Erschienen in: Explanation, prediction, and confirmation / Dennis Dieks ..., editors. - Dordrecht [u.a.] : Springer, 2011. - S. 109-121. - (The Philosophy of Science in a European Perspective ; 2). - ISBN 978-94-007-1179-2

THOMAS MÜLLER

PROBABILITIES IN BRANCHING STRUCTURES

A common, natural view about probabilities, shared by philosophers of diverse persuasions, is that probabilities are *graded possibilities*.[1] On this view, which I will presuppose, there are no probabilities without underlying possibilities, and there is room for different notions of probability at least to the extent that there are different underlying notions of possibility.

In this paper I want to trace out consequences of this view for the specific case of possibilities that are grounded in branching structures. Such structures afford a natural representation of *real possibilities*: concrete possibilities in indexically specifiable situations, like a concrete laboratory experiment or other indeterministic happenings. I will argue that branching-based probabilities lead to interesting probability structures that can shed light on, e.g., the causal Markov condition.

My paper is structured as follows: I start by discussing different types of possibility and zooming in on branching-based real possibilities, giving formal details about the theory of branching time. I then work out a theory of probabilities based on branching time structures and discuss a number of peculiarities and limitations of that approach. Finally I give a brief overview of the way in which the branching time framework is extended in the theory of branching space-times, and what an appropriate probability theory for that framework looks like.

REAL AND OTHER POSSIBILITIES

Is there just one correct analysis of the notion of possibility, or can one discern several different notions of possibility that, while sharing some common structure, are interestingly different? The logical empiricists of the 1930s held that the only viable notion of possibility was the ontologically rather innocent notion of logical possibility: the absence of formal contradiction. On the other hand, Weyl (1940) in his overview of the phenomenology of modality already distinguished several different notions of possibility and pointed out that they had their proper uses in specific contexts: apart from logical possibility he acknowledged *a priori* (conceptual) possibility, physical possibility, and technical possibility. Present-day discussions involving modality tend to focus on the notion of *metaphysical* possibility as the one notion underlying philosophical argumentation; most thought experiments in philosophy, from Descartes's *genius malignus* to Putnam's *Twin Earth*, are meant to exhibit metaphysical possibilities. The formal semantics for possibilities developed since the 1950s however shows that there is much room for

1 Cf., e.g., van Fraassen (1980, 198); Popper (1982, 70); Thomason (1984).

different accounts of possibility. How do these different notions of possibility hang together? Fine (2005) has argued convincingly that mathematical and logical possibility can be derived from the notion of metaphysical possibility, but he has also argued that physical possibility is a different matter—he shows that physical (in his terminology, "natural") possibility is a further kind of modality not reducible to the metaphysical variety. The landscape of possibilities seems to be more of a jungle than a desert.

All the notions of modality mentioned so far are abstract: they pertain to *types* of situations, not directly to concrete token-happenings. Possibilities that are important for our practical concerns are however different: they are concrete rather than abstract (I can walk to the market, along a specific route, or I can take my bike)—and as we all know, they vanish as time passes. What was possible today may not be possible any more tomorrow (the last chance to see a live Dodo has long passed; once that specific train is gone, it is impossible for me to catch it). Such *real possibilities* are at least epistemically more basic than the other, abstract notions: we encounter them in our daily lives, and they are part of our everyday conception of agency.

Branching time: a model for real possibilities

Formal models for real possibilities were developed starting in the late 1960s, in the context of semantical investigations of tense and other temporal modifiers. Arthur Prior (1957, 1967) developed his tense-logic as a formal framework in which the interaction of time and modality could be studied. The so-called Ockhamist analysis of the future tense results in a formal tempo-modal language whose models have the structure of forward-branching trees; cf. Figure 1. Formally, a *branching-time structure* is a pair $\langle T, < \rangle$ such that:

- T is a nonempty set;
- $<$ is a strict partial ordering on T, i.e., transitive (if $x < y$ and $y < z$, then $x < z$) and asymmetric (if $x < y$ then $y \not< x$), hence also irreflexive ($x \not< x$);
- $<$ is backward linear, i.e., if $x < z$ and $y < z$, then ($x < y$ or $x = y$ or $y < x$).

The elements of T are usually called *moments*, and the ordering relation $<$ is read tempo-modally: if $x < y$, we say that x occurs before y, or that y is in the future of possibilities of x. A non-strict order $\leq$ is derived from $<$ in the usual way, i.e., $x \leq y$ iff ($x < y$ or $x = y$).

In such structures one can single out so-called *histories* as maximal linear subsets (i.e., subsets h of T that are maximal w.r.t. the property that for any $x, y \in h$, either $x < y$ or $x = y$ or $y < x$). In a history h, any two distinct elements are comparable via $<$, while in general, T will contain incomparable elements (cf.,

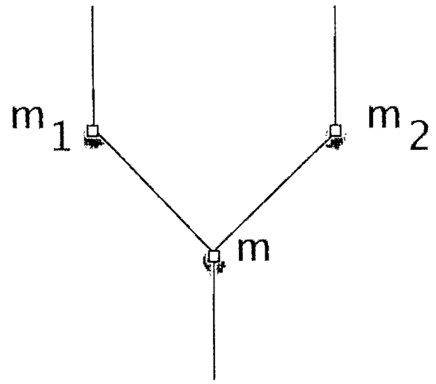

Figure 1: A branching structure

e.g., the moments m_1 and m_2 in Figure 1); such elements do not have a history in common. In view of the interpretation of the ordering relation, we can say that incomparable elements are *inconsistent*, and that a history represents a maximal possible course of events—a possible way the world depicted by $\langle T, < \rangle$ could develop.

The formal requirements for branching time structures laid out above form the bare minimum, and additional requirements may be useful. For example, in the important case of continuous structures it turns out to be convenient, for technical reasons, to constrain the topology of the branching. Basically there are two options: two histories could split such that there is a last moment of overlap and no first moment of difference, or they could have no last moment of overlap, but two different first moments of difference. The convenient choice is to have a last moment of overlap—cf. McCall (1990) for discussion. Formally we thus require:

- If $m \in h_1 - h_2$ and $m' \in h_2 - h_1$, then there is some m^* for which $m^* < m$, $m^* < m'$, and m^* is maximal in $h_1 \cap h_2$.

Note that this requirement also guarantees that any two histories are connected at some moment; any two possible developments of the world thus share a common past.

In order to allow for some small technical simplifications in what follows, we also require that the ordering $\langle T, < \rangle$ have no maximal moments:

- For every $m \in T$ there is $m' \in T$ s.t. $m < m'$.

Formal semantics for branching time

In order to explain the notion of real possibility in branching structures, and thus, the background for our discussion of probability, it will be good to have some details about the formal language that is customarily defined for branching time. The Prior-Thomason semantics for such branching treelike structures defines the following temporal and modal operators: The past tense ("it was the case that", P), the future tense ("it will be the case that", F), and the operators of real possibility and necessity (or settledness), Poss and Sett. As is customary in modal languages, formulae are to be evaluated with respect to a model $\mathfrak{M}$ (a branching structure $\langle T, < \rangle$ together with a valuation V assigning extensions to atomic propositions) at a point of evaluation. In basic modal logic such a point of evaluation consists just of a simple index (a "world"), so that one could expect the point of evaluation in branching time to be a simple index as well, i.e., a moment. The innovative element of Prior-Thomason branching time semantics is, however, to take the point of evaluation to be a moment and a history through that moment (written m/h, presupposing $m \in h$). For a stand-alone sentence uttered in a given context (formally: at a moment of context $m_C \in T$), the parameter m is initialized as $m = m_C$. This indexical link to a context makes the Prior-Thomason semantics well suited to represent real tenses and possibilities. A specific problem is that there is no similar initialization of the history parameter: metaphysically speaking, there is no "history of the context" singling out "the real future"—the future is open. We do not dwell on the difficulties of interpreting stand-alone sentences containing the future tense. As the philosophical discussion of the notion of an open future shows, these difficulties are real; they are not an artefact of the formal analysis via branching time. For an illuminating account of how these difficulties can be met formally, cf. Belnap (2002a). For the account of real possibility that is at issue here, however, the mentioned difficulty plays no significant role: the semantic clause for Poss (as well as the dual one for Sett) introduces a quantifier over histories up front, so that it makes no difference which history containing the context moment m_C is assigned to h initially.

Intuitively speaking, the tense operators P and F move the moment of evaluation along the current history of evaluation, in accordance with the idea that tense encodes location in time. The modal operators Sett and Poss, on the other hand, do not change the moment of evaluation, but the history, thus quantifying over the possible futures of the moment of evaluation. The formal semantical clauses for the mentioned operators are as follows:

- $\mathfrak{M}, m/h \models \mathsf{P}\phi$ iff there is $m' \in h$ s.t. $m' < m$ and $\mathfrak{M}, m'/h \models \phi$;
- $\mathfrak{M}, m/h \models \mathsf{F}\phi$ iff there is $m' \in h$ s.t. $m < m'$ and $\mathfrak{M}, m'/h \models \phi$;
- $\mathfrak{M}, m/h \models \mathsf{Sett}\,\phi$ iff for all h' for which $m \in h'$, $\mathfrak{M}, m/h' \models \phi$;
- $\mathfrak{M}, m/h \models \mathsf{Poss}\,\phi$ iff there is some h' for which $m \in h'$ and $\mathfrak{M}, m/h' \models \phi$.[2]

2 As usual, this clause for possibility is the dual to the clause for necessity, so an explicit definition could be avoided by declaring Poss to be an abbreviation for ¬Sett¬.

This semantics suggests a natural definition of *real possibilities at a moment*, which is based on the notion of *division of histories at a moment*:

- Let m belong to histories h_1 and h_2, i.e., $m \in h_1 \cap h_2$. h_1 and h_2 are called *undivided at* m (written $h_1 \equiv_m h_2$) iff there is some $m' \in h_1 \cap h_2$ for which $m < m'$. Being undivided at m is an equivalence relation on the set H_m of histories containing m. Reflexivity and symmetry are trivial.[3] For transitivity, let $h_1 \equiv_m h_2$ as witnessed by m', and $h_2 \equiv_m h_3$ as witnessed by m''. As $m', m'' \in h_2$, we have $m' \leq m''$ or $m'' \leq m'$ (by linearity of histories); assume the former. By backwards linearity we get $m' \in h_3$, so that m' also serves as a witness for $h_1 \equiv_m h_3$.
- We say that h_1 and h_2 *split at* m (written $h_1 \perp_m h_2$) iff m is maximal in $h_1 \cap h_2$.
- The *real possibilities at* m are the members of the partition Π_m of H_m induced by the equivalence relation of undividedness at m, $\equiv_m$.
- We say that m is a *choice point* iff Π_m has more than one member, i.e., if there are at least two histories splitting at m.

In order to have an easy way of talking about real possibilities in a branching structure, we introduce the notion of a *transition* (cf. von Wright, 1963; Belnap, 1999): a transition t is an ordered pair $\langle m, H \rangle$ consisting of an *initial* (a moment $m \in T$), together with an *outcome* (one of the real possibilities at that moment, $H \in \Pi_m$). Employing suggestive notation, we also write $t = m \rightarrowtail H$. A transition is *trivial* if there is no alternative transition with the same initial. We write TR for the set of all transitions, and TR_m for the set of all transitions with initial m. If m is not a choice point, then $TR_m = \{t\}$ with $t = m \rightarrowtail H_m$, a trivial transition. We will alternatively speak of Π_m and of TR_m as embodying the real possibilities open at m; these possibilities form an exhaustive set of mutually exclusive alternatives. A set of transitions $T = \{t_i \mid i \in I\}$, where $t_i = m_i \rightarrowtail H_i$, is *consistent* if all the outcomes can occur together in one history, i.e., if $\bigcap_{i \in I} H_i \neq \emptyset$.

Probabilities in Branching Time

How can we implement probabilities in branching time? What is the proper notion of probability based on real possibilities? Given that real possibilities at a moment m form a set Π_m corresponding to the set of transitions TR_m, a natural move is to

3 For reflexivity we require that there be no maximal moments. If T is allowed to have maximal moments, the definition of undividedness will have to treat those as a special case: h_1 and h_2 are undivided at $m \in h_1 \cap h_2$ iff, either, there is no $m' \in T$ s.t. $m < m'$ (in which case it turns out that $h_1 = h_2$), or there is some $m' \in h_1 \cap h_2$ s.t. $m < m'$.—A similar patch would be needed to improve the definition of splitting.

use the set TR_m as a sample space and to employ a suitable Boolean σ-algebra and some normalized measure to define a probability space.[4] Technically—in the finite case we are assuming—, such a probability space at m is a triple $PR_m = \langle \Omega, F, \mu \rangle$ where $\Omega = TR_m$ is the sample space, F is the set-theoretic Boolean algebra over Ω (i.e., the set of subsets of Ω with the usual set-theoretical operations), and μ is a normalized measure on F (i.e., $\mu(\Omega) = 1$ and for $a, b \in F$ with $a \cap b = \emptyset$, we have $\mu(a \cup b) = \mu(a) + \mu(b)$).

It seems plausible that in any case in which an exhaustive set of alternative possibilities is given, this will be the way to introduce probabilities as weights on these possibilities. What's so special about probabilities in branching structures then? It turns out that interesting questions arise once we try to *combine* probability spaces that are defined at different moments.

Normally, probability spaces PR_A, PR_B are combined by forming the Cartesian product $\Omega_{A,B} = \Omega_A \times \Omega_B$ of the sample spaces and $F_{A,B} = F_A \times F_B$ of the algebras, and by defining a joint measure $\mu_{A,B}$ that satisfies the *marginal property* (i.e., reduces to the single measures when plugging in a unit element):

$$PR_{A,B} = \langle \Omega_{A,B}, F_{A,B}, \mu_{A,B} \rangle \quad \text{s.t.}$$

$$\mu_{A,B}(\langle \mathbf{1}_A, b \rangle) = \mu_B(b); \quad \mu_{A,B}(\langle a, \mathbf{1}_B \rangle) = \mu_A(a).$$

This recipe rests on the presupposition that the underlying possibilities combine by forming Cartesian products. This is often the case, even if there are correlations (i.e., even if generally, $\mu_{A,B}(\langle a, b \rangle) \neq \mu_A(a) \cdot \mu_B(b)$). But in the case of branching structures, this assumption is not warranted. What is the technically appropriate way of defining something like joint probabilities in branching structures? How do the underlying (branching) possibilities constrain these probability spaces? And does the interpretation of the tree-like ordering in terms of an open future lead to any additional constraints on joint measures?

It turns out that most of the relevant observations can be made by discussing a few simple cases. When combining *two* nontrivial probability spaces PR_m and $PR_{m'}$ with $m \neq m'$, only two cases need to be discussed: we either have $m < m'$ or $m' < m$ (which cases can be discussed together as they coincide after relabeling), or m and m' are incomparable and thus inconsistent.

Combining PR_m with $PR_{m'}$, $m < m'$.

In this case the choice point m' lies in the future of possibilities of the choice point m. The Cartesian product recipe fails instructively: most combinations of results from PR_m and $PR_{m'}$ make no sense. In fact, since there is splitting at m that

4 In what follows, we will ignore well-known measure-theoretic moves that will have to be made in case the sample space is infinite. We will assume *finite branching* from now on, i.e., we will be working exclusively with branching structures $\langle T, < \rangle$ s.t. for all $m \in T$, TR_m is finite. (Note that this does not imply that T itself, or even all H_m, are finite.)

partitions the set H_m of histories through m, the moment m' occurs in only *one* of those partitions. (Otherwise it would constitute a witness for undividedness.) Thus there is just one $t = m \rightarrowtail H \in TR_m$ for which $m' \in \cup H$ (i.e., $H_{m'} \subseteq H$). Given some $t' = m' \rightarrowtail H' \in TR_{m'}$, we can understand the pair $\langle t, t' \rangle$ as expressing the fact that "first, choice point m had outcome H, enabling m', and then choice point m' had outcome H'". For some $\bar{t} \in TR_m - \{t\}$, however, the pair $\langle \bar{t}, t' \rangle$ makes no sense given the tempo-modal (if you wish, causal) interpretation of the ordering relation: there is no history in which first $\bar{t}$ happens, excluding m', and then m' has some outcome or other. Given $\bar{t}$, m' is no longer possible—even though up to (and including) m, it was possible.

This point is strengthened if we look at the options for assigning probabilities to pairs of transitions in the Cartesian product. If $\langle \bar{t}, t' \rangle$ cannot happen, then it should certainly have probability zero. By the marginal property, summing over all the alternative outcomes from $TR_{m'}$, this leads to $\mu_m(\{\bar{t}\}) = 0$, and this holds for all transitions $\bar{t} \in TR_m - \{t\}$. Accordingly, we need to have $\mu_m(\{t\}) = 1$, by normalization. This will usually not accord with the given probability measure from PR_m. Furthermore, by considering an incompatible alternative m'' to m', the same argument will lead to $\mu_m(\{t\}) = 0$, a contradiction. Dropping the marginal requirement seems quite a high prize to pay in order to avoid this—if one cannot recover the individual probabilities from the joint probability structure, in which sense can the latter be called a combination of the former any more? In the framework of the Cartesian product, the only sensible alternative seems to be to use the product measure, which fulfils the marginal property by definition. But then we have to be ready to assign non-zero probabilities to $\langle \bar{t}, t' \rangle$ and other impossible combinations. (Note that it is not an option to use the product measure only in the case in which the first component is t, and set the probabilities to be zero otherwise—such a measure will generally not be normalized.)

Isn't there a more sensible way to combine PR_m and $PR_{m'}$? In fact, moving from the abstract discussion above to a concrete example leads the way to a better framework. Assume that at the concrete moment m, I have the choice to go to the races (t) or to stay at home ($\bar{t}$), and assume for the sake of the argument that some suitable probability measure μ_m can be given.[5] Let m' be the moment at which I am at the races and bet \$5 on Silver Shadow, with outcomes win (t') and lose ($\bar{t'}$). There are two two-way splittings, but clearly, there are only three alternatives: staying at home, going and winning, and going and losing. Returning to the abstract framework and generalizing, the recipe is as follows:

- The alternatives that form the sample space of the combined probability space are *sets of transitions* from the individual sample spaces;
- only those sets that are *consistent* qualify as alternatives;

5 This should not be read as an endorsement that probabilities can be usefully assigned to the outcomes of a singular human decision. In fact this may not be possible. If you are worried, substitute your favourite quantum-mechanical set-up.

- in fact, the sample space consists of the set of maximal consistent sets of transitions.

This recipe explains why {stay at home, win} doesn't occur in the sample space and thus needn't be assigned a probability: the set is inconsistent. It also explains why {go} by itself is not one of the alternatives—the set is consistent, but not maximally so (it can be extended by "win" or by "lose"). The problems discussed above are hereby avoided. Furthermore insights are gained as to the marginal property and the question of correlations.

The marginal property. In the Cartesian product case, the marginal property can be motivated by the thought that if one ignores what happens in the other probability spaces ("elsewhere"), the probability measure under consideration ("the probabilities for a local outcome") should be recovered. The spatial metaphor is apt here: in fact the Cartesian product is well motivated if simultaneous outcomes of chance experiments are under consideration. Space (at least on a simple-minded, Newtonian conception) orders that which is simultaneous. Once a temporal dimension enters, however, things are different: Here the guiding intuition is that the *present* probabilities for outcomes should be recovered if we ignore *what happens later*. Accordingly, in the races example, we want to have

$$\mu_m(\{\text{go}\}) = \mu_{m,m'}(\{\text{go, win}\}) + \mu_{m,m'}(\{\text{go, lose}\}).$$

Correlations. Again, if we think about the Cartesian product case in terms of the spatial combination of alternatives, we have the intuition that we should normally expect to find the product measure: spatially separated happenings shouldn't influence one another. However, correlations are not unthinkable. To use Van Fraassen's slogan, correlations cry out for explanation—but such explanation can often be had, e.g., by signalling, or by a common cause. If the chance of Paul's wearing red socks on a given day is 1/7, and the same holds for Albert, but the probability of seeing them both wearing red socks is 1/7 instead of 1/49, we may well venture the guess that they have communicated about this, e.g., by declaring Monday to be Red Socks Day.

In the case of the temporal combination of probability spaces, correlations are much harder to make sense of. In fact, it seems that they always point to a flawed (and not just incomplete) model. If the chance to get from m to (later) m' is a, and the chance to get from m' to (later) m'' is b, but the chance to get from m to m'' isn't $a \cdot b$, what are we to say? Think in terms of frequencies: A certain fraction of a given population arrives at m', and from there again, a certain fraction arrives at m'', but the overall account doesn't fit together?

This consideration gives a strong motivation for the *causal Markov condition* in combining probabilities in branching structures: if m, m', and PR_m and $PR_{m'}$ are given and $m < m'$, then on the combined probability space of maximal consistent sets of transitions from TR_m and from $TR_{m'}$, the correct measure is the product measure.

Combining $PR_{m'}$ with $PR_{m''}$, m' and m'' incomparable.

In this case, no combination of an outcome of m' and an outcome of m'' is possible at all. Suppose that, continuing the races example, when staying at home I can read or cook, and that at m'', the moment of the decision, there is an appropriate probability space $PR_{m''}$ that captures the relevant probabilities. Thus, $t'' = m'' \rightarrowtail H''$ could stand for "read" and $\overline{t''}$ for "cook". The moments m' (of the decision at the races) and m'' (of the decision at home) are incomparable and thus inconsistent.[6] What would it mean to combine the two given probability spaces? Can any sense be made of a combination of, e.g., t' ("win") and t'' ("read")? It seems not, and it seems that in this case, even the strategy to look at maximal consistent sets of transitions will fail: In the union of $TR_{m'}$ and $TR_{m''}$, only the singletons are consistent sets of transitions—but if the given probabilities of the single spaces are assigned to these singletons, as we seem to have to, it turns out that normalization fails again: the sum of the respective probabilities is 2, not 1.

Again, shifting the discussion to a consideration of the underlying possibilities is helpful. Surely m' and m'' are inconsistent, but in branching structures as defined above, there is always a common past, a choice point at which the histories containing m' split off from those containing m''. In our story, this happens exactly at moment m, the moment of deciding whether to go to the races or to stay at home. If that moment is added in, the general recipe from the previous discussion can be used: the maximal consistent sets of transitions now contain one outcome of m and one of the outcomes of either m' or m'', and by the given recipe we arrive at a well-defined, normalized probability space. The general moral of this example is that inconsistent moments need to be supplemented by a common past choice point before a sensible joint probability structure can be defined.

Summary: Branching-time based probabilities

The two cases discussed above in fact cover everything that can happen in combining probabilities defined on branching time structures. Joint probabilities can be defined if the individual given probability spaces afford a combination that gives a full account of causal alternatives. No "raw" inconsistencies may occur. Somewhat more formally:

- Let a set M of nontrivial initials and for each $m \in M$ a probability space PR_m be given. (PR_m has the set of transitions with initial m as its sample space. By nontriviality, this sample space has more than one element.) A joint probability space PR_M can be defined if and only if for any incompatible $m', m'' \in M$, there is some $m \in M$ for which $m < m'$ and $m < m''$.

6 Note that these two incompatible moments could still occur at the same clock time. In order to avoid confusion it is therefore useful to call the elements of a branching structure "moments" and not "times". This is the terminology adapted in this paper.

Given such an M, the above recipe can be applied for combining the individual probability spaces:

- Let $TRM = \bigcup_{m \in M} TR_m$ be the set of all transitions with initials from M.
- The sample space of PR_M, Ω_M, is the set of maximal consistent subsets of TRM (where consistency, as stated above, means that all outcomes share at least one history). This is the set of causal alternatives that can be built up given initials from M. As one can check, in a branching time structure a set of transitions can be consistent only if its set of initials is linearly ordered (else there isn't even a history containing all the *initials*).
- In the finite case that we are assuming, the algebra is simply the algebra of subsets of that sample space.
- The probabilities of the (singletons of) elements of the sample space Ω_M are given by multiplying the respective probabilities of the transitions making up these elements: one multiplies probabilities "vertically", in accordance with the causal Markov condition.

In this way, joint probability structures can be defined in models of branching time.

EXTENDING THE ACCOUNT: BRANCHING SPACE-TIMES

In our discussion of why the Cartesian product recipe fails for probabilities based on branching possibilities, we stressed the "spatial" nature of Cartesian product-type combinations, and went for an exploration of temporal (or rather: tempo-modal) combination, resulting in the above recipe. It seems, however, that something important was lost along the way: in our above scheme, there is no place any more for Cartesian products, and thus, for combining probability spaces not just at the same time, but at the same moment (cf. note 6 above).

To be sure, one can always choose to add in such structure by labeling. If, e.g., at m there is 36-fold splitting because two dice are thrown independently at that same moment, we can choose to label the outcomes accordingly, viewing the 36 element partition of H_m effected by the gamble as itself structured in 6×6 cases. But the framework so far provides no means of representing such information internally. This is one motivation for coming up with an extended framework for combining branching possibilities: branching space-times. (In fact, considerations of how to represent simultaneous choices by independent agents were what historically triggered the development of that framework.) The matter is intricate, and page constraints make it impossible for me to give a full account of the framework, or of the resulting probability theory, in this paper. The groundwork for the framework was laid by Belnap (1992); modal correlations were analyzed by Belnap (2002b, 2003) and by Müller et al. (2008). Transitions were analyzed by Belnap (2005). The probability theory was developed by Weiner and Belnap (2006), Müller (2005), and Belnap (2007). For the most natural extension of the story about branching time given above, one should consult Müller (2005).[7]

Very briefly, in branching space-times (BST) the guiding idea is that space should be represented at the same level as time. Considerations of relativity theory lead to the choice of relativistic space-times as the appropriate candidates for histories (possible courses of development). For the Minkowski space-time of special relativity it turns out that defining histories to be *maximal directed sets* is the way to go: a history should be a set such that any two of its members have a common upper bound, and should be maximally so.[8] This makes room for space: incomparable elements can belong to one history, if they share a common future. Such elements are called *space-like related*. In a single Minkowski space-time, all future light cones intersect, so that any two incomparable (i.e., causally isolated) events have a common upper bound. In BST, which integrates, e.g., several Minkowski space-times, two elements without a common upper bound are seen to belong to modal alternatives, i.e., to different histories.

The radical extension of the history definition complicates matters significantly. Obviously, one cannot start with backwards linear structures, but has to allow for a broader class of partial orders. Continuity is a natural requirement, as well as future non-ending. The topology of branching is an especially subtle point; the requirement of maximal points in the intersection of histories, similar to what has been discussed above, turns out to be the right choice (i.e., the choice that allows for sensible definitions of undividedness and splitting, leading to local partitions of histories). BST's *prior choice principle* fixes the topological features.

Once the theory of transitions in BST has been developed, it turns out that much of what has been said above about the combination of "small" (local) alternatives to form joint causal alternatives carries over. Maximal consistent sets of transitions define these alternatives as well. Dealing with space however means that incomparable initials cannot be excluded straight away. If they correspond to space-like related events, like in tossing two dice at the same time, then we want the respective possibilities to combine, in a Cartesian product-like fashion. (It is still better, and more general, to keep talking about consistent *sets* of transitions.) Probabilistic correlations return as a live option: the causal Markov condition only constrains "vertical" combinations of transitions, not "horizontal" ones.[9]

7 In fact I believe that it is useful to read the present paper as an introduction to the mentioned paper, giving a broader motivational background for the technical apparatus that is deveoped there.

8 This will not do for some models of *general* relativity. An appropriate extension of the framework is still a desideratum.

9 The BST framework also leaves room for a broader notion of correlation called "modal funny business": it may turn out that local alternatives do not combine smoothly to form joint possibilities. Perfect correlations in quantum mechanics may be an example of such strange combinatorics. So far, BST probability theory has not been extended to cases of modal funny business. The matter certainly merits further study.

Given this framework, it becomes possible to tackle the issue of *screening off* for probabilistic correlations in an integrated and formally perspicuous manner. For an initial result, and for the technical background, see Müller (2005).

CONCLUSIONS

Probabilities are graded possibilities. Different notions of possibilities afford different probability structures. Standard ways of combining probability spaces by forming Cartesian products correspond to the *spatial* combination of possibilities. The *temporal* combination of possibilities in branching structures such as branching time leads to a different formalism, for which the notion of a maximal consistent set of transitions gives the elements of a joint sample space. By representing the temporal combination of possibilities in an explicit way, the framework contains a good motivation for the causal Markov condition. Branching time-based probabilities are however limited because the framework cannot describe the important case of spatial combination that lies at the bottom of the standard approach. The theory of branching space-times provides for an integrated framework in which both types of combination can be represented.

Acknowledgements
I would like to thank the audience and my co-symposiasts at the ESF Workshop *Physical and Philosophical Perspectives on Probability, Explanation and Time*, Zeist, 20 October 2009, for helpful discussions.

BIBLIOGRAPHY

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

Belnap, N. (1999). Concrete transitions. In Meggle, G., editor, *Actions, Norms, Values: Discussions with Georg Henrik von Wright*, pages 227–236. Berlin: de Gruyter.

Belnap, N. (2002a). Double time references: Speech-act reports as modalities in an indeterminist setting. In Wolter, F., Wansing, H., de Rijke, M., and Zakharyaschev, M., editors, *Advances in Modal Logic, Vol. 3*, pages 37–58. Singapore: World Scientific.

Belnap, N. (2002b). EPR-like "funny business" in the theory of branching space-times. In Placek, T. and Butterfield, J., editors, *Non-locality and Modality*, pages 293–315. Dordrecht: Kluwer.

Belnap, N. (2003). No-common-cause EPR-like funny business in branching space-times. *Philosophical Studies*, 114:199–221.

Belnap, N. (2005). A theory of causation: *Causae causantes* (originating causes) as inus conditions in branching space-times. *British Journal for the Philosophy of Science*, 56:221–253.

Belnap, N. (2007). Propensities and probabilities. *Studies in History and Philosophy of Modern Physics*, 38(3):593–625.

Fine, K. (2005). *Modality and Tense*. Oxford: Oxford University Press.

Gabbay, D. and Guenthner, F., editors (1984). *Handbook of Philosophical Logic, Vol. II*. Dordrecht: D. Reidel.

McCall, S. (1990). Choice trees. In Dunn, J. and Gupta, A., editors, *Truth or Consequences. Essays in Honor of Nuel Belnap*, pages 231–244. Dordrecht: Kluwer.

Müller, T. (2005). Probability theory and causation: A branching space-times analysis. *British Journal for the Philosophy of Science*, 56:487–520.

Müller, T., Belnap, N., and Kishida, K. (2008). Funny business in branching space-times: infinite modal correlations. *Synthese*, 164:141–159.

Popper, K. (1982). *Quantum Theory and the Schism of Physics*. London: Hutchinson.

Prior, A. N. (1957). *Time and Modality*. Oxford: Oxford University Press.

Prior, A. N. (1967). *Past, Present and Future*. Oxford: Oxford University Press.

Thomason, R. H. (1984). Combinations of tense and modality. In Gabbay and Guenthner (1984), pages 135–165.

van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.

von Wright, G. H. (1963). *Norm and Action. A Logical Inquiry*. London: Routledge.

Weiner, M. and Belnap, N. (2006). How causal probabilities might fit into our objectively indeterministic world. *Synthese*, 149:1–36.

Weyl, H. (1940). The ghost of modality. In Farber, M., editor, *Philosophical Essays in Memory of Edmund Husserl*, pages 278–303. Cambridge, MA: Harvard University Press.

Department of Philosophy
University of Utrecht
Janskerkhof 13a
3512 BL Utrecht
The Netherlands
Thomas.Mueller@phil.uu.nl