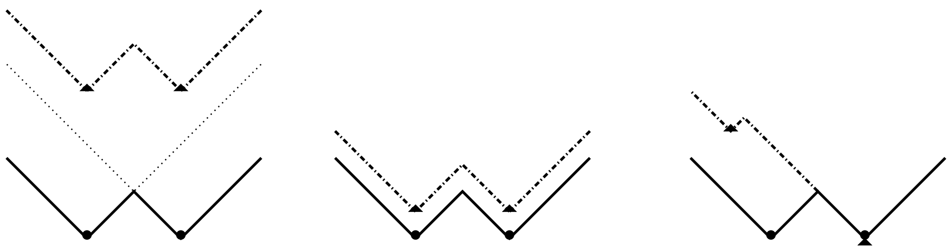

Figure 1: Illustration of the three notions of strict global comparative similarity. Circles represent splitting points of σ and γ, triangles represent splitting points of σ and η. By strong comparative similarity, only in the A history is η more similar to σ than γ. By mild comparative similarity, η is more similar to σ than γ in the A and B histories, but not in the C history. By weak comparative similarity, η is more similar to σ than γ in all three histories.

DEFINITION 2 (STRONG STRICT GLOBAL COMPARATIVE SIMILARITY)
η is more similar to σ than γ in the strong sense ($\eta \sqsubset_\sigma^S \gamma$) iff $C_{\sigma,\gamma}$ totally causally precedes $C_{\sigma,\eta}$, i.e., iff $\forall x \in C_{\sigma,\gamma} \ \ \forall y \in C_{\sigma,\eta} \ \ x <_M y$.

DEFINITION 3 (MILD STRICT GLOBAL COMPARATIVE SIMILARITY)
η is more similar to σ than γ in the mild sense ($\eta \sqsubset_\sigma^M \gamma$) iff $\forall x \in C_{\sigma,\eta} \exists y \in C_{\sigma,\gamma} \ y <_M x$.

DEFINITION 4 (WEAK STRICT GLOBAL COMPARATIVE SIMILARITY)
η is more similar to σ than γ in the weak sense ($\eta \sqsubset_\sigma^W \gamma$) iff $\forall x \in C_{\sigma,\eta} \exists y \in C_{\sigma,\gamma} \ y \leq_M x$ and for some $x' \in C_{\sigma,\eta}, y' \in C_{\sigma,\gamma} \ y' <_M x'$.

The formal properties of these three relations can be read off from the definition (i.e., transitivity and antisymmetry follow from the corresponding properties of $<_M$). The force of the respective definitions is illustrated in Figure 1.

The three orderings thus defined are partial orderings, but the relation $\sim^*$ ($* = S, M, W$), where $\eta \sim_\sigma^* \gamma$ iff neither $\eta \sqsubset_\sigma^* \gamma$ nor $\gamma \sqsubset_\sigma^* \eta$, is not transitive, since the relation SLR of being space-like separated fails to be transitive. Lewis' original definition of the counterfactual [5, pp. 48ff.] is based on a similarity preordering (i.e., an ordering that is transitive and connected). If one uses one of the partial orderings defined in this Section to define the counterfactual conditional, the formal properties of the conditional will be different from those of Lewis' standard analysis. It is not certain that this is a disadvantage, since some of Lewis' rules of counterfactual inference have been questioned independently, leading to the demand that the counterfactual be based on a partial ordering of histories (cf. [10]).

3.5.2 Local comparative similarity

In the previous Section we have not been able to follow Lewis in defining a global similarity preordering of histories. However, it is possible to define transitive and connected similarity orderings of histories by taking into account the space-time point Y at which a formula is evaluated. The basic idea, which can also be found in Lewis [5, pp. 50ff], is to define a real-valued similarity measure. We will define a local measure $D(\sigma, Y, \eta)$ that somehow represents numerically the distance between histories σ and η, *as judged from space-time point* Y. Once such a measure is given, a transitive and connected similarity ordering can be defined via

$$\eta \sqsubseteq_{\sigma,Y} \gamma \quad \text{iff} \quad D(\sigma, Y, \eta) \leq D(\sigma, Y, \gamma). \tag{20}$$

The crucial question is how such a distance measure can be introduced. We have not been able to come up with a fully satisfactory proposal. A first try might be to use the distance of splitting points from Y:

$$D(\sigma, Y, \eta) = - \sum_{Z \in C_{\sigma,\eta}} |Y - Z|, \tag{21}$$

where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^4$. (This definition is frame-dependent, but using the Lorentz-invariant Minkowskian space-time distance $D^2_M(Y, Z)$ will lead to counterintuitive results, as this distance can be negative.)

A version that is closer to the guiding idea of the previous section, i.e., the later histories split, the more similar they are, is given by

$$D(\sigma, Y, \eta) = \sum_{Z \in C_{\sigma,\eta}} d(Y, Z), \quad \text{where} \quad d(X, Z) = \begin{cases} |Y - Z| & \text{if } Z <_M Y, \\ 1/|Y - Z| & \text{otherwise.} \end{cases} \tag{22}$$

Any such measure will give the required formal properties for the similarity preordering, but no case for one specific definition of $D(\sigma, Y, \eta)$ can be made at this point.[11]

3.5.3 Formal definition of the counterfactual

A counterfactual $\Psi \;\Box\!\!\rightarrow \Psi'$ is to be evaluated in a history σ and at a space-time point Y. In line with Lewis' definition, we will use two auxiliary notions: the set Σ_σ of histories accessible from σ, which, due to historical connection, is simply $\Sigma_\sigma = \Sigma$, and the partial ordering $\sqsubset_\sigma$ from the family $\sqsubset$ defined above in Section 3.5.1, where $\eta \sqsubset_\sigma \gamma$ means "history η is more similar to σ than γ". In the following, "Ψ-history" stands for "history η such that $\mathfrak{M}, \eta, Y \models \Psi$". The formal definition of the counterfactual then reads in the case of finitely many histories:

$$\mathfrak{M}, \sigma, Y \models \Psi \;\Box\!\!\rightarrow \Psi' \quad \text{iff} \quad \text{in all closest } \Psi\text{-histories } \eta \in \Sigma_\sigma, \;\; \mathfrak{M}, \eta, Y \models \Psi'. \tag{23}$$

[11] Almost all available distance measures are frame-dependent. I assume that this fact calls for a unification of the formal semantics given here with a formal logic of points of view that accounts for changing the reference frame (cf. note 5 above).

For infinitely many histories, the following definition is appropriate (cf. [6]):

$$\begin{aligned}
\mathfrak{M}, \sigma, Y \models \Psi \;\Box\!\!\rightarrow\; \Psi' \quad \text{iff} \\
\text{for all } \Psi\text{-histories } \eta \in \Sigma_\sigma \text{ there is a } \Psi\text{-history } \gamma \in \Sigma_\sigma \text{ s.t.} \\
\text{(i) } \gamma = \eta \text{ or } \gamma \sqsubset_\sigma \eta \text{ and} \\
\text{(ii) for any } \Psi\text{-history } \delta \in \Sigma_\sigma \text{ s.t. } \delta = \gamma \text{ or } \delta \sqsubset_\sigma \gamma,\ \mathfrak{M}, \delta, Y \models \Psi'.
\end{aligned} \tag{24}$$

If a transitive local notion of comparative similarity from Section 3.5.2 is used, this definition can be simplified to the following:

$$\begin{aligned}
\mathfrak{M}, \sigma, Y \models \Psi \;\Box\!\!\rightarrow\; \Psi' \quad \text{iff} \\
\text{either there is no } \Psi\text{-world or} \\
\text{there is a } \Psi\text{-world } \eta \text{ s.t. for all } \Psi\text{-worlds } \gamma : \\
\text{if } \gamma \sqsubseteq_{\sigma,Y} \eta \text{ then } \mathfrak{M}, \gamma, Y \models \Psi'.
\end{aligned} \tag{25}$$

3.5.4 Assessment of the three global notions of similarity

Which notion of comparative similarity should be employed in the definition of the counterfactual? As the definitions of Section 3.5.2 are not fully satisfactory yet, the discussion will be limited to the global partial similarity orderings $\sqsubset^S$, $\sqsubset^M$ and $\sqsubset^W$ from Section 3.5.1.

In weighing these three alternatives, two considerations pull in different directions: (1) The definition should be obvious. Looking at Figure 1, it seems obvious that η is more similar to σ than γ in cases A and B, but it does not seem so obvious in case C. (2) The definition should not create too many ties. Recall that if the definition does not apply, two histories are rated "equally similar". If one opts for strong comparative similarity, case B is left undecided — even though we just suggested that it was obvious how things are in that case. Leaving case B undecided is therefore not acceptable: strong comparative similarity is out. Thus, one has to decide between mild and weak comparative similarity. The mild definition is more obvious, but the weak one decides more cases. Which one should be used?

Here formal considerations will not help any more. Thus, well-argued intuition is called for. I assume the following story should help (and I know of no way apart from telling stories that could help):

I flip a coin and bet on tails. The coin shows heads. I cry out: "Had I bet on heads, I would have won." Is this correct? It is important to stress that my counterfactual "had I bet on heads" refers to the *actual* situation — it does not mean "had a coin been flipped, and had I bet on heads…", but "had I, in this situation, bet on heads".

My intuition is quite clearly that a situation in which the coin does in fact show heads again is more similar to the actual one that one in which the coins shows tails. The situation was: I bet on tails, the coin shows heads. Changing this to me betting on heads, but the coin still showing heads is obviously less of a change than also changing the outcome of the toss. Thus, had I bet on heads, I would really have won. (Nothing I can do about that now, of course.)

For the setup described, case C from figure 1 applies: my choice (left splitting point) is different, but the chance event of the coin landing heads or tails up is left unchanged (right splitting point). If this situation has to be counted as a case of "η (I: heads, coin: heads) more similar to σ (actually: I: tails, coin: heads) than γ (I: heads, coin: tails)", then both strong and mild comparative similarity are out. This means that the weak notion of comparative similarity $\sqsubset^W$ from definition 4 should be used in the definition of the counterfactual conditional.

4 Application to Stapp's purported proof

One of the most puzzling features of quantum mechanics is the fact that a system consisting of more than two particles can be in a state that cannot be described through the state of the individual particles alone. Such so-called entangled states can give rise to correlations between space-like separated events, for which a causal connection is rules out by special relativity. There has been a long debate as to whether the existence of such correlations points to a failure of special relativity, to the incompleteness of quantum mechanics or to something else. It is generally agreed that the (experimentally confirmed) existence of the mentioned correlations *together with* assumptions about an underlying mechanism "behind" the correlations (so-called hidden variables) stands in conflict with the locality requirement of special relativity. This leaves open two possibilities: Either quantum mechanics is non-local, or the assumption that there are hidden variables is false.

Henry Stapp has recently given a formal proof that claims to derive non-locality from orthodox quantum mechanics alone, without an appeal to any notions of hidden variables [14]. If his proof were correct, this would constitute a major achievement. We will now analyze two premises of the proof with the aid of the formal semantics introduced in the previous Section.

A relevant part of the proof is given below. The lines whose labels start with "S" are Stapp's own (with some slight changes of notation to enhance readability); lines 15 and 16 were added to make explicit the formal contradiction that Stapp claims to have derived.

Stapp's proof is based on the quantum mechanical properties of the Hardy state [4]. In a Hardy experiment, particle pairs are created at a source S. The two particles travel to the left (L) and right (R) wing of the experimental setup, where polarizers can be set to angles α or β in order to measure spin projections of the respective particle. The outcome of a spin measurement is either "$+$" or "$-$". Thus, "$L\alpha+$" stands for the proposition "in the left wing, the polarizer was set to α, and the measurement outcome was $+$." The experiment is pictured schematically in Figure 2. Particle pairs in the Hardy state will give rise to a number of correlations: (1) If $L\alpha$ was selected and had outcome $-$, then if $R\beta$ was selected, the outcome is $+$. (2) If $L\beta$ was selected and had outcome $+$, then if $R\alpha$ was selected, the outcome is $-$. (3) If $R\beta$ was selected and had outcome $+$, then if $L\beta$ was selected, the outcome is $+$.

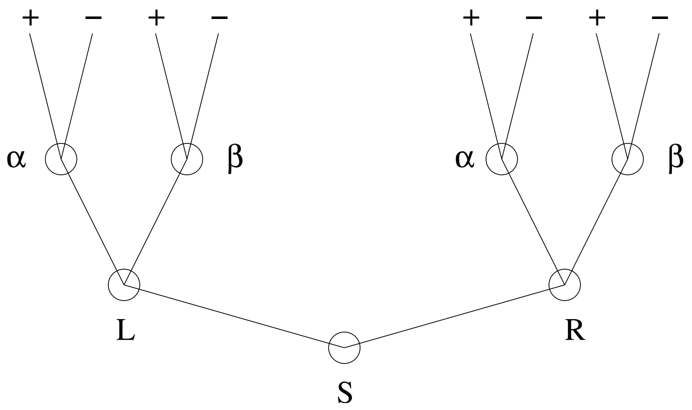

Figure 2: Schematic drawing of the Hardy experiment. S: source of particle pairs; L, R: left and right wing of the experiment, resp.; α, β: the two possible settings of the polarizers in both wings; $+, -$: possible spin measurement outcomes.

$$L\beta \wedge R\beta \wedge L\beta+ \;\mathrel{-\!\!3}\; (R\alpha \;\Box\!\!\rightarrow L\beta \wedge R\alpha \wedge L\beta+) \qquad \text{(LOC 1} \quad \text{S1)}$$

$$L\beta \wedge R\beta \wedge R\beta+ \;\mathrel{-\!\!3}\; L\beta \wedge R\beta \wedge L\beta+ \qquad \text{(QM} \quad \text{S2)}$$

$$L\beta \wedge R\alpha \wedge L\beta+ \;\mathrel{-\!\!3}\; L\beta \wedge R\alpha \wedge R\alpha- \qquad \text{(QM} \quad \text{S3)}$$

$$L\beta \wedge R\beta \wedge R\beta+ \;\mathrel{-\!\!3}\; (R\alpha \;\Box\!\!\rightarrow L\beta \wedge R\alpha \wedge R\alpha-) \qquad \text{(LOGIC, S1, S2, S3} \quad \text{S4)}$$

$$L\beta \mathrel{-\!\!3} (R\beta \wedge R\beta+ \;\mathrel{-\!\!3}\; (R\alpha \;\Box\!\!\rightarrow R\alpha-)) \qquad \text{(LOGIC ?, from S4} \quad \text{S5)}$$

$$L\alpha \mathrel{-\!\!3} (R\beta \wedge R\beta+ \;\mathrel{-\!\!3}\; (R\alpha \;\Box\!\!\rightarrow R\alpha-)) \qquad \text{(LOC 2, S5} \quad \text{S6)}$$

$$\ldots$$

$$\neg\Diamond(L\alpha \wedge R\beta) \qquad \text{(LOGIC, S11, S14} \quad \text{15)}$$

$$\Diamond(L\alpha \wedge R\beta) \qquad \text{(FREE CHOICE} \quad \text{16)}$$

In its published form [14], the proof contains formal errors resulting from an uncautionary mixing of strict and material conditionals (cf., e.g., the inference from line S4 to line S5, which is valid only if the second conditional in line S5 is changed into a material one).[12] These problems can however be circumvented.For an assessment of the proof, it is conceptually most important to analyze the premises LOC1 and LOC2 that Stapp refers to in lines S1 and S6.

[12]A very detailed and careful analysis of Stapp's proof is given in [12; 13]. Their discussion is not based on an explicitly given formal semantics, but their results agree with the results derived here. — Thus, I do not claim that the formal machinery introduced is a necessary ingredient in an analysis of Stapp's purported proof, nor do I claim that the result of this analysis is new. However, my analysis is formally rigorous and takes up little space. It is offered here as a first application of the formal language framework.

LOC1 Stapp's premise LOC1 "asserts that if under the condition that the choices were $L2$ [setting on the left] and $R2$ [on the right] the outcome in **L** at some earlier time were $L2+$, then if the (later) choice in **R** were to be $R1$, instead of $R2$, but the free choice in **L** were to remain unchanged, then the outcome $L2+$ in **L** would likewise remain unchanged" [14, p. 301]. Counterfactual statements of this kind are true if the counterfactual is based on the weak global notion of comparative similarity, but false on the mild and strong global notion: As there is a choice point common to both the actual and the counterfactual history, only definition 4 applies, being the only one of the three definitions that allows for ties in the set of splitting points (via $\leq$ instead of $<$ as in the other definitions). Thus, the validity of Stapp's first premise depends on a fine detail of the modal semantics.[13] As we have argued that the weak notion of comparative similarity is the correct one, LOC1 is seen to be true.[14]

LOC2 Stapp claims that the transition from statement S5 to statement S6 is justified by a rule of inference that he calls LOC2. Stapp gives the following description of this rule: "LOC2 asserts this: 'If SR [the statement to the right of the first strict conditional in S5, which makes assertions about the right wing of the experiment only] is proved to be true under the condition that [$L\beta$] is freely chosen in [the left wing of the experiment], then SR must be true also under the condition that [$L\alpha$] is freely chosen there instead' " [14, p. 302; glosses and slight change of notation T.M.]. In the form given in the published proof, the rule LOC2 is inappropriate, again due to a mixing up of strict and material conditionals. It can however be amended by changing the second conditional in lines S5 and S6 to a material one. Call the amended lines S5' and S6', resp. The amended rule that allows one to infer S6' from S5' will be called LOC2'.[15]

LOC2' is incorrect, as can be shown by a simple counter-model. We take the branching model $\langle B, \leq_S \rangle$ to be the model of the Hardy experiment [4] pictured in Figure 2, derived from the description of the experiment and its outcomes via the pasting construction of Section 2. Evaluating the lines of the proof at the location of the source and in any history, the antecedent of the purported rule of inference LOC2', statement S5', is true, while the consequent, statement S6', is false. This one counterexample shows that, contrary to the intuitive motivation given by Stapp in his paper, LOC2' cannot be a valid rule of inference. The purported proof thus does not show that quantum mechanics of itself is non-local.

Summing up: Stapp is not justified in his claim. In order to show a conflict between quantum mechanics and the concept of relativistic locality, assumptions about hidden variables still play a vital role.

[13] Thanks to Jeremy Butterfield for a discussion of this point.

[14] If the counterfactual is to be based on a local definition of comparative similarity, then that definition will also have to respect the intuitions spelled out in the coin story above. Thus, LOC1 will be valid on any acceptable notion of local comparative similarity as well.

[15] Thanks to Niko Strobach for a discussion on the role of LOC2.

5 Summary

In this paper, models of Belnap's branching space-time were constructed by pasting states defined on Minkowski space-time. The branching framework was used to give a formal semantics for a formal propositional language containing modal and counterfactual connectives. An appeal to intuition was used to justify a weak notion of comparative similarity as the basis for the definition of the counterfactual. The formal framework was employed to analyze two premises of Stapp's purported proof of non-locality. While premise LOC1 was justified, premise LOC2 was seen to be faulty, rendering Stapp's proof invalid.

Acknowledgments

This paper has been much inspired by work carried out jointly with Tomasz Placek, whom I wish to thank for numerous helpful discussions. I have also benefited greatly from comments on a previous version of this paper that I delivered at the workshop "Modality, Probability, and Bell's Theorems", Cracow, Poland, August 2001. Special thanks to Nuel Belnap, Jeremy Butterfield, and Tomasz Placek.

References

[1] Belnap, N.D. (1992). Branching space-time. *Synthese*, 92:385–434.

[2] Bressan, A. (1980). On physical possibility. In M. L. Dalla Chiara, ed., *Italian Studies in the Philosophy of Science*, pages 197–210. Dordrecht: D. Reidel.

[3] Finkelstein, J. (1999). Space-time counterfactuals. *Synthese*, 119:287–298.

[4] Hardy, L. (1992). Quantum-mechanics, local realistic theories and Lorentz-invariant realistic theories. *Physical Review Letters*, 68:2981–2984.

[5] Lewis, D.K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press, 1973.

[6] Lewis, D.K. (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic*, 10:217–234.

[7] Müller, T. and T. Placek (2001). Against a minimalist reading of Bell's theorem: Lessons from Fine. *Synthese*, 128:343–379.

[8] Placek, T. (2000). *Is Nature Deterministic?* Cracow: Jagiellonian University Press.

[9] Placek, T. (2000). Stochastic outcomes in branching space-time. An analysis of Bell's theorem. *British Journal for the Philosophy of Science*, 51:445–475.

[10] Pollock, J.L. (1976). *Subjunctive Reasoning*, Dordrecht: D. Reidel.

[11] Prior, A.N. (1968). Tense logic and the logic of earlier and later. In *Papers on Time and Tense*. Oxford: Clarendon Press, pp. 116–134.

[12] Shimony, A. and H. Stein (2001). Comment on “Nonlocal character of quantum theory,” by Henry P. Stapp. *American Journal of Physics*, 69:848–853.

[13] Shimony, A. and H. Stein (2001). On quantum non-locality, special relativity, and counterfactual reasoning. In Ashtekar, A., R.S. Cohen, D. Howard, J. Renn, S. Sakar, and A. Shimony, eds., *Space-Time, Quantum Entanglement and Critical Epistemology. Essays in Honour of John Stachel*. Dordrecht: Kluwer, forthcoming.

[14] Stapp, H.P. (1997). Nonlocal character of quantum theory. *American Journal of Physics*, 65:300–304.

[15] Xu, M. (1997). Causation in branching time (I): Transitions, events and causes. *Synthese*, 112:137–192.