# The coat problem. Counterfactuals, truth-makers, and temporal specification

*Francesco Gallina*
*Giuseppe Spolaore*
*University of Padua, Italy*
*francescogallina@hotmail.it, giuseppe.spolaore@unipd.it*

**Abstract**

Standard semantic treatments of counterfactuals appeal to a relation of similarity between possible worlds. Similarity, however, is a vague notion. Lewis suggests reducing the vagueness of similarity by adopting a principle known as 'late departure' (LD): the more the past two worlds share, the more they are similar. LD has several virtues. However, as Bennett points out, a standard semantics based on LD suffers from the so-called coat problem. In a nutshell, we are led to assign counterintuitive truth-values to counterfactuals whose antecedent time is left underspecified. In the present paper, we argue that the coat problem may be solved by defining a time-sensitive notion of similarity. To illustrate, we assume a Priorean, tensed language, interpreted on branching-time frames in the usual, 'Ockhamist' way, and we enrich it with a counterfactual connective. Within this framework, we define a time-sensitive relation of similarity, based on Yablo's work on truth-makers and partial truth. In the resulting semantics, which has independent interest, the coat problem does not arise.

**Keywords:** Branching-time, Counterfactuals, Truth-maker, Partial truth

## 1 A problem for the late departure principle

By a *counterfactual* we mean a subjunctive conditional 'If $\mathcal{A}$ were the case, $\mathcal{B}$ would be the case' ($\mathcal{A} \Box\!\!\rightarrow \mathcal{B}$) with false antecedent (see, e.g., [13, 173]). Here is the standard semantic clause for counterfactuals, as developed in [12]:

**Standard counterfactual semantics** (**CS**):
$\mathcal{A} \Box\!\!\rightarrow \mathcal{B}$ is true at a world $w$ iff:

(a) either $w$ has no access to any $\mathcal{A}$-world (the vacuous case); or
(b) some $(\mathcal{A} \wedge \mathcal{B})$-world is more similar to $w$ than any $(\mathcal{A} \wedge \neg\mathcal{B})$-world.

Comparative similarity (with the world of evaluation) is a key ingredient of **CS**. Similarity, however, may be unsatisfying, for it displays a high degree of vagueness. As Lewis [10] points out, the vagueness of similarity cannot be entirely dispelled, for counterfactuals are intrinsically vague. Nevertheless, it can be significantly attenuated.

A first step in this direction is that of characterising comparative similarity as a reflexive and transitive relation, that is, as a preorder. Moreover, as Lewis [11] points out, the vagueness of similarity can be further reduced. Let us focus on *non-backtracking* counterfactuals, that is, roughly, counterfactuals whose antecedents are about eventualities that obtain before (or simultaneously with) the eventualities the consequents are about, as in "If John missed the train, he would still be in London". Non-backtracking counterfactuals, Lewis contends, are evaluated against possible antecedent scenarios that keep, as far as possible, the actual past fixed (for a study providing empirical support to Lewis's insight, see [7]). According to Lewis, in assessing these counterfactuals, we should focus on those worlds that share with the actual world as much of their past as possible, except for the stretch of time strictly needed to make the antecedent true. If Lewis is right, then, it is natural to require that the similarity relation satisfy the following principle (this and the other principles we discuss are assumed to hold for non-backtracking counterfactuals and to impose *ceteris paribus* conditions on similarity):

**Late departure principle** (**LD**):
The more the past two worlds share, the more they are similar.[1]

[1]See [13] for a counterfactual semantics that encodes principle **LD**.

It is easy to refine principle **LD** in formal terms. A very natural mathematical background against which **LD** can be refined is that of the so-called branching-time frames. In this paper, we discuss versions of principle **LD** that are made precise within these frames.

To sum up, principle **LD** helps to reduce the vagueness of similarity, it appears to be involved in counterfactual reasoning, and it has rigorous, formal translations. One may be tempted, then, to assume **LD** with full confidence. Unfortunately, **LD** is problematic.

As Bennett [2] shows, if **LD** is taken to determine the relation of similarity (and assuming something like a branching conception of time), the resulting counterfactual semantics suffers from the so-called *coat problem*. Suppose that John's coat was not stolen from the restaurant where he left it. Moreover, assume that there were just two, equally good, chances for it to be stolen, the former at 1 pm and the latter at 1:30 pm. If the coat were stolen at 1:30 pm, then the thief would sell the coat to a pawnbroker named Fence. Since the latter chance for theft is the one that would obtain later, it obtains in the latest world departing from actuality, among those where the coat is stolen. Thus, if **LD** is adopted, **CS** predicts the truth of:

(1) If John's coat had been stolen, it would now be in Fence's shop.

Intuitively, this result is unwelcome. If we take **CS** for granted, we can conclude that **LD** leads to counterintuitive results. This is the *coat problem*.

In this paper we argue that the coat problem does nothing to undermine the key motivations that justify principle **LD**. Rather, it just shows that **LD** is not general enough, that there are limits in the kind of cases it can be sensibly applied to. In [13, 191] it is suggested that the the coat problem highlights a difficulty of **LD** that eludes formal treatment. In our view, this suggestion is too hasty. As we shall see, once the limits of **LD** are put to the fore, it is possible to define a more general principle, which retains all the virtues of **LD** and vindicates the same basic insights without falling prey to the coat problem.

A disclaimer. It is not part of our aim to offer a variant of **LD** that escapes *any* known difficulty. To be sure, the coat problem is not the only problem that must be solved (or at least put into perspective) if **LD** and similar principles are to be applied to a sufficiently wide range of cases (see, for instance, the so-called late departure problem discussed in [13] and [16]). However, as far as we can see, there is no problem affecting our variant, which does not affect **LD** as well.

## 2 A Priorean revision of the late departure principle

In the celebrated paper *Identifiable individuals* [14], Prior discusses the view that a person could have had different parents than the ones he actually had. According to Prior, in assessing counterfactual hypotheses of this kind, philosophers are often led to neglect a key issue. In his own words:

> It is always a useful exercise (and one insufficiently practised by philosophers), when told that something was possible, i.e. could have happened, to ask '*When* was it possible?' '*When* could it have happened?' (70)

This recommendation has deep connections with other key contributions from Prior, such as his seminal works in the semantics and metaphysics of historical modality.

We suggest that the coat problem depends precisely on the oversight Prior mentions in the above quote, that is, neglecting the role of time in assessing modal claims. Exactly as the modal properties of certain individuals may change across time (as Prior observes), so may change the relative similarity of certain worlds. With this in mind, consider the following principle.

**LD*** The more the past two worlds *do not* share, the more they are *dis*similar.

Under reasonable assumptions, **LD*** is equivalent to **LD**. Nonetheless, **LD*** and **LD** interact in a different way with Prior's recommendation, as the former is much easier to turn into a principle connecting time and (dis)similarity. Here is the principle, that is, the time-relativised version of **LD***:

**Dynamic late departure principle** (**DLD**): The more the past two worlds do not share *at a given time*, the more they are dissimilar *at that time* (or, if you prefer, *up to* that time).

To make **DLD** precise, we need to define a formally respectable, time-sensitive relation of past-(not-)sharing. To this aim, we adopt a branching-time conception, in which worlds are identified with *histories*, that is, spatially and temporally complete courses of physical events. A branching time structure or *tree* is, roughly, a bunch of histories that share an initial, 'past' part and divide afterwards, yielding different branches.

If $h$ is the actual history (i.e., the history of evaluation) and $h'$ is a history divided from $h$ at time $t$, we call a *non-actual stretch* of $h'$ the part of $h'$ that goes from $t$ to the time of branching between $h$ and $h'$ (which we indicate as $h \curlyvee h'$, see figure 1).[2] By **DLD**, the shorter the non-actual stretch of $h'$ at $t$ is, the more similar $h'$ is to actuality at $t$.

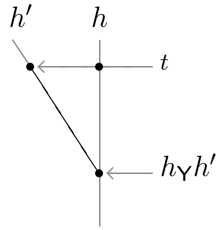


Figure 1: The non-actual stretch of $h'$ at $t$ (highlighted in black).

Before returning to the coat problem, let us spend a few words of comment on principle **DLD**. Admittedly, **DLD** is conceptually more costly than **LD**, for it requires that different histories be temporally comparable or 'synchronised' (see below, p. 4). However, there are independent reasons to pay the price (see, e.g., [4, 265–266]). Once synchronised trees are adopted, it is hard to see why someone who leans towards **LD** may be willing to reject its time-relativised version **DLD**. After all, for any time $t$, if we use **LD** and **DLD** to measure the comparative similarity of two histories that are divided at $t$, then the two principles yield precisely the same verdict. In addition to that, **DLD** allows us to compare distinct histories relative to different times—a kind of comparison that makes perfect sense within the branching-time conception. As we shall argue in the next section, this feature of **DLD** is key to solving the coat problem.

[2]We assume a discrete temporal ordering. For a formal definition of trees, see § 4.

# 3 Outline of a solution to the coat problem

Let us familiarly speak of an *antecedent truth-maker* to indicate an event that would make the antecedent true, and let us speak of an *antecedent time* to indicate the time at which a truth-maker of the antecedent obtains (for more precise characterisations of these notions, see below, §§ 5 and 6).[3]

Consider again the scenario described by Bennett, as represented in figure 2. John's coat is not stolen in the actual history $h$ but might have been in two specific occasions, $O_1$ (which would obtain only at 1 pm on $h_1$) and $O_2$ (only at 1:30 pm on $h_2$). Both $O_1$ and $O_2$ are antecedent truth-makers. Let us assume that $h_1$ is the $O_1$-history that has the shortest non-actual stretch at the relevant antecedent time 1 pm. Analogously, $h_2$ is the $O_2$-history that has the shortest non-actual stretch at antecedent time 1:30 pm. In figure 2, the non-actual stretch of $h_1$ at 1 pm is the distance from 12:45 pm (time $h \curlyvee h_1$) to 1 pm (the antecedent time corresponding to $O_1$), and the non-actual stretch of $h_2$ at 1:30 pm is the distance from 1:15 pm (time $h \curlyvee h_2$) to 1:30 pm (the antecedent time corresponding to $O_1$). Both stretches are 15 minutes long. The coat now would be in Fence's shop (scenario $S$) if, and only if, $O_2$ had obtained.

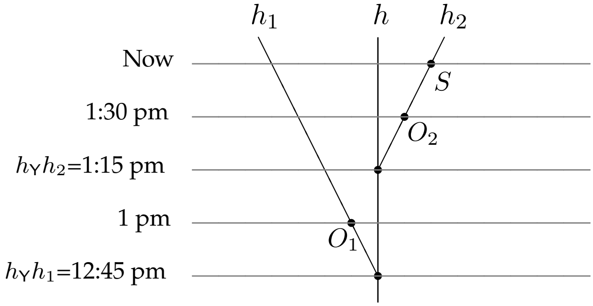


Figure 2: A partial representation of Bennett's scenario.

This is Bennett's counterfactual:

(1) If John's coat had been stolen, it would now be in Fence's shop.

It has been noted that the antecedent of (1) is underspecified, in that

[3] In [6], Kit Fine adopts a counterfactual semantics based on the so-called truth-maker semantics. Our proposal differs from Fine's account, for it mixes possible worlds semantics with truth-maker semantics. For a similar approach, see [5].

there are two distinct antecedent truth-makers $O_1$ and $O_2$ (see, e.g., [8]). More importantly for us, the antecedent is *temporally* underspecified, as there are two different antecedent times, 1 pm and 1:30 pm. We can make this temporal underspecification explicit by replacing (1) with:

(2) If John's coat had been stolen either at 1 pm or at 1:30 pm, it would now be in Fence's shop.[4]

When the antecedent time is uniquely specified, **LD** measures precisely what it is supposed to measure: how long a counterfactual history has to be divided from actuality to make the antecedent true. When the antecedent time is underspecified, however, **LD** may go astray. This is precisely what happens in the coat problem. By **LD**, $h_2$ is deemed closer to actuality than $h_1$ just because the time of $O_2$ is later than that of $O_1$. Thus, by **CS**, it turns out that (2) is true, against common intuitions. Now, let us see why **DLD** fares better than **LD** in this respect. As just seen, (2) is temporally underspecified in that it has two antecedent times. Accordingly, the task of evaluating (2) boils down to the following tasks:

(i) consider 1 pm, that is, the antecedent time of the truth-maker $O_1$, and assign to the $O_1$-histories a measure of similarity with actuality at 1 pm;
(ii) repeat the same operation with the antecedent time of the truth-maker $O_2$;
(iii) apply **CS** to all antecedent histories, using the measures of similarity assigned at points (i), (ii).

It should be clear that, differently from **LD**, the variant **DLD** we are proposing is perfectly adequate for tasks (i) and (ii). In a branching-time structure, if $h$ is the actual history, for each antecedent history $h'$ with antecedent time $t$, there exists a unique non-actual stretch of $h'$ at $t$. By **DLD**, the length $n$ of this stretch counts as a measure of dissimilarity from actuality. Thus, we are in a position to apply **CS** in the usual way: a counterfactual $\mathcal{A} \mathrel{\Box\!\!\rightarrow} \mathcal{B}$ is presently true iff $\mathcal{B}$ is true in all $\mathcal{A}$-worlds whose measure $n$ is smaller than that of any $\mathcal{A}$-world in which $\mathcal{B}$ is false.

[4] It has been recently argued that counterfactuals with disjunctive antecedent pose serious problems to any semantics based on **CS** (see, e.g., [3]). Albeit we think that these problems are worth considering, discussing them is beyond the scope of this paper.

**DLD** enables to get the coat story right: since the non-actual stretch of $h_1$ at the (antecedent) time of $O_1$ is as long as that of history $h_2$ at the (antecedent) time of $O_2$, both histories count as equally similar to actuality. Since $S$ is true at $h_2$ only, however, statement (2) (or, equivalently, (1)) is false at the actual history. Notice that when the antecedent is not temporally underspecified—when a unique antecedent time exists—**DLD** and **LD** agree on the similarity ordering. In the next section we outline a formal version of this preliminary analysis.

# 4 Counterfactuals and branching-time structures

As said above, our solution requires that all histories in the tree are temporally comparable. Accordingly, we choose a suitable brand of branching-time structures, viz., *sinchronised* trees (see [4, 269–273], see also [1, 195–196]).

A synchronised tree $\mathcal{T}$ is defined as a tuple $(M, \preceq, d)$, where $M$ is a nonempty set of entities called moments, and $\preceq$ is a partial order on $M$, which corresponds to the (improper) precedence relation on $M$ ($\prec$, $\succeq$, and $\succ$ are defined in the obvious way). To keep the formal complexity to a minimum, we assume that $\preceq$ is a *discrete* ordering. Nothing philosophically important hinges on this assumption. A *history* is defined as a maximal set of moments in $\preceq$. Moreover, letting $m, m'$ and $m''$ vary over moments:

(a) if $m' \npreceq m''$ and $m'' \nprec m'$, some $m$ is such that $m \prec m'$ and $m \prec m''$;
(b) if $m' \prec m$ and $m'' \prec m$, either $m' \preceq m''$, or $m'' \prec m'$;
(c) histories are all isomorphic;
(d) a *time* $t$ is a set of moments that intersects each history at precisely one moment. We also write $t_h$ to indicate the moment in $t \cap h$.
(e) times preserve the order of the corresponding moments (that is, $t_h \prec t'_h$ entails $t_{h'} \npreceq t'_{h'}$). We shall say that $t$ is earlier than $t'$ iff $t_h \prec t'_h$ for some $h$.

$d$ is a metric function that assigns to any pair $(t, t')$ of times a non-negative number $n$ expressing the temporal distance between $t$ and $t'$ (see [9] for details). We require that $t' = t''$ iff $t_h \prec t'_h, t''_{h'}$ entails $d(t, t') = d(t, t'')$.

The language $\mathcal{L}_T$ is a standard tensed propositional language endowed with a set $Atom$ of countable atoms $p, q, p_1, \ldots$, and with two sentential operators P ('Sometimes in the past') and F ('Sometimes in the future'). We indicate as H and G the duals of P, F, respectively. As usual, atoms are thought of as simple, present tensed sentences that contain no 'trace of futurity'.

A standard semantics for $\mathcal{L}_T$ is Prior's [15] *Ockhamist semantics*, which evaluates sentences of $\mathcal{L}_T$ at moment-history pairs. An Ockhamist model is a tuple $\mathfrak{M}_{\mathsf{O}} = (\mathcal{T}, \mathcal{I})$, where $\mathcal{T}$ is a tree and $\mathcal{I}$ an interpretation function from $Atom \times M$ onto the set $\{0, 1\}$ of truth-values. Elements of $\mathcal{I}$ are called *assignments*. Ockhamist truth is defined in the usual, recursive way (clauses for booleans and reference to models are omitted; we abbreviate $\{t_h, h\}$ as $t/h$):

$$t/h \vDash p \quad \text{iff} \quad \mathcal{I}(p, t_h) = 1;$$
$$t/h \vDash \mathsf{P}\mathcal{A} \quad \text{iff} \quad \exists t'(t'_h \prec t_h \;\&\; t'/h \vDash \mathcal{A});$$
$$t/h \vDash \mathsf{F}\mathcal{A} \quad \text{iff} \quad \exists t'(t_h \prec t'_h \;\&\; t'/h \vDash \mathcal{A}).$$

Now, let us enrich $\mathcal{L}_T$ with connective $\Box\!\!\rightarrow$, yielding language $\mathcal{L}_{TC}$. The syntax of $\mathcal{L}_C$ is defined in the obvious way. We ignore complications that depend on embedding counterfactuals one into another.

To interpret $\mathcal{L}_{TC}$ in accordance with the Ockhamist semantics, counterfactuals are themselves to be evaluated relative to moment-history pairs. As argued in [16, 182–185], the moments that are relevant for assessing $\mathcal{A} \Box\!\!\rightarrow \mathcal{B}$ at a point $t/h$ are those located at $t$. Clause **CS** must be modified accordingly:

**CS*** $t/h \vDash \mathcal{A} \Box\!\!\rightarrow \mathcal{B}$ iff

(a) either no $t/h'$ satisfies $\mathcal{A}$, for any history $h'$, or
(b) some $(\mathcal{A} \wedge \mathcal{B})$-point $t/h'$ is strictly more similar to $t/h$ than any $(\mathcal{A} \wedge \neg\mathcal{B})$-point $t/h''$.

For our purposes, claims about the similarity of two points $t/h$ and $t/h'$ boils down to claims concerning the similarity of histories $h$ and $h'$ up to time $t$.

## 5 Similarity, truth-makers, and antecedent times

Clause **CS*** requires that a notion of comparative similarity between moment-history pairs be provided. Both **LD** and **DLD** can be used to perform this task.

Let us start with **LD**. Recall that histories branch off from one another only towards the future. Thus, to say that the more the past two worlds share the more they are similar amounts to saying that the more the moments two histories share, the more they are similar. Let us call LD-similarity the notion of comparative similarity that we can distil along these lines. Formally, $h'$ is at least as LD-similar to $h$ as $h''$ iff $h' \cap h \supseteq h'' \cap h$ (see [13] and [18]). Based on LD-similarity, it is straightforward to define a comparative notion of LD*-similarity between antecedent moment-history pairs: $\mathcal{A}$-point $t/h'$ is as LD*-similar to $t/h$ as $\mathcal{A}$-point $t/h''$ if and only if $h'$ is as LD-similar to $h$ as $h''$.

To get a semantics for counterfactuals, it is sufficient to identify the similarity relation in **CS*** with LD*-similarity. As expected, the resulting semantics falls prey to the coat problem. If $h$, $h_1$ and $h_2$ are as in figure 2, then the following formal version of (2) turns out to be true at Now$/h$:

(3) $(\mathsf{P}o_1 \vee \mathsf{P}o_2) \mathrel{\Box\!\!\rightarrow} s$

Let us now consider **DLD**. We start by offering a formal counterpart of the above, intuitive notion of a truth-maker, based on Yablo's work [17, Chap. 4]. We identify a truth-maker of $\mathcal{A}$ at $t/h$ ($\mathsf{tmk}(\mathcal{A}, t/h)$ in symbols) with a *minimal model* of $\mathcal{A}$ at $t/h$. In turn, a minimal model is a set of assignments that is, intuitively, as small as is strictly necessary to make $\mathcal{A}$ true at $t/h$. More formally, given a model $\mathfrak{M}_O = (\mathcal{T}, \mathcal{I})$, a truth-maker $\mathsf{tmk}(\mathcal{A}, t/h)$ is a set of assignments such that:

(i) $\mathsf{tmk}(\mathcal{A}, t/h) \subseteq \mathcal{I}$;
(ii) if an Ockhamist model $\mathfrak{M}'_O = (\mathcal{T}, \mathcal{I}')$ is such that $\mathsf{tmk}(\mathcal{A}, t/h) \subseteq \mathcal{I}'$, then $\mathfrak{M}'_O, t/h \vDash \mathcal{A}$;
(iii) if $f$ is a set of assignments such that $\mathsf{tmk}(\mathcal{A}, t/h) \supset f$, then some Ockhamist model $\mathfrak{M}'_O = (\mathcal{T}, \mathcal{I}')$ is such that $\mathcal{I}' \supset f$ and $\mathfrak{M}'_O, t/h \nvDash \mathcal{A}$.

If a truth-maker assigns a value to a pair $(p, t_h)$, we shall say that it *covers* time $t$ (on $h$). Moreover, a set $f$ is a *possible truth-maker* of $\mathcal{A}$ at $t/h$ in $\mathfrak{M}_O = (\mathcal{T}, \mathcal{I})$ if, for some Ockhamist model $\mathfrak{M}'_O = (\mathcal{T}, \mathcal{I}')$, $f$ is a truth-maker of $\mathcal{A}$ at $t/h$ in $\mathfrak{M}'_O$. We indicate a possible truth-maker of $\mathcal{A}$ at $t/h$ as $\Diamond\mathsf{tmk}(\mathcal{A}, t/h)$.

Some comments are in order. First, if a sentence $\mathcal{A}$ is true at a point, there exists at least one truth-maker of $\mathcal{A}$ at that point. However, a sentence can have more than one truth-maker at a point. For instance, there are exactly two possible truth-makers of $(p \vee \neg q)$ at $t/h$, namely, $\{(p, t_h) \mapsto 1\}$ and $\{(q, t_h) \mapsto 0\}$. If $(p \vee \neg q)$ is true at $t/h$, at least one of these possible truth-makers is also actual. But nothing forbids both from being actual.

Second, $\mathsf{tmk}(\mathcal{A}, t/h)$ need not cover time $t$. Consider, for instance, $\mathsf{P}p$ at point $t/h$, and suppose that $\mathcal{I}(p, t'_h) = 1$, with $t'_h \prec t_h$. Then, $\{(p, t'_h) \to 1\}$ is a truth-maker of $\mathsf{P}p$ at $t/h$ but does not cover $t$ at all.

Third, a truth-maker may cover more than one time. Suppose, for instance, that $(p \wedge \mathsf{P}q)$ is true at $t/h$. If so, its truth-makers at $t/h$ must cover exactly two times, that is, be of form $\{(p, t_h) \mapsto 1, (q, t'_h) \mapsto 1\}$ (with $t'_h \prec t_h$).

To see how truth-makers enable us to deal with temporal underspecification, let us start by assessing counterfactual (3) against the tree in figure 2. There are two truth-makers of the antecedent of (3) at time Now:

(i) $\{(o_1, 1\,\text{pm}_{h1}) \mapsto 1\}$, corresponding to occasion $O_1$; and
(ii) $\{(o_2, 1{:}30\,\text{pm}_{h2}) \mapsto 1\}$, corresponding to $O_2$.

The measure of similarity with actuality of $\text{Now}/h_1$ and $\text{Now}/h_2$ is given by the distance between the time covered by the antecedent truth-maker on each history and the time at which that history divides from actuality. Since the distance is the same (15 minutes), point $\text{Now}/h_1$ is as similar to actuality as point $\text{Now}/h_2$.

Let us generalise this account. We define $\mathsf{antec}'$ as a function from triples $(\mathcal{A}, t/h', h)$ (where $t/h'$ is assumed to be an antecedent $\mathcal{A}$-point and $h$ is the actual history) to times such that:

- if the earliest time $t^*$ covered by some truth-maker of $\mathcal{A}$ at $t/h'$ is later than the time of branching $h \curlyvee h'$, then $\mathsf{antec}'(\mathcal{A}, t/h', h) = t^*$;
- $\mathsf{antec}'(\mathcal{A}, t/h', h)$ is undefined otherwise.

Intuitively, $\mathsf{antec}'$ is (our first shot at) a formal counterpart of the notion of an antecedent time. Let us note that $\mathsf{antec}'(\mathcal{A}, t/h', h)$, if defined, is required to be later than the time of branching between $h$ and $h'$. Since we characterised counterfactuals as conditionals with false antecedents, we may assume that the antecedent time on a history $h'$ should be later than the time at which $h'$ divides from the actual history $h$. Now, let us define a similarity mapping $\mathsf{siml}'$, as follows, where $t/h'$ is an $\mathcal{A}$-point:

$$\mathsf{siml}'(\mathcal{A}, t/h', h) = d(h \curlyvee h', \mathsf{antec}'(\mathcal{A}, t/h', h)).$$

The values $\mathsf{antec}'(\mathcal{A}, t/h', h)$ and $\mathsf{siml}'(\mathcal{A}, t/h', h)$ are defined only if the earliest time covered by some truth-maker of $\mathcal{A}$ at $t/h'$ is after the time of branching between $h$ and $h'$. We are now in a position to specify a relation of comparative similarity between antecedent points:

**DLD′-similarity** $\mathcal{A}$-point $t/h'$ is at least as similar to actuality $t/h$ as $\mathcal{A}$-point $t/h''$ iff $\mathsf{sml}'(\mathcal{A}, t/h', h) \leq \mathsf{sml}'(\mathcal{A}, t/h'', h)$.

If the similarity relation in **CS*** is DLD-similarity, the resulting semantics does not fall prey to the coat problem.

Thus, it is tempting to suppose that DLD′-similarity is precisely the notion we were looking for. This is false, however, as we are going to argue in the next section.

## 6 Antecedent times and partial truth

To see why DLD′-similarity is unsatisfying, an example may be useful. Let then $\mathsf{H}p$ be the formal version of "Always in the past, the Moon was free from human footprints", and consider counterfactual $\mathsf{H}p \mathrel{\Box\!\!\rightarrow} \mathcal{B}$. Let $t_h$ be the present, actual point, and let $t'_h$ be the moment on July 20, 1969 of Armstrong's celebrated 'one small step' on the Moon. Clearly, on $h$, sentence $p$ is always true up to $t'$, it becomes false at $t'$, and is always false from $t'$ onwards. Now, consider an antecedent point $t/h'$. Since $p$ is always true in the past of $t_{h'}$, there exists no $\mathsf{antec}'(\mathsf{H}p, t/h', h)$ that is later than $h'_{\curlyvee}h$. As a consequence, $\mathsf{sml}'$ is undefined for argument $(\mathsf{H}p, t/h', h)$, and we cannot assign to $t/h'$ any measure of similarity to actuality. The problem is that our 'official' notion of antecedent time, $\mathsf{antec}'$, is in wait of substantial refinement.

To best appreciate the reason why $\mathsf{antec}'$ goes astray, it is useful to introduce a novel notion, that of *partial truth* (see also [17, Ch. 5]). Let us say that a sentence $\mathcal{A}$ is partially true at a point $t/h$ in model $\mathfrak{M}_O$ if, for some assignment $f \in \mathcal{I}$, $f$ is an element of a possible truth-maker of $\mathcal{A}$ at $t/h$ in $\mathfrak{M}_O$. More formally:

**Partial truth** $\mathcal{A}$ is *partially true* at a point $t/h$ in model $\mathfrak{M}_O$ iff, for some $\Diamond\mathsf{tmk}(\mathcal{A}, t/h)$ in $\mathfrak{M}_O$, we have that $\Diamond\mathsf{tmk}(\mathcal{A}, t/h) \cap \mathcal{I} \neq \emptyset$. Intersection $\Diamond\mathsf{tmk}(\mathcal{A}, t/h) \cap \mathcal{I}$ is called a *partial truth-maker* of $\mathcal{A}$ at $t/h$.

A *non-empty* partial truth-maker of $\mathcal{A}$ is also called a *true part* (of content) of $\mathcal{A}$.

The antecedent of counterfactual $\mathsf{H}p \;\Box\!\!\rightarrow \mathcal{B}$, unlike the other examples we have considered thus far, is partially true at $t/h$. Its true part is the set of assignments $\{(t'_h, p) \mapsto 1 : t'_h \prec t_h\}$. In general, DLD′-similarity does not work with counterfactuals whose antecedent is *not completely* false. The reason is that, intuitively, DLD′-similarity is only sensitive to the antecedent truth-makers that cover times on antecedent histories. In counterfactual reasoning, however, we are not interested only in what happens on antecedent histories. Rather, we are crucially interested in the *differences* between antecedent and actual histories—more specifically, in the differences that are *determined by the truth of the antecedent*. Antecedent times, in turn, may be thought of as times at which such differences (begin to) surface. Thus, when singling out antecedent times, we must forget about the parts of the antecedent (as it were) that are true in both the actual and the antecedent histories, for these parts correspond to no genuine difference.

Let us try to make these casual remarks more precise. To this aim, the following definitions are useful:

**Completion** If $g$ is a possible truth-maker of $\mathcal{A}$ at $t/h$ and $f \subseteq g$ is a partial truth-maker of $\mathcal{A}$ at $t/h$, we call *completion* of $\mathcal{A}$ at $t/h$ the difference $g - f$.

**Copy** If $f$ is a set of assignments on history $h'$, the *copy* of $f$ on $h$ is the set obtained by replacing, for each time $t$ covered by some $a \in f$, the moment $t_{h'}$ in the argument of $a$ with $t_h$.

**Difference-maker** A $h$-difference-maker of $\mathcal{A}$ at $t/h'$ is a subset $f$ of a truth-maker of $\mathcal{A}$ at $t/h'$ such that the copy of $f$ on $h$ is a completion of $\mathcal{A}$ at $t/h$.

To illustrate, let us assume that $h$ is the actual history and $t/h'$ is an antecedent $\mathcal{A}$-point. A completion of $\mathcal{A}$ at $t/h$ is, intuitively, the minimal set of assignments that need to be added to a partial truth-maker of $\mathcal{A}$ at $t/h$ to turn it into a possible truth-maker of $\mathcal{A}$ at $t/h$. The $h$-difference-maker of $\mathcal{A}$ at $t/h'$ is a minimal set of assignments on $h'$ that, intuitively, if copied on the actual history $h$, would make $\mathcal{A}$ true at the actual point

$t/h$. Note that if $\mathcal{A}$ is completely false at $t/h$, then all truth-makers of $\mathcal{A}$ at $t/h'$ are $h$-difference-makers of $\mathcal{A}$ at $t/h'$.

Let us turn to the resulting, refined notion of antecedent time. If $h$ is the actual history and $t/h'$ is an $\mathcal{A}$-point, we let $\mathsf{antec}(\mathcal{A}, t/h', h)$ be the earliest time covered by some $h$-difference-maker of $\mathcal{A}$ at $t/h'$. In the Moon example discussed above, $\mathsf{antec}(\mathsf{H}p, t/h', h)$ is, intuitively, the time of the earliest assignment that we must modify in $\mathcal{I}$ to make $\mathsf{H}p$ actually true at $t$. Of course, $\mathsf{antec}(\mathsf{H}p, t/h', h)$ is the time of Armstrong's 'one small step'.

By appeal to this difference-sensitive notion of antecedent time, we may define the following, refined counterparts of $\mathsf{siml}'$ and of **DLD**′-similarity, respectively:

$$\mathsf{siml}(\mathcal{A}, t/h', h) = d(h_\curlyvee h', \mathsf{antec}(\mathcal{A}, t/h', h));$$

**DLD-similarity** $\mathcal{A}$-point $t/h'$ is at least as similar to actuality $t/h$ as $\mathcal{A}$-point $t/h''$ iff $\mathsf{sml}(\mathcal{A}, t/h', h) \leq \mathsf{sml}(\mathcal{A}, t/h'', h)$.

Since DLD-similarity of an $\mathcal{A}$-point $t/h'$ to actuality is inversely proportional to the length of the non-actual stretch of $h'$ at the antecedent time, DLD-similarity vindicates the basic insights behind **DLD**. Besides, DLD-similarity helps to get the role of partial truth in counterfactual reasoning right.

Partial truth antecedents, moreover, highlight an interesting difference between the notions of antecedent time encoded by $\mathsf{antec}$ and by $\mathsf{antec}'$. When $\mathcal{A}$ is completely false at $t/h$ and the antecedent time on $h'$ is later than $h_\curlyvee h'$, then $\mathsf{siml}(\mathcal{A}, t/h', h) = \mathsf{siml}'(\mathcal{A}, t/h', h)$. If $\mathcal{A}$ is completely false at $t/h$, indeed, every copy on $h$ of a truth-maker of $\mathcal{A}$ at $t/h'$ is a completion of the unique partial truth-maker of $\mathcal{A}$ at $t/h$ (viz., a completion of $\emptyset$). Accordingly, every truth-maker $\mathsf{tmk}(\mathcal{A}, t/h')$ is a $h$-difference-maker of $\mathcal{A}$ at $t/h'$, and so the antecedent times we get by $\mathsf{antec}'$ and $\mathsf{antec}$ coincide.

In the face of it, one may be tempted to hold that, if the antecedent time we get by $\mathsf{antec}'$ on $h'$ is later than $h_\curlyvee h'$, then it must coincide with the antecedent time we get by $\mathsf{antec}$ on $h'$. This is not so, however. Consider the tree in figure 3, where $(\mathsf{P}p \wedge \mathsf{P}q)$ is false at $t/h$ (for $q$ never holds on $h$), but not completely so (for $p$ holds at point $t''/h$, which lies in the past of $t/h$).

The (unique) truth-maker of $(\mathsf{P}p \wedge \mathsf{P}q)$ at $t/h'$ is the set of assignments $\{(p, t''_{h'}) \mapsto 1, (q, t'_{h'}) \mapsto 1\}$, and the earliest time it covers is

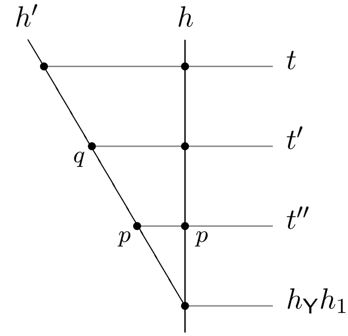


Figure 3: A partial representation of a tree, where atoms $p$ and $q$ never hold, except at the specified points.

$t''$. Accordingly, $\mathsf{antec}'((\mathsf{P}p \wedge \mathsf{P}q), t/h', h) = t''$. However, the only $h$-difference-maker of $(\mathsf{P}p \wedge \mathsf{P}q)$ at $t/h'$ is $\{(q, t'_{h'}) \mapsto 1\}$, for copying it on $h$ is sufficient for making $(\mathsf{P}p \wedge \mathsf{P}q)$ true at the evaluation point $t/h$. Therefore, $\mathsf{antec}((\mathsf{P}p \wedge \mathsf{P}q), t/h', h) = t'$, and thus $\mathsf{antec}((\mathsf{P}p \wedge \mathsf{P}q), t/h', h) \neq \mathsf{antec}'((\mathsf{P}p \wedge \mathsf{P}q), t/h', h)$.

This result highlights that antec is sensible to differences between actuality and antecedent histories that antec′ cannot detect. Intuitively, for each actual point $t/h$ and antecedent point $t/h'$, antec picks up the time at which a difference between $h$ and $h'$ surfaces, which explains why the antecedent is true at $t/h'$ as opposed to $t/h$.

# 7 Conclusions

In this paper, we have introduced and discussed a 'dynamic' version **DLD** of the late departure principle. **DLD** allows us to distil a notion of comparative similarity between histories, which we called DLD-similarity. As we have shown, a counterfactual semantics based on DLD-similarity retains all the virtues of that based on LD-similarity, but it also accounts for counterfactuals whose antecedent time is underspecified.

We have argued that this difference is key to solving Bennet's coat problem.

## Bibliography


[1] Belnap, N., Perloff, M. & Xu, M. (2001). *Facing the future: agents and choices in our indeterminist world*. Oxford: Oxford University Press.

[2] Bennett, J. F. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.

[3] Champollion, L., Ciardelli, I. & Zhang, L. (2016). Breaking de Morgan's law in counterfactual antecedents. In M. Moroney, C.-R. Little, J. Collard, and D. Burgdorf, editors, *Proceedings of the 26$^{th}$ Semantics and Linguistic Theory Conference*, pages 304–324.

[4] Di Maio, M. C. & Zanardo, A. (1994). Synchronized histories in Prior-Thomason representation of branching time. In D. Gabbay and H. Ohlbach (Eds.), *Proceedings of the First International Conference on Temporal Logic* (265–282). Dordrecht: Springer-Verlag.

[5] Embry B. (2014). Counterfactuals without possible worlds? A Difficulty for Fine's exact semantics for counterfactuals. *The Journal of Philosophy*. 111(5), 276–287.

[6] Fine, K. (2012). Counterfactuals without possible worlds. *The Journal of Philosophy*, 190(3), 221–246.

[7] Gerstenberg, T., Bechlivanidis, C. & Lagnado, D. A. (2013). Back on track: Backtracking in counterfactual reasoning. In M. Knauff, M. Pauen, N. Sebanz & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, (2386–2391). Austin, TX: Cognitive Science Society.

[8] Khoo, J. (2016). Backtracking counterfactuals revisited. *Mind*, 126 (503), 841–910.

[9] Koymans, R. (1990). Specifying real-time properties with metric temporal logic. *Real-Time Systems*, 2(4), 255–299.

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

[11] Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, 3(4), 455–476.

[12] Lewis, D. (1981). Ordering semantics and premise semantics for counterfactuals. *Journal of philosophical logic*, 10(2), 217–234.

[13] Placek, T. & Müller, T. (2007). Counterfactuals and historical possibility. *Synthese*, 154(2), 173–197.

[14] Prior, A. (1960). Identifiable individuals. *Review of Metaphysics*, 13(4), 684–696. Reprinted in his *Papers on Time and Tense* (66–77). Oxford: Oxford University Press.

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

[16] Thomason, R. H., & Gupta, A. (1980). A theory of conditionals in the context of branching time. *Philosophical Review*, 89(1), 65–90.

[17] Yablo, S. (2014). *Aboutness*. Princeton: Princeton University Press.

[18] Zanardo, A. (2006). Moment/history duality in Prior's logics of branching-time. *Synthese*, 150(3), 483–507.