Journal of Logic, Language and Information (2023) 32:677–706
https://doi.org/10.1007/s10849-023-09400-6

From *Counterfactual* Conditionals to *Temporal* Conditionals

Yuichiro Hosokawa[1]

Accepted: 16 April 2023 / Published online: 3 June 2023
© The Author(s) 2023

Abstract
Although it receives less attention, (Lewis in Noûs 13:455–476, 1979. https://doi.org/10.2307/2215339) admitted that the branching-time(-like) model fits a wide range of counterfactuals, including (Nix) *'If Nixon had pressed the button, there would have been a nuclear war'*, which was raised by (Fine in Mind 84:451–458, 1975). However, Lewis then claimed that similarity analysis is more general than temporality analysis. In this paper, we do not scrutinise his claim. Instead, we re-analyse (Nix) not only model-theoretically but also proof-theoretically from the 'meaning-as-use' and 'inferentialist' points of view. Then, we re-formalise (Nix) in a natural extension of hybrid tense logic, which we refer to as hybrid tense logic for temporal conditionals (HTL_{TC}). Consequently, we find that not only among counterfactuals, but also among indicatives, there is a wide range of conditionals whose formalisation in HTL_{TC} is appropriate. We refer to these conditionals as *temporal conditionals*. This suggests *a new logical generality* that temporality analysis has but similarity analysis does not, from which emerges *a new logical perspective* on conditionals in general: *temporal ones and others*.

Keywords Counterfactuals · Indicatives · Conditionals · Hybrid tense logic · Branching-time model

1 Introduction

1.1 Fine Problem Revisited

Lewis (1973) symbolised counterfactual conditionals in the form *'If it were the case that* φ*, then it would be the case that* ψ*'* as $\varphi \mathrel{\Box\!\!\rightarrow} \psi$. He then presented the truth condition of $\varphi \mathrel{\Box\!\!\rightarrow} \psi$, which roughly states:

✉ Yuichiro Hosokawa
hosokawayuichiro@mail.gpwu.ac.jp

[1] Gunma Prefectural Women's University, Tamamura, Gunma, Japan

(SA)[1] In all possible worlds that are most similar to the actual world except that φ holds, ψ also holds.

Thus, the meaning of $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ was explained in terms of 'similarity' among possible worlds. Since then, Lewis's analysis of counterfactuals has been influential and standard, particularly in analytic philosophy and linguistics.

However, Fine (1975) pointed out that (SA) had a serious problem, which we refer to as the 'Fine Problem'. The problem is shown perspicuously by the following sentence:

(Nix) If Nixon had pressed the button, there would have been a nuclear war.

Although this is a counterfactual conditional that is surely true, it may be thought to be false according to (SA). If we take the meaning of 'being similar' literally, the situations in possible worlds in which a nuclear war has occurred because of Nixon's pressing the button should be entirely different from the situation in our actual world in which no nuclear war has occurred yet. Instead, the situations in the possible worlds in which Nixon did press the button but miraculously no nuclear war has occurred may well remain largely unchanged compared to the actual situation. If so, we would have to say that the following sentence is true:

($\overline{\text{Nix}}$) Even if Nixon had pressed the button, there would not have been a nuclear war.

To avoid this absurd consequence, Lewis made the following reply (1979), which was so natural that anyone who wanted to maintain (SA) would adopt it: in the case of (Nix), such a miracle that Nixon pressed the button but no nuclear war occurred quite contradicts important laws, such as physical, social, political, and military laws, of the actual world. The deviations from such laws are much heavier than the disagreements with particular facts (the relatively peaceful course of events) in the actual world. Accordingly, there is a smaller difference from reality in the succedent of (Nix), which is only against particular facts, than in the succedent of ($\overline{\text{Nix}}$), which is against actual important laws. Therefore, (Nix) is true, and ($\overline{\text{Nix}}$) is false, as expected.

This is a fairly famous story in the literature. However, there is another story that seems to have received less attention. Before replying to the Fine Problem in this manner, Lewis temporarily took into consideration branching-time-*like* model of counterfactuals,,[2] which he referred to as 'Analysis 1', and admitted that Analysis 1 seems to fit a wide range of counterfactuals ((Lewis, 1979), p. 463). We quote his formulation of it here[3]:

1 (SA) is the abbreviation for 'Similarity Analysis'.

2 In Lewis's terms, strictly speaking, the structure given below should not be said to be a branching-time model, but instead a branching-time-*like* or *diverging*-time model. In the former, it is assumed that two worlds (histories) share one and the same initial spatiotemporal segment in common, and therefore there are two overlapping worlds (histories). In the latter, by contrast, it is assumed that there is no overlap: two worlds (histories) have two duplicate initial segments, not one that they share in common. Lewis officially rejects the former in favour of the latter. See Lewis (1986, p. 206). However, the distinction between branching-time and branching-time-*like* (or *diverging*-time) does not affect the discussion below.

3 Cited from Lewis (1979, p. 462).

ANALYSIS 1. Consider a counterfactual "If it were that A, then it would be that C" where A is entirely about affairs in a stretch of time t_A. Consider all those possible worlds w such that:

(1) A is true at w;
(2) w is exactly like our actual world at all times before a transition period beginning shortly before t_A;
(3) w conforms to the actual laws of nature at all times after t_A; and
(4) during t_A and the preceding transition period, w differs no more from our actual world than it must to permit A to hold.

The counterfactual is true if and only if C holds at every such world w.

By contrast, his similarity analysis, the (SA) above, was referred to as 'Analysis 2', and restated as follows[4]:

ANALYSIS 2. A counterfactual "If it were that A, then it would be that C" is (non-vacuously) true if and only if some (accessible) world where both A and C are true is more similar to our actual world, overall, than is any world where A is true but C is false.

He then claimed that Analysis 2 is more general than Analysis 1 because there are many counterfactual suppositions that seem to have no connection with particular times, such as *'If kangaroos had no tails...'* and *'If gravity went by the inverse cube of distance...'*.

1.2 Our Course of Logical Analysis of Counterfactuals

In this paper, we do not scrutinise Lewis's above claim. Instead, we begin with a re-analysis of (Nix), following a paradigm of *logical* analysis of natural language in analytic philosophy: *Tarski's theory of truth* (Tarski, 1944).

Our re-analysis of (Nix) proceeds as follows. In Sect. 2, we review Tarski's theory of truth (Tarski, 1944) as a paradigm we follow for *logical* analysis of natural language, including *counterfactuals*. On a closer look, the paradigm turns out to be along the lines of the 'meaning-as-use' approach (Wittgenstein, 2001) and what later has been called 'inferentialism' (Brandom, 1994). In Sect. 3, following the procedure of the paradigm, we observe that there are at least three types of use of counterfactuals: historical/real, rhetorical, and theoretical. In Sect. 4, we select historical/real counterfactuals, of which (Nix) is typical, and specify an inference schema that constrains their use, i.e. *counterfactual transitive inference* (CT), which was dismissed as *the fallacy of transitivity* in one manner of formalisation by Lewis (1973) but can be validated in another manner of formalisation in his own system. In Sect. 5, we present another formalisation and deduction of (CT) in a natural extension of hybrid tense logic, i.e. a version of hybrid logic improved proof-theoretically by Braüner (2011), which we refer to as hybrid tense logic for temporal conditionals (HTL_{TC}).

[4] Cited from Lewis (1979, p. 465).

This procedure may give one the impression that the result shall only have much lower generality than that obtained using the Stalnaker–Lewis approach, which is based on Stalnaker–Lewis reasoning: there is *a uniformity of grammatical form* that is classified as *counterfactual* in natural language; *therefore*, there should be *a uniformity of logical form* that is shared amongst almost all sentences in the grammatical form.[5] (We shall briefly examine Stalnaker–Lewis reasoning in Sect. 3.) However, this impression turns out to be wrong. Instead, we shall see that the procedure brings *a new generality*, thereby suggesting the possibility of a logical reclassification of conditionals in general that might deconstruct such a grammatical dichotomy as indicative vs. subjunctive. In Sects. 6 and 7, we provide evidence for this claim by presenting a formalisation of an *indicative version* of (CT) in HTL_{TC}. In Sect. 8, we list some of many ramifications that our results would have for related works in philosophical logic and linguistics. Finally, in Sect. 9, we conclude this paper with a few more remarks about a crucial difference between our new analysis and Lewis's classical analysis.

The "Appendix" presents a natural deduction system for HTL_{TC} as a special case of Braüner's systems (Braüner, 2011).

2 Our Paradigm of Logical Analysis of Natural Language: Tarski's Theory of Truth

We can say that what Tarski's historic work (Tarski, 1944) attained is *a logical refinement* of our informal conception of the word *'true'* in natural language, which involves two points of view along the lines of the 'meaning-as-use' approach (Wittgenstein, 2001) and what later has been called 'inferentialism' (Brandom, 1994).[6][7]

[5] We can find the uniformity-oriented mindset of Stalnaker and Lewis in many instances. For example, see Stalnaker (1968, pp. 40–41) for the former and Lewis (1979, pp. 40–41) for the latter.

[6] The detection of 'meaning-as-use' and 'inferentialist' points of view in *Tarski's theory of truth* (Tarski, 1944) was inspired by Okamoto (1999), which clarifies similar points of view in *Frege's theory of quantification* (1879, 1979), preceding Brandom (1994), Tarski (1944), and Wittgenstein (2001).

[7] As an anonymous reviewer aptly pointed out, the claimed connection between Tarski and inferentialism may surprise the reader, because inferentialism is often seen to be in opposition to Tarski's notion of truth, reflecting the divide between model-theory and proof-theory. However, there is textual evidence in Tarski (1944) that would corroborate the claimed connection. Tarski (1944) considers the following suspicion of a 'vicious circle' involved in his definition of truth:

In formulating the definition we use necessarily sentential connectives, i.e., expressions like *"if... , then," "or,"* etc. They occur in the definiens; and one of them, namely, the phrase *"if, and only if"* is usually employed to combine the definiendum with the definiens. However, it is well known that the meaning of sentential connectives is explained in logic with the help of the words *"true"* and *"false"'*; for instance, we say that an equivalence, i.e., a sentence of the form *"p if, and only if, q,"* is true if either both of its members, i.e., the sentences represented by *'p'* and *'q,'* are true or both are false. Hence the definition of truth involves a vicious circle. (Tarski, 1944, p. 356)

Clearing this suspicion of the 'vicious circle', Tarski (1944) states that:

It is undoubtedly the case that a strictly deductive development of logic is often preceded by certain statements explaining the conditions under which sentences of the form *"if p, then q,"* etc., are considered true or false. (Such explanations are often given schematically, by means of the so-called truth-tables.) However, [...] these statements do *not* influence the deductive development of logic in any way. For in such

First, Tarski (1944) notes that there are *different uses* of the word *'true'* and, accordingly, *different conceptions* of it. Further, he admits that he does not understand such a question as *'What is the right conception of truth?'* (Tarski, 1944, p. 355). Then, from the various uses he selects *a use* that he thinks is intelligible and specifies what constrains the selected use, i.e. what he refers to as T-schema:

$$\text{(T) } X \textit{ is true if}, \textit{ and only if}, \ p,$$

where 'p' stands for an arbitrary sentence and 'X' is the name of the sentence that 'p' stands for. Here, we can detect a 'meaning-as-use' point of view: to specify the meaning (conception) of a word or a sentence, we should specify the use of the word or the sentence.

Second, we can say that by (T), he also specifies *a simple inference* involving *'true'*: (i) we can infer the proposition that 'p' *is true* from the proposition that p; conversely, (ii) we can infer the proposition that p from the proposition that 'p' *is true*, where 'p' is substituted for X in (T). We can then regard the *'if'* direction (i) as *the introduction rule* of *'true'* and the *'only if'* direction (ii) as *the elimination rule* of it. Here, we can detect an 'inferentialist' point of view: to specify the *logical* use of a word or a sentence, we should specify the *inference* (i.e. the *logical* context) in which the word or the sentence is used.

Thus, he first specified the requirement that a *logical* analysis of *'true'* should satisfy[8] the result should imply all the sentences of the form (T). He then proceeded to (show the outline of a way to) construct a formal language in which this requirement is met[9] one in which the selected use of *'true'* can be reconstructed formally (i.e. recursively defined in this case) so that all of the sentences of the form (T) can be asserted consistently.[10] [11]

---

a development we do *not* discuss the question whether a given sentence is true, we are *only* interested in the problem whether it is *provable*. (Tarski, 1944, p. 357; emphasis mine)

Here, Tarski (1944) seems to base his definition of truth upon *'provability'* to avoid the alleged 'vicious circle'.

8 Tarski referred to the property that is attained on completion of the first process as *'material adequateness'*: (Tarski, 1944, p. 341).

9 Tarski referred to the property that is attained on completion of the second process as *'formal correctness'*: (Tarski, 1944, p. 342).

10 His well-known distinction between the object-language and the meta-language is introduced at this stage to prevent the constructed language from giving rise to the liar paradox, i.e. such a sentence as '*"s" is true if, and only if, "s" is not true*', where "*s*" is the name of the sentence "'*s*" is not true'.

11 As an anonymous reviewer helpfully suggested, Tarski's view detected here could be extended to a more general view on inferentialism, namely, what Belnap (1962) refers to as the *synthetic* mode of explanation:

> Among formal logicians, use of the synthetic mode in logic is illustrated by Kneale and Popper (cited by Prior), as well as by Jaskowski, Gentzen, Fitch, and Curry, all of these treating the meaning of connectives as arising from the role they play in the context of formal inference. It is equally well illustrated, I think, by aspects of Wittgenstein and those who learned from him to treat the meanings of words as arising from the role they play in the context of discourse. (Belnap, 1962, pp. 130-131)

Belnap (1962) contrasts the synthetic mode with the *analytic* mode, which he thinks is represented by model-theoretic procedure (Belnap, 1962, p. 130), and states:

3 Three Types of Use of Counterfactuals: Historical/Real, Rhetorical and Theoretical

Our re-analysis of (Nix) follows the procedure of Tarski's theory of truth (Tarski, 1944) involving the 'meaning-as-use' and 'inferentialist' points of view.

From the 'meaning-as-use' point of view in Tarski (1944), we suspend the reasoning that we referred to as *Stalnaker–Lewis* reasoning in Sect. 1.2: there is *a uniformity of grammatical form* that is classified as *counterfactual* in natural language; *therefore*, there should be *a uniformity of logical form* that is shared amongst almost all sentences in the grammatical form.[12]

In fact, it is clear that Stalnaker–Lewis reasoning is not valid for cases other than counterfactuals. We present a traditional counterexample: the indefinite article. Consider *'Jemima is waiting for a mouse who lives in that hole'* $\cdots$ (a).[13] The indefinite article *'a'* in (a) can be construed as an *existential* quantifier as well as a *universal* quantifier. Notoriously, the grammatical form of a sentence containing indefinite articles is far from a decisive evidence for determining its logical form.[14] After all, we cannot rely on uniformity of the grammatical form of sentences or words, even though it is a non-negligible hint as to their logical form in many cases.

Curiously enough, at a level of conditionals in general, people seem to be fully conscious of this lesson. Today, most logicians, linguists, and analytic philosophers would never reason, *'Conditionals have the uniform grammatical form "if..., then...", so there should be a uniform logical form of conditionals in general'*.[15] On the contrary, there have been considerable efforts devoted to distinguishing conditionals as material, strict, intuitionistic, relevance, and, for that matter, indicative and counterfactual.

Why, then, do not we suspect that there might be logical differences among counterfactuals even though they have past-tense verbs in their antecedents and succedents in many languages quite uniformly? Why do not we apply the lesson about conditionals in general to counterfactual conditionals in particular *recursively* even if there should be family resemblances?

In fact, Placek and Müller (2007), Xu (1997), and Lewis himself (Lewis, 1996) detect a remarkable subclass of counterfactuals, which are referred to as '*historical* counterfactuals' by Placek and Müller (2007) and '*real* counterfactuals' by Xu (1997) and Lewis (1996). They are counterfactuals concerning '*historical* possibility' (Placek

It seems to me nearly self-evident that employment of *both* modes of explanation is important and useful. (Belnap, 1962, p. 131; emphasis mine)

Thus, we can find a view according to which model-theory and proof-theory do not exclude each other. We can then find that Tarski's procedure might also be a remarkable paradigm of such a non-exclusive view, because it suggests that Tarski's theory of truth, an origin of model-theoretic semantics, is based upon a proof-theoretic point of view. For the dependence of Tarski's definition of truth upon the notion of provability, see footnote 7 above.

[12] For this uniformity-oriented mindset of Stalnaker and Lewis, see footnote 5 above.

[13] Cited from Geach (1980, p. 91).

[14] In this regard, consider also indefinite articles in donkey sentences; for example, the *'a'* in *'Any man who owns a donkey beats it'* (cited from Geach (1980, p. 143)).

[15] We refer to this reasoning *not* as *Stalnaker–Lewis* reasoning *but* as *Stalnaker* reasoning, as, whereas Stalnaker (1968, 1975) tried to extend his theory of counterfactuals to conditionals in general, Lewis (1973) did not. In this regard, they seem to part company.

& Müller, 2007) or '*real* possibility' (Xu, 1997; Lewis, 1996). Although the characterisation of historical/real possibility is not strict but intuitive, it will suffice to borrow the characterisation stated by Lewis (1996): it is the possibility that conforms to actual history up to some past moment and the actual laws of nature as ever (Lewis, 1996, p. 552).

Then, (Nix), as repeated below, is thought to be a typical case of historical/real counterfactuals:

(Nix) If Nixon had pressed the button, there would have been a nuclear war.

In fact, (Nix) supposes historical due courses that coincide with actual history up to some past moment and subsequently proceed otherwise while conforming to the actual laws of nature.[16] Another example *'If Oswald had not killed Kennedy, then someone else would have'*··· (Osw), which is referred to in Lewis (1973), is also typical of this class; however, the sentence may very well be an example of *false* historical/real counterfactuals according to Lewis's opinion (Lewis, 1973, p. 3).

By contrast, examples of counterfactuals that are obviously *not* historical/real include *'If I were a bird, I could fly there!'*··· (b) and *'If I were you, I would retort!'*··· (y). Example (b) may well be a rhetorical expression close to a sigh showing both the fact that the speaker wants to go there but cannot and his or her regret in that regard. As for (y), it may well be more realistic to consider it a euphemistic expression for both a criticism of the other's failure to retort and advice for him or her to retort even after that time than to assign any philosophical interpretation to the phrase *'I were you'*. We loosely refer to such a class of counterfactuals as '*rhetorical* counterfactuals'.

A boundary case between historical/real and rhetorical counterfactuals is Lewis's familiar one, i.e. *'If kangaroos had no tails, they would topple over'*··· (k).[17] As pointed out by Placek and Müller (2007) exactly, it may be possible to understand the sentence by seriously considering a possibility of evolution of kangaroos without tails, but it seems again more realistic to consider that it highlights the importance of their tails for their balance (Placek & Müller, 2007, p. 177).

Now, what becomes of another of Lewis's examples, i.e. *'If gravity went by the inverse cube of distance,...'* ··· (g)?[18] This can represent a theoretical thought experiment (in physics, in this case) that might be comparable to, say, *'If this axiom in this system were replaced by another,...'*··· (ax) or *'If we modified the value of the parameter in this system,...'*··· (para), as when a logician, mathematician, or computer scientist is investigating the significance of an axiom or a parameter in a given system or developing variations of the system. If so, we may loosely refer to such counterfactuals as '*theoretical* counterfactuals'.

[16] We can also give another interpretation of (Nix): it may be a rhetoric that stresses the heavy responsibility of Nixon for military matters at the time.

[17] Cited from Lewis (1973, p. 1).

[18] Cited from Lewis (1979, p. 464).

4 Counterfactual Transitive Inference

From the (at least) three uses of counterfactuals listed in Sect. 3, we select a historical/real use, of which (Nix) is to be typical.[19] Then, from the 'inferentialist' point of view in Tarski (1944),[20] we ask the question: what inference can constrain the historical/real use of counterfactuals? We think that *counterfactual transitive inference*, i.e. *transitive inference comprising counterfactual conditionals*, can[21]:

$$\frac{\begin{array}{c}\textit{If it were the case that } \varphi\textit{, then it would be the case that } \psi. \\ \textit{If it were the case that } \psi\textit{, then it would be the case that } \chi.\end{array}}{\textit{If it were the case that } \varphi\textit{, then it would be the case that } \chi.}\ (CT)$$

The conceptual reason for our selection of this inference patten is as follows. We find that historical/real counterfactuals should be essentially transitive in a *two-fold* sense. *First*, their *temporality* should be *certainly* transitive in the sense that if the time of φ is as early as the time of ψ and the time of ψ is as early as the time of χ, then the time of φ is as early as the time of χ.[22] *Second*, the *influentialness* of the state of affairs in their antecedent should be *probably* transitive in the sense that if the state φ influences the state ψ and the state ψ influences the state χ, then the state φ influences the state χ.[23]

However, some readers familiar with Lewis's work (Lewis, 1973) might suspect that (CT) is just what he referred to as *the fallacy of transitivity* and *in fact* proved to be *invalid*, as his example intuitively shows[24]:

$$\frac{\begin{array}{c}\textit{If Otto had gone to the party, then Anna would have gone.} \\ \textit{If Anna had gone to the party, then Waldo would have gone.}\end{array}}{\textit{If Otto had gone to the party, then Waldo would have gone.}}\ (OAW)$$

The background to (OAW) is that Otto and Waldo are rivals for Anna's affections. Anna likes Otto; hence, the first premise is true. Waldo usually follows Anna around; hence, the second one is also true. However, Waldo never runs the risk of meeting Otto; hence, the conclusion is false.

However, this way of speaking is quite misleading. A little reflection will make us realise this simple fact: it is true that there are cases in which (CT) *does not* hold, whereas there are *also* cases in which it *does* hold. For example, normally, we would not bother to contrive counterexamples to the following instance of (CT), whose first

[19] Placek and Müller (2007) remarked that historical/real counterfactuals are important in real life situations. We completely agree with this remark. See Placek & Müller (2007, p. 174).

[20] For the 'inferentialist' point of view, see Sect. 2.

[21] (CT) is the abbreviation for 'counterfactual transitivity'.

[22] Here, by 'the time' we mean *not* 'the event time' *but* 'the reference time' in the sense of H. Reichenbach (1947). See Sect. 5 below.

[23] From this perspective, we can consider that Judea Pearl's probabilistic causal model (Pearl, 1999, p. 98) focusses on *probable transitivity* of counterfactuals in the latter sense; it can be conceived as a theory of *probabilistic* assessment of how the antecedent *influences* the succedent.

[24] (OAW) is the abbreviation for 'Otto, Anna, and Waldo'.

premise is (Nix), although it is possible to do so if one wishes to[25]:

$$\frac{\begin{array}{c}\textit{If Nixon had pressed the button, there would have been a nuclear war.}\\ \textit{If there had been a nuclear war, there would have been a depopulation.}\end{array}}{\textit{If Nixon had pressed the button, there would have been a depopulation.}}\ (NND)$$

Then, we claim that what is more important both logically and philosophically is to explain *why* (CT) *holds when it does hold* rather than *why it does not hold when it does not*.[26] In any case, the fact is that Lewis's system can symbolise (CT) such that it *does* hold *and* such that it *does not* hold. Here are two ways.

As is well known, Lewis symbolised '*if it were that* φ, *then it would be that* ψ' as '$\varphi \mathrel{\Box\!\!\rightarrow} \psi$'. Accordingly, the simplest way to symbolise (CT) as a whole is as follows[27]:

$$\frac{\begin{array}{c}\varphi \mathrel{\Box\!\!\rightarrow} \psi\\ \psi \mathrel{\Box\!\!\rightarrow} \chi\end{array}}{\varphi \mathrel{\Box\!\!\rightarrow} \chi}\ (FCT)$$

Thus, it is not (CT) itself but (FCT), *a symbolisation of* (CT) in Lewis's system, that is rightly referred to as *the fallacy of transitivity*, as (FCT) is *in fact invalid* in Lewis's semantics.[28] Then, (OAW) should be an instance of (FCT).

Meanwhile, there is *another symbolisation of* (CT) in Lewis's system under which the resulting schema becomes *valid* in Lewis's semantics. To see this, we interpret (CT) as an abbreviation for the following[29]

$$\frac{\begin{array}{c}\textit{If it were the case that } \varphi\textit{, then it would be the case that } \varphi \textit{ and } \psi.\\ \textit{If it were the case that } \varphi \textit{ and } \psi\textit{, then it would be the case that } \chi.\end{array}}{\textit{If it were the case that } \varphi\textit{, then it would be the case that } \chi.}\ (CT')$$

The quite natural idea behind (CT') is that in the succedent of the first premise, the condition ψ *is accumulated on* the condition φ; hence, the antecedent of the second

[25] (NND) is the abbreviation for 'Nixon, a nuclear war, and a depopulation'.

[26] The literature seem to have been focussed mainly on the latter *fallacious* cases of (CT), while this paper focusses on the former *successful* cases of (CT).

[27] (FCT) is the abbreviation for 'false counterfactual transitivity'.

[28] For the invalidity of (FCT), see Lewis (1973, pp. 32–34).

[29] If one feels that (CT') as an interpretation of (CT) is unnatural, he/she should recall *enthymeme*:: *'Socrates is human, so he is mortal'*, which should be an abbreviation for *'Socrates is human and all humans are mortal, so he is mortal'*, as is stated in Geach (1976, p. 48):

It may be necessary to add a truistic premise that can be taken as generally admitted: e.g. 'Every man is an animal' is a truism needed to reduce 'Socrates is a man, therefore Socrates is an animal' to schema (4) [Every F is G; *a* is F; therefore *a* is G].

Complementing logical elements appropriately in this manner is often an essential part of logical analysis of inferences in natural language. It often leads to a discovery of a logical mechanism underlying natural language.

premise must take over the accumulated conditions φ and ψ. Accordingly, (CT') is straightforwardly symbolised in Lewis's system as follows[30]:

$$\frac{\begin{array}{c}\varphi \mathrel{\Box\!\!\rightarrow} (\varphi \,\&\, \psi) \\ (\varphi \,\&\, \psi) \mathrel{\Box\!\!\rightarrow} \chi\end{array}}{\varphi \mathrel{\Box\!\!\rightarrow} \chi}\ (VCT)$$

Then, (VCT) is *perfectly valid* in Lewis's semantics,[31] and can be deduced in the most basic version V of Lewis's axiomatic system.[32] Thus, (NND) should be an instance of (VCT) in Lewis's system.

5 Deduction of Transitivity of Counterfactuals in Hybrid Tense Logic

By accepting (VCT) as *a* correct formalisation of (CT), we can say that Lewis's analysis certainly fulfils *a* logical refinement of our informal use and conception of historical/real counterfactuals. However, we claim that there is *another* logical refinement of historical/real counterfactuals that is much more fine-grained than Lewis's version and has *another* logical generality. The main purpose of this paper is, among other things, to demonstrate this claim.

5.1 Making Explicit Reference Times in Historical/Real Counterfactuals

In Fig. 1, we present our informal branching-time model for (Nix), which is a representative example of historical/real counterfactuals, using only classical temporal concepts. Although our model is based on the idea in Tsai (2014), which is relatively recent research on Japanese counterfactual conditionals, it is quite similar in particular to Thomason and Gupta's (Thomason & Gupta, 1980). Hence, for comparison, we also display another informal branching-time model for (Nix) in Thomason and Gupta's manner in Fig. 2 just beneath ours in Fig. 1. After explaining our informal model we shall briefly look at correspondences and differences between ours (Fig. 1) and theirs (Fig. 2) at this informal level.

In Fig. 1, the middle arrow represents the actual time series, on which the point Val represents 'the valuation time' at which we make a valuation of whether the sentence (Nix) is true or false.[33] The arrows branching (up and down, respectively) from the middle represent possible time series starting to follow different courses from the

30 (VCT) is the abbreviation for 'valid counterfactual transitivity'.

31 For the validity of (VCT), see Lewis (1973, p. 35).

32 For the deduction of (VCT), see Hosokawa (2012, pp. 21–22).

33 We deliberately avoid referring to 'the speech time' here, because when evaluating a tensed sentence, it seems unnecessary to refer to the 'speech' act *recursively* Consider '*it was raining*' $\cdots (r_1)$ and '*it is raining*' $\cdots (r_2)$. In terms of classical tense logic, (r_1) is true at the present moment if, and only if, (r_2) is true at some past moment. However, it seems unnatural to interpret this truth condition of (r_1) as saying that (r_1) is true at the time of the utterance of (r_1) if, and only if, (r_2) is true at a time of the utterance of (r_2) earlier than the time of the utterance of (r_1), as (r_1) can be true without someone uttering (r_2) at some past moment. Instead, it seems much more straightforward to interpret it as saying that (r_1) evaluates to the truth at the present moment if, and only if, (r_2) evaluates to the truth at some past moment.

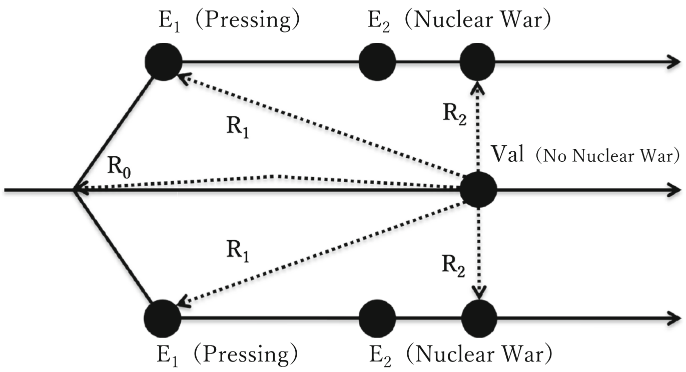

Fig. 1 Informal branching-time model of (Nix) based on the idea in Tsai (2014)

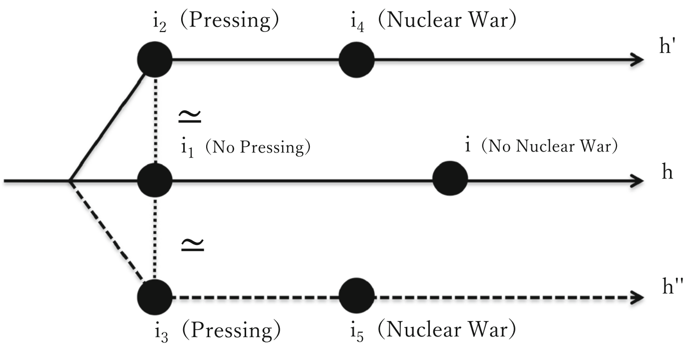

Fig. 2 Informal branching-time model of (Nix) in the manner of Thomason and Gupta (1980)

actual one at some past moment.[34] E_1 and E_2 on these arrows represent 'the event times' at which Nixon presses the button and a nuclear war occurs respectively.[35]

The hypothesis proposed by Tsai (2014) is that, in addition to these event times, there exists 'a reference time', which is a classical concept from Reichenbach (1947), in each of the antecedent and the succedent in a Japanese counterfactual conditional. The significance of this hypothesis is that, by considering the reference time in the succedent in particular, we can explain variations in tense and aspect of the succedent of Japanese counterfactual conditionals.[36]

34 Lewis (1979) refers to this 'some past moment' as 'a transition period' in Analysis 1.

35 Lewis (1979) denotes them as 't_A' and 't_C', respectively.

36 The following sentences are examples supporting the hypothesis by Tsai (2014).

(3a) Kono shigoto ga nakattara, asu ha tsuri ni *iku* noni. (If I did not have this work, I would *go* fishing tomorrow.)
(3b) Kono shigoto ga nakattara, asu ha tsuri ni *itta* noni. (If I did not have this work, I would *have gone* fishing tomorrow.)

While the succedent of (3a) is in the future tense ('tsuri ni *iku*'), that of (3b) is in the past tense ('tsuri ni *itta*'). In these cases, indicating the reference time that locates the event time E_2 of going fishing as R_2, the former can be explained by E_2 lying to the *right* of R_2, and the latter can be explained by E_2 lying to the *left* of R_2.

In this paper, we presuppose that this hypothesis holds completely for *English* counterfactual conditionals *as well*. In the present case (Nix), for example, the perfective aspect of *'have been'* in *'there would have been a nuclear war'* can be explained by the configuration shown in Fig. 1 in which E_2 lies to the *left* of R_2. In addition, our hypothesis is that there exists *another* implicit reference time in (Nix), in addition to R_1 and R_2: a diverging time point R_0.

However, Fig. 1 involves technical improvements to the original version of model representation in Tsai (2014). Although Tsai (2014) represents the two reference times R_1 (in the antecedent) and R_2 (in the succedent) as *points* straightforwardly, we represent them as *relations*, which we refer to as *'time reference relations'*. This is because, while the original version of model representation in Tsai (2014) depicts only a single possible time series other than the actual one, the branching-time model *qua* the strict mathematical model can have multiple possible divergences from the actual time series. In such a case, a reference time *may* be located on *each* of the multiple possible time series. More specifically, in the present case, each R_1 and R_2 as respective *points* lies on each of the possible time series; hence, there are multiple R_1s and R_2s as *points* in the above-mentioned model as a whole. The valuator of (Nix) then refers to all multiple R_1s and R_2s located on different time series at the valuation time. To this end, in this paper we refine *reference time points* R_1, R_2 into *time reference relations* R_1, R_2 from the valuation time, which represent *the distribution* of multiple R_1s, R_2s as *respective points*. Similarly, we technically interpret the diverging time *point* R_0 as the diverging time reference *relation* R_0 from the valuation time Val.

Thus, we can see that (Nix) can be interpreted intuitively as 'in any possible time series branching off from the actual one at some past moment referred to by R_0, if a moment is referred to by R_1 and E_1 occurs at that moment, then, thereafter, if a moment is referred to by R_2, then E_2 has occurred by that moment', as shown in Fig. 1.

Now, let us compare our model (Fig. 1) with one represented in Thomason and Gupta's manner (Fig. 2) at the informal level.

In Fig. 2, (Nix) is evaluated at moment i in history h at which no nuclear war has occurred. Moment i_1 on h is a past moment of i at which Nixon had a chance to press the button but did not do so. The counterfactual histories h' and h'' are those in which Nixon pressed the button at moments i_2 and i_3, respectively, which are the 'alternative presents' to i_1 and said to be 'copresent' with i_1. This 'copresence' relation is indicated by $\simeq$. The solid line indicating h' through i_2 and the dotted line indicating h'' through i_3 show that the pair $< i_2, h' >$ is selected to be the closest pair to $< i_1, h >$ at which *'Nixon presses the button'* is true, and the pair $< i_3, h'' >$ is not. ("The closest pair to $< i_1, h >$ at which *'Nixon presses the button'* is true" is determined by the "selection function".) Then, by the existence of i_4 on h', at which there is a nuclear war, *'There will have been a nuclear war'* is true at the closest pair $< i_2, h' >$ to $< i_1, h >$. Because such a moment i_1 is found to be in the past of i on h, (Nix) is considered to be true at i on h in Fig. 2.

We can then list some correspondences and differences between Figs. 1 and 2 that would already be clear. *Correspondences*: Val corresponds to i, E_1 corresponds to i_2 and i_3, E_2 corresponds to i_4 and i_5, and R_1 is comparable to $\simeq$ in the sense that both relate the valuation time (Val, i) to the counterfactual event time of Nixon's pressing the button (E_1, i_2/i_3), where $\simeq$ does so via some past moment i_1 of i. *Differences*:

R_0 and R_2 have no counterparts in Fig. 2; conversely, i_1 has no counterpart in Fig. 1; furthermore, while 'the closest' counterfactual history (time series) to the actual history is selected in Fig. 2, it is not selected in Fig. 1.

After presenting our formal language and model in Sects. 5.2 and 5.3, we shall make a more formal comparison between our formalism and Thomason and Gupta's in Sect. 5.4.

5.2 Symbolising Historical Counterfactuals by Time Reference Operators

Note that Fig. 1, regarded as a branching-time model, comprises (1) the temporal order relation among moments, and (2) the time reference relations from the valuation time to the reference times. We then find that classical tense logic from Prior (1955), as it is, applies to (1). For (2), we note that tense logic can be regarded as a type of multi-modal logic that sets the label of the future relation as F and that of the past relation as P and, accordingly, introduces the indexed modal operators $\langle F \rangle$ ('at some time in the future') and $\langle P \rangle$ ('at some time in the past')—the abbreviations of which are F and P, respectively, as modal operators—into language. In accordance with this precedent, we can also convert (2) into symbols in a straightforward way; we prepare the labels R_n ($n \in \{0, 1, 2, \ldots\}$) of the time reference relations and then introduce the indexed modal operators $\langle R_n \rangle$ ($n \in \{0, 1, 2, \ldots\}$), which we refer to as *'time reference operators'*, into language.

However, there are some complications to this approach. As seen from Fig. 1, a time reference relation is a relation that starts from the current valuation time Val and crosses time series; hence, it is connected independently of the temporal order. Although we omit the explanation of the difficulties and the details of the trial-and-error approach while designing the syntax,[37] we mention that it is difficult to express the behaviour of a time reference relation sufficiently by means of a simple form of classical tense logic.

To describe its behaviour adequately, we use (i) the downarrow binder $\downarrow$, an early version of which was introduced by Valentin Goranko (1994,1996), and (ii) the inverse operator $\langle R_n^- \rangle$ of the time reference operator $\langle R_n \rangle$. By means of the former, we can construct a sentence of the form '$\downarrow a.\varphi$' with which we can express 'Name the current state 'a' and φ holds at a', where φ is an arbitrary formula that is possibly open; hence, a can occur in φ such that we can make a reflexive reference to the current state a. For example, '$\downarrow a.GPa$' is a (valid closed) formula that reads as 'Name the current state 'a' and it henceforth holds that a held once'.

The latter, $\langle R_n^- \rangle$, corresponds to the inverse relation R_n^- of the time reference relation R_n, where R_n^- expresses the *passive* reference relation '*being referred to* by R_n' and R_n expresses the active reference relation 'referring to by R_n' (for the strict satisfaction

[37] Our trial-and-error approach is based on Blackburn (1990; 1994) and Blackburn and Jørgensen (2016) ('Reichenbach, Prior and Hybrid Tense Logic'), which use nominals for temporal references. However, nominals are true at one and only one time, i.e. each nominal refers to (or 'name') a unique time. Therefore, by definition, nominals *cannot* refer to multiple reference time points located on multiple possible time series at once *by themselves*, i.e. we *cannot* express the distribution of multiple reference time points across multiple possible time series by using *only* nominals. This is why we use nominals *combined with* the downarrow binder $\downarrow$ and the inverse $\langle R_n^- \rangle$ of the time reference operator $\langle R_n \rangle$ introduced below.

relation, see Sect. 5.3 below). Note that this extension is also only an imitation of the precedent set by Prior (1955, 1967) in the sense that the pair of F and P is a typical instance of pairs of modal operators that are the inverse operators of each other.

Through the extension of classical tense logic by the two items (i) and (ii) above, for instance, '$\downarrow a.PF\langle \mathrm{R}_n^- \rangle a$' can express 'Name the current state 'a' and at some moment in the past there exists some moment in the future that is being referred to by R_n from a', i.e. 'In some possible future course starting from some moment in the past of the current state a, there exists a reference time R_n from a'.

Using this expressive power, we consider a symbolisation of (Nix) that reflects Fig. 1. The most straightforward way to describe it is as follows, where p:= '*Nixon presses the button*' q:= '*There is a nuclear war*', and G is the future-necessity tense operator meaning '*At any moment in the future,*'.

$$\begin{aligned} &\downarrow a.P(\langle \mathrm{R}_0^- \rangle a \ \wedge \\ &\qquad \text{(a)} \ \ F(\langle \mathrm{R}_1^- \rangle a \ \wedge \ p \ \wedge \ F\langle \mathrm{R}_2^- \rangle a) \ \wedge \\ &\qquad \text{(b)} \ \ G(\langle \mathrm{R}_1^- \rangle a \rightarrow (p \rightarrow G(\langle \mathrm{R}_2^- \rangle a \rightarrow Pq)))) \qquad \cdots \text{(Nix-S)} \end{aligned}$$

(Nix-S) says that there is a past moment R_0 such that (a) in a future course proceeding from it, p holds at R_1 and, thereafter, R_2 will come, and (b) in any future course proceeding from it, if p holds when R_1 comes, then q will have occurred when R_2 comes.

(Nix-S) is surely the strongest version of symbolisation of (Nix) (where 'S' denotes 'strong'). The speaker of (Nix) may *not* assume the existence of all or some of the reference time points R_0, R_1, and R_2. For instance, the weakest version is surely the following, where H is the past-necessity tense operator meaning '*At any moment in the past,*':

$$\downarrow a.H(\langle \mathrm{R}_0^- \rangle a \rightarrow G(\langle \mathrm{R}_1^- \rangle a \rightarrow (p \rightarrow G(\langle \mathrm{R}_2^- \rangle a \rightarrow Pq)))) \qquad \cdots \text{(Nix-W)}$$

(Nix-W) says that *if* there is a diverging time point R_0 in the past, *then*, *if* R_1 comes in the future after R_0 and p holds at R_1, *then*, *if* R_2 comes in the future after R_1, *then* q will have occurred at R_2. Thus, we can find that there are various options for the symbolisation of (Nix) depending on which of R_0, R_1, and R_2 is assumed by the speaker to exist in his/her model.

In the next section, we specify our language and its model formally.

5.3 HTL$_{TC}$ and Its Branching-Time Model

Let $\mathsf{Prop} = \{p, q, r, \ldots\}$ and $\mathsf{Nom} = \{a, b, c, \ldots\}$ be countably infinite sets of propositional variables and nominals (which we may regard as names of moments), respectively, which are mutually disjoint. Further, let $\{\mathrm{R}_n\}_{n \in \mathbb{N}}$ be a countably infinite set of relation symbols.

We then define the language to symbolise (Nix-S, W) using the syntax below, where $N \in \{G, H, [\mathrm{R}_n], [\mathrm{R}_n^-]\}$:

$$\varphi ::= p \mid a \mid \varphi \wedge \psi \mid \varphi \rightarrow \psi \mid \bot \mid N\varphi \mid @_a\varphi \mid \ \downarrow a.\varphi$$

For other propositional connectives, we define $\neg\varphi \equiv \varphi \rightarrow \bot$, $\top \equiv \neg\bot$, $\varphi \vee \psi \equiv \neg(\neg\varphi \wedge \neg\psi)$, and $\varphi \leftrightarrow \psi \equiv (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. For the future-possibility operator F and the past-possibility operator P, we define $F\varphi \equiv \neg G\neg\varphi$ and $P\varphi \equiv \neg H\neg\varphi$ derivatively from G and H, respectively. Similarly, we set $\langle \mathrm{R}_n\rangle\varphi \equiv \neg[\mathrm{R}_n]\neg\varphi$, $\langle \mathrm{R}_n^-\rangle\varphi \equiv \neg[\mathrm{R}_n^-]\neg\varphi$.

This language is only a small extension of hybrid tense logic $\mathcal{H}(\downarrow)$ given by Braüner (2011), which can be obtained by adding time reference operators and their inverses ($[\mathrm{R}_n]$, $[\mathrm{R}_n^-]$), as well as classical tense operators (G, H), to it. We refer to this language as *hybrid tense logic for temporal conditionals* and denote it by HTL_{TC}.

We next present a standard Kripke model for HTL_{TC}. First, we prepare a Kripke frame $\mathcal{T} = (T, \{<\} \cup \{\mathrm{R}_n\}_{n\in\mathbb{N}})$ comprising a set T of moments and relations $<$, R_n ($n \in \mathbb{N}$) among these moments. Given a valuation V for propositional variables, the pair $\mathcal{M} = (\mathcal{T}, V)$ is a Kripke model for HTL_{TC}.

In the above, $<$ represents the strict temporal order on T, which is assumed to be transitive.[38]

It is also assumed to be *past-ward* linear, i.e. it satisfies the following condition:

Past-ward Linearity

$$\forall t, t', t'' \in T \ ((t' < t \text{ and } t'' < t) \Rightarrow (t' < t'' \text{ or } t' = t'' \text{ or } t'' < t'))$$

This expresses a general constraint on the branching-time model in that it can branch forward but not backward, reflecting the idea that the past is determined and therefore linear, whereas the future is undetermined and branches into many possible time series.

Furthermore, we assume the following condition for the interaction between $<$ and R_n[39]

Intermediateness of Time Reference

$$(\mathrm{R}_n <^{\exists} \mathrm{R}_{n+1} < \mathrm{R}_{n+2})$$

$$\forall t, t', t'' \in T \ (t\mathrm{R}_n t' \text{ and } t' < t'' \text{ and } t\mathrm{R}_{n+2}t''$$

$$\Rightarrow \exists t''' \in T \, (t' < t''' \text{ and } t\mathrm{R}_{n+1}t''' \text{ and } t''' < t'')),$$

which is equivalent to the R_n-*inverse* version:

$$(\mathrm{R}_n^- <^{\exists} \mathrm{R}_{n+1}^- < \mathrm{R}_{n+2}^-)$$

38 By this, we do not mean that time structure should generally be strict and transitive. Rather, this only means that for our analysis it is sufficient to assume strictness and transitivity besides the past-ward linearity and intermediateness of time reference introduced below.

39 We can consider this condition to be a type of relativisation of the condition of time being *dense*: by time reference relations: if we omit R_n, R_{n+1}, and R_{n+2} from this condition, we obtain the condition of $<$ *in itself* being *dense*, i. e.

$$\forall t', t'' \in T \ (t' < t'' \Rightarrow \exists t''' \in T \, (t' < t''' \text{ and } t''' < t'')).$$

I owe this observation to a suggestion by an anonymous reviewer.

$$\forall t, t', t'' \in T \ (t'\mathrm{R}_n^- t \text{ and } t' < t'' \text{ and } t''\mathrm{R}_3^- t \\ \Rightarrow \exists t''' \in T (t'' < t''' \text{ and } t'''\mathrm{R}_2^- t \text{ and } t''' < t'')).$$

An instance of this condition is

$$(\mathrm{R}_1^- <^{\exists} \mathrm{R}_2^- < \mathrm{R}_3^-) \\ \forall t, t', t'' \in T \ (t'\mathrm{R}_1^- t \text{ and } t' < t'' \text{ and } t''\mathrm{R}_3 t' \\ \Rightarrow \exists t''' \in T (t'' < t''' \text{ and } t'''\mathrm{R}_2^- t \text{ and } t''' < t'')),$$

whose corresponding inference rule is $(R_1^- <^{\exists} R_2^- < R_3^-)$ shown in Table 1 in Sect. 5.5, which, in turn, plays a crucial role in our formal deduction of counterfactual transitivity.

With the language and its model above, we state the satisfaction conditions only of the downarrow binder $\downarrow$, the time reference operator $\langle \mathrm{R}_n \rangle$, and its inverse $\langle \mathrm{R}_n^- \rangle$. The remaining items are defined as usual.

For $\downarrow$, we introduce a function g from nominals to moments and relativise the evaluation of formulas to g as follows:

$$\mathcal{M}, g, t \models \downarrow a.\varphi \text{ iff } \mathcal{M}, g[a \mapsto t], t \models \varphi,$$

where $g[a \mapsto t]$ assigns the same moments to the nominals as g except that $g[a \mapsto t]$ assigns t to a.

The satisfaction conditions for $[\mathrm{R}_n]$ and $[\mathrm{R}_n^-]$ are given as follows:

$$\mathcal{M}, g, t \models [\mathrm{R}_n]\varphi \ \text{ iff } \forall t' \in T \ (t\mathrm{R}_n t' \Rightarrow \mathcal{M}, t' \models \varphi), \\ \mathcal{M}, g, t \models [\mathrm{R}_n^-]\varphi \ \text{ iff } \forall t' \in T \ (t'\mathrm{R}_n t \Rightarrow \mathcal{M}, t' \models \varphi).$$

A natural deduction system for HTL_{TC}, which is, again, only a special case of Braüner's system (Braüner, 2011), is presented in "Appendix A". In particular, note that all rules corresponding to conditions on the accessibility relation, listed in Table 5 in "Appendix A", are instances of the form for *geometric theories*,[40]

5.4 Comparison with Thomason and Gupta's Theory

Our resulting formalism is similar to that of Thomason and Gupta (1980), which analyses counterfactual conditionals (and conditionals in general) using branching-time structures. We list *three* similarities between our approach and theirs that illuminate the respective formalisms:

First, we share with Thomason and Gupta (1980) the idea of using basic formalism of classical tense logic originating from Prior (1955, 1967) to analyse counterfactual conditionals (and conditionals in general).

[40] For the form, see Braüner (2011, p. 29). For the definition of *a geometric theory* see also Braüner (2011, pp. 28–30).

Second, their symbolisation of a supposedly typical example of historical/real counterfactuals shows signs of our Reichenbachian idea based on Tsai (2014): a historical/real counterfactual conditional generally has *an* implicit reference time in *each* of the antecedent and the succedent.[41]

Third, there is a technical correspondence. In the semantics of Thomason and Gupta (1980), the 'copresence' relation $\simeq$ is used to refer to the 'alternative presents' at which the antecedent of a conditional holds.[42] We can think that this 'copresence' relation $\simeq$ is the counterpart of our 'time reference' relation R_n.

However, there are crucial differences between our formalism and that of 1980, of which we list *three* that also illuminate both formalisms:

First, the semantics of Thomason and Gupta (1980) involves quantification over histories (or scenarios), i.e. sets of moments, which ours does not.[43] This implies that their semantics departs from the standard Kripke semantics, while ours does not.

Second, following Stalnaker (1968), Thomason and Gupta (1980) used the special single connective $>$ for the conditional, and present complex semantics with it.[44] (In short, the syntax is simple but the semantics is complex.) By contrast, we dispense with such a special connective. Instead, we use material implication $\rightarrow$ *combined with* the downarrow binder, the classical tense operators, and their natural (Reichenbachian) extensions (namely the time reference operators and their inverses) and present straightforward Kripke semantics with them.[45] (In short, the syntax is complex but the semantics is straightforward.)

Third, our formalism has a specific purpose not shared by that of Thomason and Gupta (1980)[46]: to explain *why counterfactual transitive inference* (CT) *holds when it does hold*.[47] We pursued this proof-theoretically as well as model-theoretically, as discussed in the next section.

---

[41] The example is

(Max) If Max missed the train, he would have taken the bus,

which Thomason and Gupta (1980) symbolise as

(Max') $P(p > Fq)$,

where P is the past tense operator, F is the future tense operator, and $>$ is the conditional operator originating from Stalnaker (1968). If we interpret p:= '*Max misses the train*' and q:= '*Max* <u>*has taken / took*</u> *the bus*' (the authors do not specify the interpretation of q as well as p precisely, but we may interpret q as *perfective aspect or past tense*, as, in the succedent of (Max), 'have taken' follows 'would', which the authors suggest that corresponds to F in the scope of P in (Max')), then we can regard the *absence* and the *presence* of F in the antecedent p and the succedent Fq of the conditional $>$ respectively as reflecting that there are reference times in the antecedent and the succedent of (Max) that locate the two events p and q, respectively. See Thomason & Gupta (1980, pp. 299–300).

[42] For the 'copresence' relation, see Thomason & Gupta (1980, p. 305).

[43] For the use of quantification over histories in Thomason and Gupta's semantics, see Thomason & Gupta (1980, pp. 305–306).

[44] For the semantics of Thomason and Gupta's version of $>$, see Thomason & Gupta (1980, pp. 305–306, 310–311).

[45] For our syntax and semantics, see Sects. 5.1–5.3 above.

[46] This does not imply that the theory of Thomason and Gupta (1980) cannot fulfil this purpose.

[47] For Lewis's explanation as to this question, see Sect. 4.

5.5 Deduction of Counterfactual Transitivity in HTL_{TC}

Now, we are prepared for a formalisation of (CT) in a manner different from Lewis's approach. We symbolise three sentence-schemas constituting (CT) in HTL_{TC} in the same way as that for (Nix) into (Nix-W), which is much simpler than (Nix-S). The resulting schema is as follows[48]:

$$\frac{\begin{array}{c}\downarrow a.H(\langle \mathrm{R}_0^- \rangle a \to G(\langle \mathrm{R}_1^- \rangle a \to (\varphi \to G(\langle \mathrm{R}_2^- \rangle a \to \psi)))) \\ \downarrow a.H(\langle \mathrm{R}_0^- \rangle a \to G(\langle \mathrm{R}_2^- \rangle a \to (\psi \to G(\langle \mathrm{R}_3^- \rangle a \to \chi))))\end{array}}{\downarrow a.H(\langle \mathrm{R}_0^- \rangle a \to G(\langle \mathrm{R}_1^- \rangle a \to (\varphi \to G(\langle \mathrm{R}_3^- \rangle a \to \chi))))}\ (SCT)$$

We note two points. First, all three sentence-schemas of (SCT) share *the same* diverging time point R_0. This implies that the speaker of (CT) *keeps the context unchanged throughout his/her inference*. Second, (SCT) involves *three* reference times — R_1, R_2 and R_3 — in addition to R_0, where R_2 is *both* the reference time in the *succedent* of *the first* premise *and* the one in the *antecedent* of *the second*. This reflects an obvious idea on temporal relation in the whole inference structure: R_2 *should be located between* R_1 *and* R_3.

We can then construct the formal proof of (SCT) in the natural deduction system for HTL_{TC}, which is a special case of Braüner's system (Braüner, 2011). In Fig. 3, Φ, Ψ, and X are subformulas of the three formulas of (SCT), which exclude $\downarrow a.$, and $\Phi[i/a]$ is the result of replacing all the occurrences of a in Φ by i. Figure 4 elucidates the meaning of each step of the proof.

Let us observe the deduction in Fig. 3. Every inference rule used in each step has been made explicit. Thus, we find that, although the deduction is lengthy, nearly all of the steps are routine practices of standard natural deduction and the rules on the accessibility relation are used in only two steps. One is common in modal logic, i.e. $(Trans\text{-}F)$, corresponding to *transitivity* of future relation (see Table 5 in "Appendix A"). The other is unique to the present context, which requires that (CT) and, accordingly, (SCT) should hold. It is as follows:
This rule (Table 1) is also an instance of the form for *geometric theories*,[49] The meaning is obvious: if reference times R_1 and R_3 are on the same timeline, then there is a reference time R_2 between them.

We note two points. (1) The routine part of the deduction comprises eliminations and introductions of $\to$ and the classical tense operators H, G in addition to two eliminations and one introduction of the $\downarrow$ binder. This means that, essentially, (SCT) is an inference involving *nested*-tense operators interposed among implications. (2) The *'nested-tense'* inference presupposes a *special* accessibility condition, i.e. *intermediateness* of R_2 between R_1 and R_3, as well as a *usual* condition, i.e. *transitivity* of the temporal order relation $<$.

This is our answer to the question *why* (CT) *holds when it does hold*. In summary, it is by virtue of (1) the elimination/introduction rules of the $\downarrow$ binder and implication

[48] (SCT) is the abbreviation for 'successful counterfactual transitivity'.

[49] For the definition of *a geometric theory* see footnote 40 above.

$$\Pi_1$$

$$\dfrac{\dfrac{\dfrac{[@_a i]^9 \qquad @_a \downarrow a.\Phi}{@_i \Phi[i/a]}\,(\downarrow E) \qquad [@_i Pb]^8}{@_b(\langle \mathrm{R}_0^- \rangle i \to G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to G(\langle \mathrm{R}_2^- \rangle i \to \psi))))}\,(HE) \qquad [@_b \langle \mathrm{R}_0^- \rangle i]^7}{@_b G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to G(\langle \mathrm{R}_2^- \rangle i \to \psi)))}\,(\to E)$$

$$\Pi_2$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\Pi_1 \qquad [@_b Fc]^6}{@_c(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to G(\langle \mathrm{R}_2^- \rangle i \to \psi)))}\,(GE) \qquad [@_c \langle \mathrm{R}_1^- \rangle i]^5}{@_c(\varphi \to G(\langle \mathrm{R}_2^- \rangle i \to \psi))}\,(\to E) \qquad [@_c \varphi]^4}{@_c G(\langle \mathrm{R}_2^- \rangle i \to \psi)}\,(\to E) \qquad [@_c Fd]^1}{@_d(\langle \mathrm{R}_2^- \rangle i \to \psi)}\,(GE)$$

$$\Pi_3$$

$$\dfrac{\dfrac{\dfrac{[@_a i]^9 \qquad @_a \downarrow a.\Psi}{@_i \Psi[i/a]}\,(\downarrow E) \qquad [@_i Pb]^8}{@_b(\langle \mathrm{R}_0^- \rangle a' \to G(\langle \mathrm{R}_2^- \rangle i \to (\psi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi))))}\,(HE) \qquad [@_b \langle \mathrm{R}_0^- \rangle i]^7}{@_b G(\langle \mathrm{R}_2^- \rangle i \to (\psi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi)))}\,(\to E)$$

$$\Pi_4$$

$$\dfrac{\dfrac{\Pi_3 \qquad \dfrac{[@_b Fc]^6 \qquad [@_c Fd]^1}{@_b Fd}\,(Trans\text{-}F)}{@_d(\langle \mathrm{R}_2^- \rangle i \to (\psi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi)))}\,(GE) \qquad [@_d \langle \mathrm{R}_2^- \rangle i]^1}{@_d(\psi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi))}\,(\to E)$$

$$\Pi_5$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\Pi_2 \qquad [@_d \langle \mathrm{R}_2^- \rangle i]^1}{@_d \psi}\,(\to E) \qquad \Pi_4}{@_d G(\langle \mathrm{R}_3^- \rangle i \to \chi)}\,(\to E) \qquad [@_d Fe]^1}{@_e(\langle \mathrm{R}_3^- \rangle i \to \chi)}\,(GE) \qquad [@_e \langle \mathrm{R}_3^- \rangle i]^2}{@_e \chi}\,(\to E)$$

The Whole Deduction

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[@_c \langle \mathrm{R}_1^- \rangle i]^5 \qquad [@_c Fe]^3 \qquad [@_e \langle \mathrm{R}_3^- \rangle i]^2 \qquad \Pi_5}{@_e \chi}\,(R_1 <^{\exists} R_2 < R_3)^1}{@_e(\langle \mathrm{R}_3^- \rangle i \to \chi)}\,(\to I)^2}{@_c G(\langle \mathrm{R}_3^- \rangle i \to \chi)}\,(GI)^3}{@_c(\varphi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi))}\,(\to I)^4}{@_c(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi)))}\,(\to I)^5}{@_b G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi)))}\,(GI)^6}{@_b(\langle \mathrm{R}_0^- \rangle i \to G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to G(\langle \mathrm{R}_3^- \rangle i \to \chi))))}\,(\to I)^7}{@_a X[i/a]}\,(GI)^8}{@_a \downarrow a.X}\,(\downarrow I)^9$$

Fig. 3 Natural deduction derivation of (SCT)

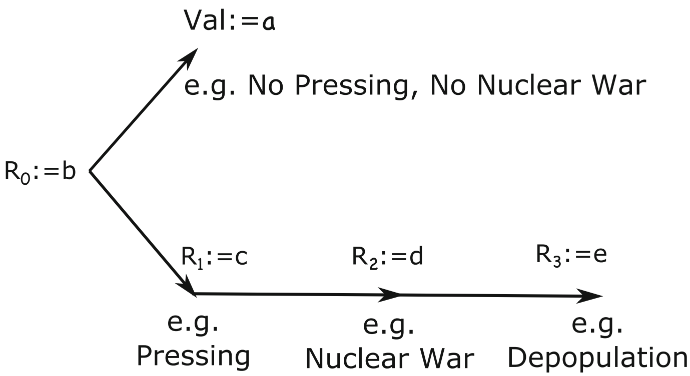

Fig. 4 Intuitive understanding of natural deduction derivation of (SCT)

Table 1 Special accessibility rule: $(R_1^- <^{\exists} R_2^- < R_3^-)$

$$\dfrac{@_c\langle \mathrm{R}_1^-\rangle a \qquad @_c Fe \qquad @_e\langle \mathrm{R}_3^-\rangle a \qquad \begin{array}{c}[@_c Fd] \quad [@_d\langle \mathrm{R}_2^-\rangle a] \quad [@_d Fe]\\ \vdots \\ \varphi\end{array}}{\varphi}\ (R_1^- <^{\exists} R_2^- < R_3^-)^*$$

*d does not occur free in φ or in any undischarged assumptions other than the specified occurrences of $@_c Fd$, $@_d\langle \mathrm{R}_2^-\rangle a$ and $@_d Fe$

and classical tense operators and (2) *transitivity* of the temporal order relation and *intermediateness* of the second reference time between the first and third reference times.

As an immediate consequence, we can translate (NND), an instance of (CT), into (NND') below, which is, in turn, an instance of (SCT). In the following, $\varphi{:=}p$, $\psi{:=}Pq$, and $\chi{:=}Pr$, where p:= '*Nixon presses the button*', q:= '*There is a nuclear war*', r:= '*There is a depopulation*', and P is the past-possibility tense operator meaning '*At some moment in the past,*'.

$$\dfrac{\begin{array}{c}\downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (p \to G(\langle \mathrm{R}_2^-\rangle a \to Pq))))\\ \downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_2^-\rangle x \to (Pq \to G(\langle \mathrm{R}_3^-\rangle a \to Pr))))\end{array}}{\downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (p \to G(\langle \mathrm{R}_3^-\rangle a \to Pr))))}\ (NND')$$

Incidentally, when $R_1{=}R_2{=}R_3$, i.e. when the reasoner of (CT) assumes that the reference times in the antecedent and succedent remain the same throughout his/her inference, the result becomes simpler. In such a case, (CT) is symbolised as

$$\dfrac{\begin{array}{c}\downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (\varphi \to \psi)))\\ \downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (\psi \to \chi)))\end{array}}{\downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (\varphi \to \chi))),}\ (SCT[R_1{=}R_2{=}R_3])$$

and can be deduced as in Fig. 5 below.

$$\Pi_1$$

$$\dfrac{\dfrac{\dfrac{[@_a i]^6 \qquad @_a \downarrow a.\Phi}{@_i \Phi[i/a]}\,(\downarrow E) \qquad [@_i Pb]^5}{@_b(\langle \mathrm{R}_0^- \rangle i \to G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to \psi)))}\,(HE) \qquad [@_b \langle \mathrm{R}_0^- \rangle i]^4}{@_b G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to \psi))}\,(\to E)$$

$$\Pi_2$$

$$\dfrac{\dfrac{\dfrac{\Pi_1 \qquad [@_b Fc]^3}{@_c(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to \psi))}\,(GE) \qquad [@_c \langle \mathrm{R}_1^- \rangle i]^2}{@_c(\varphi \to \psi))}\,(\to E) \qquad [@_c \varphi]^1}{@_c \psi}\,(\to E)$$

$$\Pi_3$$

$$\dfrac{\dfrac{\dfrac{[@_a i]^6 \qquad @_a \downarrow a.\Psi}{@_i \Psi[i/a]}\,(\downarrow E) \qquad [@_i Pb]^5}{@_b(\langle \mathrm{R}_0^- \rangle i \to G(\langle \mathrm{R}_1^- \rangle i \to (\psi \to \chi))))}\,(HE) \qquad [@_b \langle \mathrm{R}_0^- \rangle i]^4}{@_b G(\langle \mathrm{R}_1^- \rangle i \to (\psi \to \chi))}\,(\to E)$$

$$\Pi_4$$

$$\dfrac{\dfrac{\Pi_3 \qquad [@_b Fc]^3}{@_c(\langle \mathrm{R}_1^- \rangle i \to (\psi \to \chi))}\,(GE) \qquad [@_c \langle \mathrm{R}_1^- \rangle i]^5}{@_c(\psi \to \chi)}\,(\to E)$$

The Whole Deduction

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Pi_2 \qquad \Pi_4}{@_c \chi}\,(\to E)}{@_c(\varphi \to \chi))}\,(\to I)^1}{@_c(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to \chi))}\,(\to I)^2}{@_b G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to \chi))}\,(GI)^3}{@_b(\langle \mathrm{R}_0^- \rangle i \to G(\langle \mathrm{R}_1^- \rangle i \to (\varphi \to \chi)))}\,(\to I)^4}{@_i X[i/a]}\,(GI)^5}{@_a \downarrow a.X}\,(\downarrow I)^6$$

Fig. 5 Natural deduction derivation of $(SCT[R_1 = R_2 = R_3])$

As seen in Fig. 5, $(Trans\text{-}F)$ and $(R_1^- <^{\exists} R_2^- < R_3^-)$ are not used anymore, and therefore our answer can be simplified as follows: when the distinction between the reference times in the antecedent and succedent collapses, it is by virtue of the elimination/introduction rules of the ↓ binder and implication and the classical tense operators that (CT) holds when it does hold. In other words, in such a case we can say that (CT) is almost purely a *'nested-tense'* inference with no additional accessibility conditions required.

In addition, by the soundness result in Braüner (2011), $(SCT[R_1 = R_2 = R_3])$ is valid in the class of *all* frames of $<$ and R_n, as is (SCT) in the class of frames satisfying *transitivity* of $<$ (corresponding to the rule $(Trans\text{-}F)$) and *intermediateness* of R_2 between R_1 and R_3 (corresponding to the rule $(R_1^- <^{\exists} R_2^- < R_3^-)$).[50]

6 Deduction of Indicative Transitivity in HTL$_{TC}$

We now consider the following syllogism, assuming that it was spoken in 1962:

$$\frac{\begin{array}{c}\textit{If Nixon presses the button, there will be a nuclear war.}\\ \textit{If there is a nuclear war, there will be a depopulation.}\end{array}}{\textit{If Nixon presses the button, there will be a depopulation.}}\ (NND\text{-}I)$$

As is seen, this is an *indicative* version of (NND). We can symbolise this in HTL$_{TC}$ as follows, where again p:= '*Nixon presses the button*' q:= '*There is a nuclear war*', and r:= '*There is a depopulation*':

$$\frac{\begin{array}{c}\downarrow a.G(\langle \mathrm{R}_1^- \rangle a \to (p \to G(\langle \mathrm{R}_2^- \rangle a \to q)))\\ \downarrow a.G(\langle \mathrm{R}_2^- \rangle a \to (q \to G(\langle \mathrm{R}_3^- \rangle a \to r)))\end{array}}{\downarrow a.G(\langle \mathrm{R}_1^- \rangle a \to (p \to G(\langle \mathrm{R}_3^- \rangle a \to r)))}\ (NND\text{-}I')$$

$(NND\text{-}I)$ and $(NND\text{-}I')$ can be regarded as instances of (IT) and (SIT), respectively[51][52]:

$$\frac{\begin{array}{c}\textit{If it is the case that } \varphi, \textit{ then it will be the case that } \psi.\\ \textit{If it is the case that } \psi, \textit{ then it will be the case that } \chi.\end{array}}{\textit{If it is the case that } \varphi, \textit{ then it will be the case that } \chi.}\ (IT)$$

$$\frac{\begin{array}{c}\downarrow a.G(\langle \mathrm{R}_1^- \rangle a \to (\varphi \to G(\langle \mathrm{R}_2^- \rangle a \to \psi)))\\ \downarrow a.G(\langle \mathrm{R}_2^- \rangle a \to (\psi \to G(\langle \mathrm{R}_3^- \rangle a \to \chi)))\end{array}}{\downarrow a.G(\langle \mathrm{R}_1^- \rangle a \to (\varphi \to G(\langle \mathrm{R}_3^- \rangle a \to \chi)))}\ (SIT)$$

(SIT), in turn, can be presented with a formal deduction in HTL$_{TC}$, which we omit, as it can be executed in a manner similar to and simpler than (SCT).

[50] For the soundness result, see Braüner (2011, pp. 32–33).

[51] (IT) is the abbreviation for 'indicative transitivity'.

[52] (SIT) is the abbreviation for 'successful indicative transitivity'.

7 A New Generality: Emergence of 'Temporal' Conditionals

Compare (SCT) with (SIT):

$$\frac{\begin{array}{c}\downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (\varphi \to G(\langle \mathrm{R}_2^-\rangle a \to \psi))))\\ \downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_2^-\rangle a \to (\psi \to G(\langle \mathrm{R}_3^-\rangle a \to \chi))))\end{array}}{\downarrow a.H(\langle \mathrm{R}_0^-\rangle a \to G(\langle \mathrm{R}_1^-\rangle a \to (\varphi \to G(\langle \mathrm{R}_3^-\rangle a \to \chi))))}\ (SCT)$$

$$\frac{\begin{array}{c}\downarrow a.G(\langle \mathrm{R}_1^-\rangle a \to (\varphi \to G(\langle \mathrm{R}_2^-\rangle a \to \psi)))\\ \downarrow a.G(\langle \mathrm{R}_2^-\rangle a \to (\psi \to G(\langle \mathrm{R}_3^-\rangle a \to \chi)))\end{array}}{\downarrow a.G(\langle \mathrm{R}_1^-\rangle a \to (\varphi \to G(\langle \mathrm{R}_3^-\rangle a \to \chi)))}\ (SIT)$$

Obviously, (SIT) is a special case of (SCT), where 'the head' of the three formulas constituting the latter, i.e. '$H(\langle \mathrm{R}_0^-\rangle a \to$', is removed.[53] This implies that there are many uses of conditionals in which the *indicative* transitivity

$$\frac{\begin{array}{c}\textit{If it is the case that } \varphi\textit{, then it will be the case that } \psi.\\ \textit{If it is the case that } \psi\textit{, then it will be the case that } \chi.\end{array}}{\textit{If it is the case that } \varphi\textit{, then it will be the case that } \chi.}\ (IT)$$

is just a special case of the *counterfactual* transitivity

$$\frac{\begin{array}{c}\textit{If it were the case that } \varphi\textit{, then it would be the case that } \psi.\\ \textit{If it were the case that } \psi\textit{, then it would be the case that } \chi.\end{array}}{\textit{If it were the case that } \varphi\textit{, then it would be the case that } \chi.}\ (CT)$$

This, in turn, implies that there are many uses of conditionals in which the *indicative* conditional

$$\text{`}\textit{If it is the case that } \varphi,\ \textit{then it will be the case that } \psi\text{'}$$

is just a special case of the *counterfactual* conditional

$$\text{`}\textit{If it were the case that } \varphi,\ \textit{then it would be the case that } \psi\text{'}.$$

Thus, we discover a *new* class of conditionals in which an *indicative* conditional is a special case of a *counterfactual* conditional in its *logical* form. We refer to such a class as *'temporal'* conditionals. A distinctive characteristic of the class is given roughly as follows:

(TC)[54] For any member of the class of *'temporal'* conditionals, if it is *grammatically counterfactual*, then we can construct *its intelligible indicative counterpart* by

[53] Semantically speaking, this means that (SIT) does not go back to a diverging time point R_0 in the past.

[54] (TC) is the abbreviation for 'temporal conditionals'.

cutting 'the head' $H(\langle \mathrm{R}_0^- \rangle a \rightarrow$ of (the weak version of) its translation in HTL_{TC}, and conversely, if it is *grammatically indicative*, then we can construct *its intelligible counterfactual counterpart* by adding 'the head' $H(\langle \mathrm{R}_0^- \rangle a \rightarrow$ to (the weak version of) its translation in HTL_{TC}.

(Nix) is a paradigm of this class. We can construct its *indicative* counterpart *'If Nixon presses the button, there will be a nuclear war'* in the specified way. Similarly, this can be done for (Osw) in Sect. 3 and the three constituents of (OAW) in Sect. 4, although there are cases in which (OAW) does not hold as a whole.[55]

Meanwhile, consider (b), (y), (k), and (g) in Sect. 3. The *indicative* counterparts of (b) and (y) are complete nonsense. For (k), its reasonable counterpart seems to be *'If kangaroos lose tails, they will topple over'*; however, this is at least quite unnatural. For (g), its counterpart would be *'If gravity goes by the inverse cube of distance,...'*; however, this is again almost nonsense.

Interestingly, for this criterion, what we want to refer to as a *historical* counterfactual may nevertheless turn out to be *not 'temporal'*. Consider *'No Hitler, no A-bomb'*$\cdots$ (Hit).[56] Its *indicative* counterpart would be something such as *'If Hitler does not appear, A-bombs will not be developed'*; however, it has unnaturalness similar to that of *'If kangaroos lose tails,...'*. This suggests that the former is also close to a rhetoric that highlights the historical importance of Hitler for the actual development of atomic bombs.

In Table 2, we display the result of applying (TC) to the samples of counterfactuals in this paper:

Certainly, there will be considerable scope to scrutinise (TC) above.[57] However, I would like the reader to at least accept the following fact: there are many conditionals for which our symbolisation in HTL_{TC} is appropriate. If this fact is accepted, there seems to emerge *a new logical perspective* on conditionals in general: *temporal ones and others*.

More specifically, the above results show that, among *both* indicative *and* counterfactual conditionals, there are those that *can* be symbolised appropriately in a natural extension of classical tense logic, and there are those that *cannot*. In this case, we find out that the former become much clearer *from a logical point of view*, whereas the latter are less clear and miscellaneous; hence, *further logical investigation* is required, which, in turn, perhaps requires *different formal systems* in which *they can be reconstructed otherwise logically*.

---

55 In HTL_{TC}, the failure of (OAW) can be explained as the failure of $(R_1^- <^{\exists} R_2^- < R_3^-)$, i.e. the failure of *intermediateness* of R_2 between R_1 and R_3.

56 Cited from Lewis (1973, p. 4).

57 Among other things, we have not scrutinised (TC) for *indicatives* yet. However, consider *'If Oswald did not kill Kennedy, then someone else did'*$\cdots$ (Osw-I). Somewhat unexpectedly, we can construct its approximation in HTL_{TC} as follows, where p:=*'Oswald kills Kennedy'* and q:=*'someone else kills Kennedy'*: $\downarrow a.H(\langle \mathrm{R}_1^- \rangle a \rightarrow (p \rightarrow q))\cdots$(Osw-I'). If the reader cannot accept (Osw-I') as a symbolisation of (Osw-I) for some reason, then our purpose is fulfilled; we may exclude (Osw-I) from the class of *'temporal'* conditionals for that reason from the outset. If we tolerate (Osw-I') and 'counterfactualise' it in the specified way, then we obtain $\downarrow a.H(\langle \mathrm{R}_0^- \rangle a \rightarrow H(\langle \mathrm{R}_1^- \rangle a \rightarrow (p \rightarrow q)))\cdots$(Osw-C?). However, under the implicit presupposition that R_0 should precede R_1, an interesting result is obtained: (Osw-I') vacuously implies (Osw-C?), and conversely, (Osw-C?) implies (Osw-I') under the presupposition; hence, they become equivalent. Thus, our 'counterfactualising' (Osw-I) *idles*, and therefore it *cannot be 'temporal'*.

Table 2 Classification of counterfactual samples in this study

Temporal		
Historical/real counterfactuals	Label	Page
If Nixon had pressed the button, there would have been a nuclear war	(Nix)	2, 7, 27
If Oswald had not killed Kennedy, then someone else would have	(Osw)	7, 27
If Otto had gone to the party, then Anna would have gone	in (OAW)	8
If Max missed the train, he would have taken the bus	(Max)	18
Not temporal		
Rhetorical counterfactuals	Label	Page
If I were a bird, I could fly there!	(b)	7, 27
If I were you, I would retort!	(y)	7, 27
If kangaroos had no tails, they would topple over	(k)	7, 27
No Hitler, no A-bomb	(Hit)	27
Theoretical counterfactuals	label	page
If gravity went by the inverse cube of distance,..	(g)	8, 27
If this axiom in this system were replaced by another,..	(ax)	8
If we modified the value of the parameter in this system,..	(para)	8

8 Related Works and Discussion

Our results would have many ramifications for related works in philosophical logic and linguistics. We list some in this section.

Two previous works, Thomason and Gupta (1980) and Placek and Müller (2007), analysed counterfactuals in terms of historical/real possibility modelled by branching-time structures. One of the differences in semantics between these approaches and ours is that while Thomason and Gupta (1980) and Placek and Müller (2007) involve quantification over histories, ours does not. Among the differences in syntax, these approaches use a special single connective (Stalnaker-style $>$ for Thomason and Gupta (1980), Lewis-style $\Box\!\!\rightarrow$ for Placek and Müller (2007)), whereas ours does not. (For these differences, see Sect. 5.4.)

As a formal system, our approach, i.e. HTL_{TC}, is close to M. J. Cresswell's system (Cresswell, 2010), which enables temporal references corresponding to 'now' and 'then' in natural language. However, there are two main differences between his system and ours. First, and less important, Cresswell's logic is a type of *predicate* tense logic, whereas ours is a type of *hybrid* tense logic.[58] Second, and more importantly, Cresswell's logic is a *linear-time* tense logic, whereas ours is a (possibly) *branching-*

[58] This means that our system stays *propositional on the surface* thanks to *nominals*. However, this is less important, as hybrid logic enjoys as much expressive power as (a fragment of) first-order logic. For the details on expressivity of hybrid logic, see Ten Cate (2005).

time tense logic, which enables us to express *many possible timelines upward* and also make temporal references *in them*.[59]

Among recent literature, W. B. Starr's approach (Starr, 2014) seems antithetical to ours, as it proposes 'a uniform theory of conditionals' (its title), which is based on an extension of Stalnaker's system and, more fundamentally, what we referr to as Stalnaker reasoning: there is *a uniformity of grammatical form* among *conditionals*; *therefore*, there should be *a uniformity of logical form* among *them*.[60] However, this certainly does not mean that our results conflict with Starr's theory; instead, it only suggests that the targets of Starr's theory include *temporal* and *other* conditionals. Therefore, it would be expected that Starr's theory would turn out to be more suitable for the latter (other) class than for the former (temporal) class.

Meanwhile, Rothschild (2015) maintains propositionalism about conditionals, which is the conservative view that conditionals, as well as non-conditional declarative sentences, have propositional content identified with their truth conditions that can adequately account for their meaning; however, as Rothschild (2015) reports, there are more serious challenges to this view concerning conditionals than there are concerning non-conditionals. Our analysis could be a *partial* but *strong* support for the propositionalism about conditionals: it could be a *partial* support, as it is specific to *temporal* conditionals; it could, on the other hand, be a *strong* support because of its specificity, as it has presented their fine-grained syntactic structures and correspondingly their fine-grained truth conditions. Thus, it would be worthwhile to scrutinise how far the specified truth conditions of *temporal* conditionals can vindicate the propositionalism about *temporal* conditionals.

Finally, there seems to be an overall characteristic of our approach that is distinct from any other in the literature on counterfactuals. As we mentioned in Sect. 4, since Lewis (1973), it has been agreed that counterfactual transitive inference (CT) does not necessarily hold. However, this is just another way of saying that there are cases in which (CT) *does not* hold *as well as* cases in which it *does* hold. Previous studies appear to have focussed mainly on the former *fallacious* cases; however, this paper focusses on the latter *successful* cases and asked the question *why does* (CT) *hold when it does hold?*

9 Concluding Remarks

In conclusion, we present a few more remarks about a crucial difference between our new analysis and Lewis's classical analysis, in the light of (CT). In this paper, we faithfully follow the 'meaning-as-use' and 'inferentialist' approach involved in the procedure of Tarski's theory of truth,[61] whereas Lewis appears to have had no intention of doing so. We specify what constrains the *use* of counterfactuals in *inference*, i.e.

[59] To the best of our knowledge, the most interesting advantage of Cresswell's system over our system is that Cresswell's system reconstructs the mysterious notion 'now' only from tense logical and first-order machinery.

[60] For Stalnaker reasoning, see Sect. 3 and footnote 15.

[61] For the 'meaning-as-use' and 'inferentialist' approach involved in the procedure of Tarski's theory of truth, see Sect. 2.

(successful) counterfactual transitive inference (CT), and, *as a result*, there emerges *a set of logical constraints* on the notion of ‘temporality’; Lewis, on the other hand, does not seem to have adopted such a procedure concerning the notion of ‘similarity’.

This difference is reflected clearly in respective formal deductions with regard to (CT). In our deduction of (SCT), there appear inference rules that are certain to constrain the notion of ‘temporality’, namely the elimination/introduction rules of tense operators, $(Trans\text{-}F)$, and $(R_1^- <^\exists R_2^- < R_3^-)$. Meanwhile, for Lewis's deduction of (VCT),[62] *no* such rules that constrain ‘similarity’ seem to be required. Other than rules and axioms of propositional logic, the only rule and axiom required there are (DWC) and $(VA(5))$:

Deduction within ConditionalsFor any $n \geq 1$,

$$\frac{(\chi_1 \ \& \ \ldots \ \& \ \chi_n) \rightarrow \psi}{((\varphi \Box\!\!\rightarrow \chi_1) \ \& \ \ldots \ \& \ (\varphi \Box\!\!\rightarrow \chi_n)) \rightarrow (\varphi \Box\!\!\rightarrow \psi)} \ (DWC)$$

Axiom (5)

$$\vdash (\varphi \Box\!\!\rightarrow \neg\psi) \vee (((\varphi \ \& \ \psi) \Box\!\!\rightarrow \chi) \leftrightarrow (\varphi \Box\!\!\rightarrow (\psi \rightarrow \chi))) \quad \cdots (VA(5))$$

However, these do not seem to have any direct connection with ‘similarity’. Roughly speaking, (DWC) is a type of necessitation of actual implication,[63] and $(VA(5))$ is a weakened analogy of the theorem $((\varphi \wedge \psi) \rightarrow \chi) \leftrightarrow (\varphi \rightarrow (\psi \rightarrow \chi))$ in propositional logic. After all, at least proof-theoretically, it seems that the notion of ‘similarity’ is not required for deduction of (VCT).[64]

Thus, the alleged success of our *new* analysis can be attributed to the two *classical* points of view, namely the ‘meaning-as-use’ and ‘inferentialist’ points of view, detected in Tarski's theory of truth in Sect. 2. We believe that it is also a demonstration of their significance for *logical* analysis of natural language.[65]

Funding Funding was provided by JSPS KAKENHI (Grant No. 20K12784). I would like to take this opportunity to thank Kengo Okamoto for years of advice. I would also like to thank Katsuhiko Sano for discussion and suggestions on the early stages of this work. Finally, I am grateful to two anonymous reviewers for helpful and constructive comments. This work was supported by JSPS KAKENHI Grant Number 20K12784.

[62] A Hilbert-style deduction of (VCT) in the most basic version V of Lewis's axiomatic system is displayed in Hosokawa (2012, pp. 21–22). For the system V, see Lewis (1973, p. 132).

[63] With a little care, we observe that (DWC) is nonstandard as a modal rule. Consider the case $n = 1$ for simplicity, and we have $\chi \rightarrow \psi \vdash (\varphi \Box\!\!\rightarrow \chi) \rightarrow (\varphi \Box\!\!\rightarrow \psi)$. Its counterpart in the form of a *strict conditional* is $\chi \rightarrow \psi \vdash \Box(\varphi \rightarrow \chi) \rightarrow \Box(\varphi \rightarrow \psi)$. However, the latter is not valid in normal modal logics. A simple way to make the latter valid, say, in **K**, is to add to the premise $(\chi \rightarrow \psi) \rightarrow \Box(\chi \rightarrow \psi)$, which means that actual implication can be necessitated. From this point of view, we can say that (DWC) involves a type of necessitation of actual implication. Meanwhile, deduction of (SCT) in HTL_{TC} does not involve such a nonstandard rule at all. See Sect. 5.5.

[64] In fact, it seems to be quite difficult to specify inferences that constrain the notion of ‘similarity’, especially in our miscellaneous ordinary discourses involving the ‘vague’ notion, as Lewis himself frequently refers to this notion in that manner and the item ‘vagueness of similarity’ is listed in the index in his own book. See Lewis (1973, p. 149).

[65] I owe my realisation of the significance of these points of view involved in Tarski's theory of truth to Okamoto (1999).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A: A Natural Deduction System for $\mathbf{HTL}_{TC}$

A natural deduction system for HTL_{TC}, which is only a special case of Braüner's system (Braüner, 2011), is presented below (Tables 3, 4, 5).

Table 3 Natural deduction rules for connectives

$$\frac{@_a\varphi \qquad @_a\psi}{@_a(\varphi\wedge\psi)}\,(\wedge I) \qquad \frac{@_a(\varphi\wedge\psi)}{@_a\varphi}\,(\wedge E1) \qquad \frac{@_a(\varphi\wedge\psi)}{@_a\psi}\,(\wedge E2)$$

$$\frac{\begin{array}{c}[@_a\varphi]\\ \vdots\\ @_a\psi\end{array}}{@_a(\varphi\to\psi)}\,(\to I) \qquad \frac{@_a\varphi \qquad @_a(\varphi\to\psi)}{@_a\psi}\,(\to E)$$

$$\frac{\begin{array}{c}[@_a\neg\varphi]\\ \vdots\\ @_a\bot\end{array}}{@_a\varphi}\,(\bot 1)^{*} \qquad \frac{@_a\bot}{@_c\bot}\,(\bot 2)$$

$$\frac{@_a\varphi}{@_c@_a\varphi}\,(@I) \qquad \frac{@_c@_a\varphi}{@_a\varphi}\,(@E)$$

$$\frac{\begin{array}{c}[@_aFc]\\ \vdots\\ @_c\varphi\end{array}}{@_aG\varphi}\,(GI)^{\star} \qquad \frac{@_aG\varphi \qquad @_aFe}{@_e\varphi}\,(GE) \qquad \frac{\begin{array}{c}[@_aPc]\\ \vdots\\ @_c\varphi\end{array}}{@_aH\varphi}\,(HI)^{\dagger} \qquad \frac{@_aH\varphi \qquad @_aPe}{@_e\varphi}\,(HE)$$

$$\frac{\begin{array}{c}[@_a\langle \mathrm{R}_n\rangle c]\\ \vdots\\ @_c\varphi\end{array}}{@_a[\mathrm{R}_n]\varphi}\,([\mathrm{R}_n]I) \qquad \frac{@_a[\mathrm{R}_n]\varphi \qquad @_a\langle \mathrm{R}_n\rangle e}{@_e\varphi}\,([\mathrm{R}_n]E)$$

$$\frac{\begin{array}{c}[@_ac]\\ \vdots\\ @_c\varphi[c/b]\end{array}}{@_a\downarrow b.\varphi}\,(\downarrow I)^{\ddagger} \qquad \frac{@_a\downarrow b.\varphi \qquad @_ae}{@_e\varphi[e/b]}\,(\downarrow E)$$

$*$ φ is a propositional symbol (ordinary or nominal)
$\star$ c does not occur free in $@_aG\varphi$ or in any undischarged assumptions other than the specified occurrences of $@_aFc$
† c does not occur free in $@_aH\varphi$ or in any undischarged assumptions other than the specified occurrences of $@_aPc$.
‡ c does not occur free in $@_a\downarrow b.\varphi$ or in any undischarged assumptions other than the specified occurrences of $@_ac$

Table 4 Natural deduction rules for nominals

$\dfrac{}{@_a a}\ (Ref\text{-}@)$	$\dfrac{@_a c \qquad @_a \varphi}{@_c \varphi}\ (Nom1)^*$	$\dfrac{@_a c \qquad @_a Mb}{@_c Mb}\ (Nom2)^\star$

$*$ φ is a propositional variable (ordinary or nominal)
$\star$ $M \in \{F, P, \langle R_n \rangle, \langle R_n^- \rangle\}$

Table 5 Rules corresponding to conditions on the accessibility relation

$$\dfrac{@_a Mb \qquad @_b Mc}{@_a Mc}\ (Trans\text{-}M)^*$$

$$\dfrac{@_a Fc}{@_c Pa}\ (Sym\text{-}FP) \qquad \dfrac{@_a Pc}{@_c Fa}\ (Sym\text{-}PF)$$

$$\dfrac{@_a \langle R_n \rangle c}{@_c \langle R_n^- \rangle a}\ (Sym\text{-}R_n^{+-}) \qquad \dfrac{@_a \langle R_n^- \rangle c}{@_c \langle R_n \rangle a}\ (Sym\text{-}R_n^{-+})$$

$$\dfrac{@_a Pc \qquad @_a Pd \qquad \begin{array}{c}[@_c Pd]\\ \vdots \\ \varphi\end{array} \qquad \begin{array}{c}[@_c d]\\ \vdots \\ \varphi\end{array} \qquad \begin{array}{c}[@_d Pc]\\ \vdots \\ \varphi\end{array}}{\varphi}\ (Lin\text{-}P)$$

$$\dfrac{@_c \langle R_n^- \rangle a \qquad @_c Fe \qquad @_e \langle R_{n+2}^- \rangle a \qquad \begin{array}{c}[@_c Fd] \quad [@_d \langle R_{n+1}^- \rangle a] \quad [@_d Fe]\\ \vdots \\ \varphi\end{array}}{\varphi}\ (R_n^- <^\exists R_{n+1}^- < R_{n+2}^-)^\star$$

$*$ $M \in \{F, P\}$
$\star$ d does not occur free in φ or in any undischarged assumptions other than the specified occurrences of $@_c Fd$, $@_d \langle R_2^- \rangle a$ and $@_d Fe$

References

Belnap, N. D. (1962). Tonk, plonk and plink. *Analysis, 22*(6), 130–134.

Blackburn, P. (1990). *Nominal tense logic and other sorted intensional frameworks*. PhD thesis, The University of Edinburgh.

Blackburn, P. (1994). Tense, temporal reference, and tense logic. *Journal of Semantics, 11*(1–2), 83–01. https://doi.org/10.1093/jos/11.1-2.83

Blackburn, P., & Jørgensen, K. F. (2016). Reichenbach, prior and hybrid tense logic. *Synthese, 193*, 3677–3689. https://doi.org/10.1007/s11229-015-0920-0

Brandom, R. B. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard University Press.

Braüner, T. (2011). *Hybrid logic and its proof-theory*. Springer.

Cresswell, M. J. (2010). Temporal reference in linear tense logic. *Journal of Philosophical Logic, 39*(2), 173–200.

Fine, K. (1975). Critical notice of counterfactuals. *Mind, 84*, 451–458.

Frege, G. (1879a). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle an der Saale: Verlag von Louis Nebert.

Frege, G. (1979b). Boole's logical calculus and the concept-script. In H. Hermes, F. Kambartel, & F. Kaulback (Eds.), *Posthumous writing* (pp. 9–46). Oxford University Press.

Geach, P. T. (1976). *Reason and argument*. University of California Press.

Geach, P. T. (1980). *Reference and generality*. Cornell University Press.

Goranko, V. (1994). Temporal logic with reference pointers. In Gabbay, D., & Ohlbach, H. (Eds). *Proceedings of the 1st international conference on temporal logic, lecture notes in artificial intelligence* (Vol. 827, pp. 133–148).
Goranko, V. (1996). Hierarchies of modal and temporal logics with reference pointers. *Journal of Logic, Language, and Information, 5*, 1–24.
Hosokawa, Y. (2012). Formalization of counterfactual inference by dynamic logic. *The Philosophy of Science Society, Japan, 45*(1), 17–33.
Lewis, D. (1973). *Counterfactuals*. Oxford Basil Blackwell.
Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs, 13*, 455–476. https://doi.org/10.2307/2215339
Lewis, D. (1986). *On the plurality of worlds*. Oxford Basil Blackwell.
Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy, 74*, 549–567. https://doi.org/10.1080/00048409612347521
Okamoto, K. (1999). Early Frege on the language of arithmetic and the concept-script. *Historia Philosophiae, 41*, 20–37.
Pearl, J. (1999). Probabilities of causation: Three counterfactual interpretations and their identification. *Synthese, 121*, 49–93. https://doi.org/10.1145/3501714.3501735
Placek, T., & Müller, T. (2007). Counterfactuals and historical possibility. *Synthese, 154*(2), 173–97.
Prior, A. N. (1955). Diodoran modalities. *Philosophical Quarterly, 5*, 205–213. https://doi.org/10.2307/2957434
Prior, A. N. (1967). *Past, present and future*. Oxford University Press.
Reichenbach, H. (1947). *Elements of symbolic logic*. The Macmillan Company.
Rothschild, D. (2015). Conditionals and propositions in semantics. *Journal of Philosophical Logic, 44*(6), 781–791.
Stalnaker, R. (1968). A theory of conditionals. Reprinted in Jackson F. (Ed). *Conditionals*, 28–45.
Stalnaker, R. (1975). Indicative conditionals. *Philosophia, 5*, 269–286.
Starr, W. B. (2014). A uniform theory of conditionals. *Journal of Philosophical Logic, 43*(6), 1019–1064.
Tarski, A. (1944). The semantical concept of truth and the foundations of semantics. *Philosophy and Phenomenological Research, 4*, 341–376. https://doi.org/10.2307/2102968
Ten Cate, B. (2005). *Model theory for extended modal languages*. PhD thesis, ILLC Dissertation Series DS-2005-01.
Thomason, R., & Gupta, A. (1980). A theory of conditionals in the context of branching time. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 299–322). Dordrecht: Reidel.
Tsai, C.-J. (2014). *Japanese counterfactual conditionals*. PhD thesis, Soochow University.
Wittgenstein, L. (2001) [1953]. *Philosophical investigations*. Blackwell Publishing.
Xu, M. (1997). Causation in branching time (I): Transitions, events and causes. *Synthese, 112*, 137–92. https://doi.org/10.1023/A:1004948726086

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.