

McTaggart's Argument for the Unreality of Time: *A Temporal Logical Analysis**

Hunan Rostomyan

In this paper I want to examine McTaggart's [1908] argument for the unreality of time. McTaggart starts with the observation that we distinguish events in time in one of two ways. The first way, which he calls the 'B-Series', involves the ordering of events as *earlier than*, *simultaneous with*, or *later than* other events. The second way, which he calls the 'A-Series', involves the ordering of events as *past*, *present*, or *future* with respect to a point of reference. Based on this distinction McTaggart gives a two-part argument to the conclusion that time is unreal. First he argues that time requires the A-Series. Then he tries to show that the A-Series leads to a contradiction, effectively reducing time to absurdity.

Before we look at the argument in detail, in the rest of this introduction I want to settle the question of what McTaggart's conclusion even means. McTaggart's text doesn't determine a particular answer to this question. Consequently, *we* have to specify what exactly it means to say that time is unreal. To meet this interpretive challenge we should start by specifying the *nature* of the B/A-Series distinction. Should we understand it as a *phenomenological* distinction about two ways of perceiving or experiencing events in time? Or perhaps as a *scientific* distinction about two physico-mathematical conceptions of events in time? While these and many other interpretations might lead to very fruitful analyses of the argument, in this paper I take the distinction to be one between two ways of *talking* about events in time, between two temporal *languages*. The argument then, I take to be a threat to the coherence of our talk of time. Although our focus will be on languages as opposed to say perceptions or cognitions, our approach will be not linguistic, but *logical*. This means that we will make significant use of temporal logical machinery both in the exposition of the argument and when deciding the truth of various claims.¹

*I would like to thank Justin Bledin, Jeff Kaplan, and Matthew Ramirez for very helpful comments on earlier drafts of this paper. I would also like to extend my gratitude to Professors Wesley Holliday and John MacFarlane for showing me the many ways of modal logic.

¹We will only talk of languages and models, so basic familiarity with semantics is all that is presupposed.

§1 Preliminary Explanations

We start with McTaggart’s informal characterization of the distinction between the ‘B-Series’ and the ‘A-Series’ (§1.1). We then explicate these notions in terms of two different *languages*, and give their semantics with respect to transitive, irreflexive flows of time which we’ll simply call *temporal models* (§1.2). We continue by noting McTaggart’s presupposition that the two languages *exhaust* the ways in which we can talk about events as taking place in time (§1.3).

§1.1 Explicanda: B-Series & A-Series

McTaggart’s (hereafter McT) argument for the unreality of time is centered around a distinction between two ways of ordering positions in time or ‘moments’. The first, which he calls ‘B Series,’ is the ordering of moments according to the three binary relations: *earlier than*, *later than*, and *simultaneous with*. Suppose we have three events: Washington’s becoming the president (w), Lincoln’s Becoming the president (l), and Obama’s becoming the president (o). We can order these events in a B-Series, obtaining the ordered sequence: w, l, o . The series thus obtained has two interesting characteristics. First, its determination is fixed, i.e., for any two moments m and m' , if m ever precedes m' in such a series, then m always has and forever will precede m' . E.g., Lincoln’s becoming a president (l) occurs at a moment before the occurrence of Obama’s becoming the president (o), so even if we expand the series by including in it more and more moments, the fact will remain that l precedes o . The second characteristic of this series is that it is what McT calls ‘absolute,’ and by which he means that when evaluating a claim like m precedes m' or m is simultaneous with m' , etc., it doesn’t matter at which point in time we do so. E.g., Lincoln’s becoming the president precedes Obama’s, whether we evaluate that claim at 5 BC, now, or in the future.

But McT recognizes that we also talk of events as being present, past, or future. To allow us to make such reference-time-dependent claims, McT introduces the ‘A Series,’ which is the ordering of moments according to three unary predicates: *past*, *present*, and *future*. We can order the aforementioned events, with respect to the reference time now, in an A-Series (by years) as follows: $Past_{224}(w)$, $Past_{153}(l)$, $Past_4(o)$. The meaning of this is that Washington’s presidency, for example, is Past, and the subscript indicates what is implicit, namely that it is now 224 years past (years because we chose that as our unit of measure). As the years (or days or whatever unit we choose) progress, these subscripts are either incremented (for Past events) or decremented (for Future events). E.g., If our reference time was five years ago, then Obama’s becoming the president would be $Future_1(o)$. This reference-sensitive nature of the A-Series will play a crucial role in McT’s argument in §2.1.

§1.2 Explicata: \mathcal{L}_B & \mathcal{L}_A

These being the informal characterizations of the two series, we will now attempt to give them precise temporal logical explications. We start by defining languages \mathcal{L}_B and \mathcal{L}_A for the two series, respectively. Then we give the semantics for each language. When we turn to McT's argument in §2, we will use the definitions that follow in deciding various claims.

Definition 1.1 (LANGUAGE \mathcal{L}_B) *Given a set of event letters $\mathcal{E} = \{e_1, e_2, \dots\}$, and a vocabulary $V_B = \mathcal{E} \cup \{\neg, \wedge, \triangleleft, =, \}$ we define the B-Series language \mathcal{L}_B as $(\forall e \in \mathcal{E})$:*

$$\phi ::= (e \triangleleft e) \mid (e = e) \mid \neg\phi \mid (\phi \wedge \phi)$$

Where the intended interpretation of $(e \triangleleft e')$ is that (e precedes e'), and of $(e = e')$ is that (e is simultaneous with e'). We can complete the full range of B-Series predicates by defining the relation of succession $(e \triangleright e')$ in terms of the precedence relation as $(e' \triangleleft e)$, with the intended interpretation of $(e \triangleright e')$ as (e succeeds e'). Since negation and conjunction are truth-functionally complete, we will simply assume the rest of the vocabulary V_B to have been given their obvious definitions in terms of those two.²

Definition 1.2 (SEMANTICS OF \mathcal{L}_B) *We define B-Series models as triples $\mathcal{T} = \langle T, <, V \rangle$ with a valuation function $V : \mathcal{E} \rightarrow \mathcal{P}(T)$ and an accessibility relation ($<$) that is transitive and irreflexive. We define truth in a B-Series model, at a point, recursively as follows:*

$$\begin{aligned} \langle \mathcal{T}, t \rangle \models (e \triangleleft e') & \quad \text{iff} \quad \exists s, s' : s < s', s \in V(e) \text{ and } s' \in V(e') \\ \langle \mathcal{T}, t \rangle \models (e = e') & \quad \text{iff} \quad \exists s, s' : s \in V(e) \text{ and } s \in V(e') \end{aligned}$$

The Boolean clauses are set up in the usual way.³ Notice that we could have given stronger truth-conditions for precedence and simultaneity, quantifying universally over time points instead.⁴ But for the purposes of this essay, what matters is that the truth-conditions do not depend on the point t of reference.

Definition 1.3 (LANGUAGE \mathcal{L}_A) *Given a set of event letters $\mathcal{E} = \{e_1, e_2, \dots\}$, and a vocabulary $V_A = \mathcal{E} \cup \{\neg, \wedge, H, G\}$ we define the A-Series language \mathcal{L}_A as $(\forall e \in \mathcal{E})$:*

$$\phi ::= e \mid \neg\phi \mid \phi \wedge \phi \mid G\phi \mid H\phi$$

²E.g., $(\phi \rightarrow \psi)$ in terms of $(\neg\phi \vee \psi)$, and so on.

³I.e., $\langle \mathcal{T}, t \rangle \models \neg\phi$ iff $\langle \mathcal{T}, t \rangle \not\models \phi$, $\langle \mathcal{T}, t \rangle \models (\phi \wedge \psi)$ iff $\langle \mathcal{T}, t \rangle \models \phi$ and $\langle \mathcal{T}, t \rangle \models \psi$.

⁴Professor Holliday introduced a worry that I haven't been able to resolve. On one hand, I want to say that my life has been simultaneous with my grandfather's (for some point we've both been alive at that point). On the other hand, I want to say that my grandfather's life has been simultaneous with my grandmother's in the stronger sense of at all points them both being alive. Which is the more appropriate explication of simultaneity I haven't been able to decide.

$G\phi$ ('it will always be the case that ϕ ') and $H\phi$ ('it has always been the case that ϕ ') allow us to quantify over all future and past moments, respectively. On the basis of these we can define $P\phi$ ('it was sometime the case that ϕ ') and $F\phi$ ('it will sometime be the case that ϕ ') as $(\neg H\neg\phi)$ and $(\neg G\neg\phi)$, respectively.⁵

Definition 1.4 (SEMANTICS OF \mathcal{L}_A) *We define A-Series models exactly as we defined B-Series models above. We define truth in an A-Series model, at a point, as follows:*

$$\begin{aligned} \langle \mathcal{T}, t \rangle \models e & \quad \text{iff} \quad t \in V(e) \\ \langle \mathcal{T}, t \rangle \models G\phi & \quad \text{iff} \quad \forall t' : t < t', \langle \mathcal{T}, t' \rangle \models \phi \\ \langle \mathcal{T}, t \rangle \models H\phi & \quad \text{iff} \quad \forall t' : t' < t, \langle \mathcal{T}, t' \rangle \models \phi \end{aligned}$$

Again, the Boolean clauses are what you would expect. Since A-Series are B-Series models are the same, in the rest of the paper we'll simply refer to them as our *temporal models*.

§1.3 Exhaustiveness Presupposition

It's important to note at this point a presupposition of McT's, namely that, the two ways of speaking about events in time (B-Series, A-Series), and therefore the two languages (\mathcal{L}_B , \mathcal{L}_A) are *exhaustive*. He does consider briefly what he calls the 'C-Series,' but he dismisses it because the 'C-Series' orders events in time, but doesn't give them directionality. Consider the set of events $\{a, b, c\}$. The B-Series could say that (a precedes b), (b precedes c), giving us the series: a, b, c. The A-Series could say that (when a is present b is future), (when b is present c is future), giving us the same series: a, b, c. The C-Series would say that (a R b), (b R c), for a transitive, irreflexive R, leaving it ambiguous the direction in which the series should be read. Does it look like (a, b, c) or (c, b, a)? The C-Series doesn't tell us. McT dismisses the C-Series, and does indeed think that the B-Series and A-Series exhaust the space of possibilities. But should this presupposition be granted? In the case of the two *series*, it's hard to tell whether there could be another one, say some 'D-Series.' But in the case of the two *languages*, it's easy to see that there could be and indeed is another one: a language, say \mathcal{L}_K , that takes Hans Kamp's binary *since* and *until* as its primitive temporal operators. This language is able to simulate both \mathcal{L}_B and \mathcal{L}_A .⁶ As far as the argument against the reality of time is concerned, one might think that we could stop here, thinking that even if McT indeed does succeed in reducing the A-Series to

⁵The rest of V_A are again given their obvious definitions in terms of negation and conjunction.

⁶First, \mathcal{L}_A can simulate \mathcal{L}_B : define ' $a \triangleleft b$ ' as $\diamond(a \wedge Fb)$ where $\diamond\phi \equiv (P\phi \vee \phi \vee F\phi)$, and ' $a = b$ ' as $\diamond(a \wedge b)$. Second, \mathcal{L}_K can simulate \mathcal{L}_A : define ' $P\phi$ ' as $S\phi\top$, and ' $F\phi$ ' as $U\phi\top$, where S and U are Kamp's *since* and *until* operators. Therefore, \mathcal{L}_K can simulate both languages. For the details see Kamp [1971].

absurdity, it would no longer imply that time itself doesn't exist because now there is this other language that McT needs to show to be incoherent. But this might be too quick a verdict, for if the A-Series is indeed incoherent, since \mathcal{L}_K simulates \mathcal{L}_A , this might threaten the coherence of the Kamp-Series as well. So we seem to have two options: either we need to find a language that's acceptable as way of speaking about events in time *and* does not simulate \mathcal{L}_A , or we grant the presupposition that the A/B-Series are exhaustive and continue with McT's argument. We will take the second route, examining McT's argument in the following section.

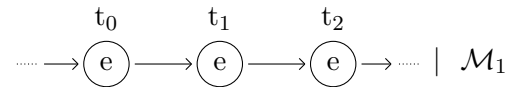
§2 The Argument

In this section we examine McTaggart's two-part argument. The first part of his argument tries to establish that (I) time requires the 'A-Series' (§2.1). He supports this claim with two premises from which it follows: (P1) time requires change, and (P2) change requires the 'A-Series.' The second part of his argument then tries to show that (II) the 'A-Series' leads to a contradiction, effectively reducing time to absurdity (§2.2). He supports this claim with two premises from which it follows: (P3) past, present, and future cannot simultaneously hold of an event, but (P4) the 'A-Series' leads to every event being both past, present, *and* future.

§2.1 Time \rightarrow A-Series

McT begins the first part of his argument with the observation that 'time involves change' (459). He motivates this by saying that 'a universe in which nothing whatever changed ... would be a timeless universe' (Ibid). McT then concludes that (P1): time *requires* change. Thus, beginning with something obvious, he ends with something controversial.⁷ What do we think about this? Are there any good reasons to grant McT this premise? The answer depends on the definition of *change* that one has in mind. Consider the model \mathcal{M}_1 below, where $V(e) = T$, and $V(e') = \emptyset$ for all event letters $e' \neq e$.

This is a model in which event e is taking place at every moment in time, and no other event is taking place any-



where. Can we *describe* such a model as changing? Well, that depends on what *language* we want this description to be made in. Now, previously we distinguished between two languages for describing time: the language \mathcal{L}_B of the B-Series and the language \mathcal{L}_A of the A-Series. Can we describe this model as changing using either of these languages? The answer, for this particular model, is a negative for both languages. Which makes sense, given

⁷The view that time requires change has found a number of able supports, e.g. Aristotle, but it is not immediately obvious why it's impossible to give an explication of time that doesn't require change. In fact, one might be more inclined to think of change in terms of time or to think of time *as* change itself. For a persuasive account of change-less time see Shoemaker [1969].

that this model intuitively doesn't seem to exhibit change. We will need a few more definitions to show this formally. After giving the definitions and demonstrating this trivial claim for this particular model, we will state a less trivial, general result about the relationship between our models, the two languages, and change.

Definition 2.1 (\mathcal{L} -DESCRIPTION) *Let the \mathcal{L} -description $Desc_{\mathcal{L}}(\mathcal{T}, t)$ of a point t in a model \mathcal{T} be the set of all formulas $\phi \in \mathcal{L}$ s.t. $\langle \mathcal{T}, t \rangle \models \phi$.*

Although inspired by Carnap's notion of a state-description, \mathcal{L} -descriptions don't have the property of uniquely characterizing a point in a model, i.e., two different points in the same model can have the same \mathcal{L} -description, but they cannot have the same state-description.

Definition 2.2 (\mathcal{L} -CHANGEABILITY) *A temporal model $\mathcal{T} = \langle T, <, V \rangle$ is \mathcal{L} -changing iff for some $t, t' \in T : t \neq t', Desc_{\mathcal{L}}(\mathcal{T}, t) \neq Desc_{\mathcal{L}}(\mathcal{T}, t')$.*

To say that a temporal model is \mathcal{L} -changing is to say that it can be described as changing in language \mathcal{L} , which comes down to checking whether there are at least two points in the model that differ in their \mathcal{L} -descriptions.

Fact 2.3 *Model \mathcal{M}_1 is neither \mathcal{L}_B nor \mathcal{L}_A -changing.*

Proof. Since \mathcal{M}_1 has e true at each of its points, the atomic \mathcal{L}_B -formulas ($e \triangleleft e$) and ($e = e$) will be true at every point in it, and so will their Boolean combinations. Therefore, every point in it will have the same \mathcal{L}_B -description, and therefore \mathcal{M}_1 will not be \mathcal{L}_B -changing. Similarly, since e is true at every point, every \mathcal{L}_A -formula will be true of every point because (i) each verifies e , (ii) each precedes an e , (iii) each follows an e , so any G or H-formulas or their Boolean combinations will be true at each point. There won't be any \mathcal{L}_A -formulas that any two points can disagree on, so \mathcal{M}_1 will not be \mathcal{L}_A -changing either. \square

From this fact nothing interesting immediately follows, for it simply confirms our intuitions that \mathcal{M}_1 exhibits no change, and so cannot correctly be described as changing. But now let's consider less degenerate, non-uniformly-valued, i.e., changing models.

Definition 2.4 (CHANGING MODELS) *A temporal model $\mathcal{T} = \langle T, <, V \rangle$ is changing iff it is not uniformly-valued, i.e., there is an event letter $e \in \mathcal{E}$ such that $V(e) \neq T$ and $V(e) \neq \emptyset$.*

If we now consider changing models, we'll notice that although they are \mathcal{L}_A -changing, they are *not* \mathcal{L}_B -changing. Proving this claim would demonstrate that \mathcal{L}_B is not able to describe a model that is actually changing *as* changing, whereas \mathcal{L}_A is, providing evidence for McT's second premise, (P2): change requires the A-Series. Thus, to establish (P2) it will suffice

to show that changing models are \mathcal{L}_A -changing, but not \mathcal{L}_B -changing.

Fact 2.5 *Changing models $\mathcal{M} = \langle T, <, V \rangle$ are \mathcal{L}_A -changing.*

Proof. Since \mathcal{M} is changing (non-uniformly-valued), by (Definition 2.4), there is an event letter $e \in \mathcal{E}$ s.t. $V(e) \neq T$ and $V(e) \neq \emptyset$. This means that there are at least two points t and t' in T s.t. $\langle \mathcal{M}, t \rangle \models e$ but $\langle \mathcal{M}, t' \rangle \not\models e$. But since $e \in \mathcal{L}_A$, we know that $\text{Desc}_{\mathcal{L}_A}(\mathcal{M}, t) \neq \text{Desc}_{\mathcal{L}_A}(\mathcal{M}, t')$, which by (Definition 2.2) means that \mathcal{M} is \mathcal{L}_A -changing. \square

Fact 2.6 *Changing models $\mathcal{M} = \langle T, <, V \rangle$ are not \mathcal{L}_B -changing.*

Proof. To show that \mathcal{M} is not \mathcal{L}_B -changing, by (Definition 2.2), it will suffice to show that for all $t, t' \in T : t \neq t'$, $\text{Desc}_{\mathcal{L}_B}(\mathcal{T}, t) = \text{Desc}_{\mathcal{L}_B}(\mathcal{T}, t')$. Pick arbitrary points $s, s' \in T$ s.t. $s \neq s'$ in order to show that $\text{Desc}_{\mathcal{L}_B}(\mathcal{T}, s) = \text{Desc}_{\mathcal{L}_B}(\mathcal{T}, s')$. The \mathcal{L}_B -descriptions of s and s' contain all the \mathcal{L}_B -formulas true at them. Since by (Definition 1.2) the truth-conditions of \mathcal{L}_B -formulas don't depend on the point of reference, s and s' will both have the same \mathcal{L}_B -formulas true at them, so they will both have the same \mathcal{L}_B descriptions. Since s and s' are arbitrary, we generalize, obtaining the required result. \square

These facts show that change requires the A-Series. Thus, based on two premises, namely that (P1) time requires change, and that (P2) change requires the A-Series, McT establishes the first part of his argument, namely that (I) time requires the A-Series. Now we turn to the second part of his argument, to the claim that (II) the A-Series is incoherent.

§2.2 A-Series $\rightarrow \perp$

McT starts the reductio with the premise that (P3) the predicates 'past', 'present', and 'future' are contraries and therefore cannot be simultaneously true of any event in time. With that premise, McT's strategy is to show that the A-Series entails that any two of these predicates is simultaneously satisfied by some moment, effectively reducing the A-Series to absurdity. In fact, McT claims that all three predicates hold of every event (468). He reasons in the following way. Take a moment m that's in the future, like the due date of this paper. In such a case we would say not only that m is future, but since it's future, then m will be present (on the date it's due) and then m will be past (after it's due). Generalizing what is true of m to all events McT concludes that (P4) 'all three incompatible terms are predicable of each event' (468). As promised, since the A-Series leads to a violation of the aforementioned incompatibility premise, the A-Series is reduced to absurdity (467).

But is McT right about (P4)? I want to argue that (\star): *m is future, m will be present, m will be past* are *not* incompatible determinations because 'was' \neq 'will be' \neq 'is.' McT

has fallen victim to the identification of tensed ‘is’ with ordinary ‘is’. If tensed ‘is’ is not ordinary ‘is’, then can one give an account of (\star) that will allow us to avoid violating the incompatibility premise (P3)? I think we can, and in the rest of this section I will show how. We can define tensed variants of ‘is’ as in the table below. We can then prove that McT’s claim is not valid on our temporal models.

Fact 2.7 $\phi \rightarrow (P\phi \wedge F\phi)$ is not valid on temporal models.

Proof. Consider a counterexample model $\mathcal{T} = \langle T, <, V \rangle$ with valuation $V(e) = \{t_1\}$. For such a model, $\langle \mathcal{T}, t_1 \rangle \models e$. But since e is only true at t_1 , $\langle \mathcal{T}, t_1 \rangle \not\models Pe$, and therefore $\langle \mathcal{T}, t_1 \rangle \not\models (Pe \wedge Fe)$. \square

| | Past | Present | Future |
|---------|----------|---------|----------|
| Is | $P\phi$ | ϕ | $F\phi$ |
| Was | $PP\phi$ | $P\phi$ | $PF\phi$ |
| Will be | $FP\phi$ | $F\phi$ | $FF\phi$ |

Using the tense-table we can formalize (\star) correctly as: $\phi \rightarrow (PF\phi \wedge FP\phi)$, which says that ‘what *is* present, *was* future, and *will be* past.’ We can then prove that it is in fact valid.

Fact 2.8 $\phi \rightarrow (PF\phi \wedge FP\phi)$ is valid on left and right-unbounded temporal models.⁸

Proof. Pick an arbitrary pointed model $\langle \mathcal{T}, t \rangle$ that has neither a beginning nor an end point, and s.t. $\langle \mathcal{T}, t \rangle \models \phi$. Since the model is not bounded on the left, there is a point $t' < t$ s.t. $\langle \mathcal{T}, t' \rangle \models F\phi$. But since t comes after t' , we have it that $\langle \mathcal{T}, t \rangle \models PF\phi$. By symmetric reasoning we get also that $\langle \mathcal{T}, t \rangle \models FP\phi$. Combining these, we get $\langle \mathcal{T}, t \rangle \models (PF\phi \wedge FP\phi)$. Since $\langle \mathcal{T}, t \rangle$ was arbitrary, we know that the principle is valid on the specified models. \square

These facts constitute a rejection of premise (P4), and therefore of McT’s argument. McT, however, anticipates our diagnosis, accusing it of circularity, for, in making tense-distinctions, he says, ‘it assumes the existence of time in order to account for the way in which moments are past, present, and future’ (468). The argument is about the reality of time, so he worries that we’re not entitled to distinguish the different tenses of ‘is’, for if there is no time, then the distinction collapses. However, McT is not very clear in his objection, so to avoid misrepresenting him I have looked at a number of commentaries. Markosian [2008] agrees with my reading of McT’s claim, but doesn’t take a stand on its correctness (§4). Dummett [1960], in his defense of McT’s argument, also doesn’t take a firm stand on this point. Prior [1967], however, agreeing with both analyses, disagrees with McT’s ‘perverse conclusion’ that the tensed analysis is circular (5-6). I too am compelled to think that there is no such problem with the tensed analysis, not because that’s the strategy

⁸The restriction to the class of unbounded temporal models is supposed to capture the intuition that the principle is valid intuitively only if we’re not evaluating it at the beginning or the end of time, for at those points no one would say that what is true was future because everyone would agree that there is no past point with respect to which the beginning of time would be future.

that I've taken, but because I'm not convinced that there is any circularity involved in it.

Conclusion

We started with McTaggart's informal characterization of the two series (§1.1). We formalized the distinction (§1.2), and took note of McTaggart's presupposition that the two series are exhaustive (§1.3). We got to McTaggart's argument in §2. We motivated his first two premises, viz., (P1) time requires change, and (P2) change requires the 'A-Series' (§2.1). We also accepted his third premise, viz., (P3) past, present, and future are contraries, but we showed that his fourth premise, viz., (P4) past, present, and future simultaneously hold of every event, is false. We concluded by noting an objection that might be raised against our analysis, an objection that McTaggart himself seems to have advanced.

References

- [1] Dummett, M. (1960) "A Defense of McTaggart's Proof of the Unreality of Time."
In *The Philosophical Review*, Vol. 69, No. 4. (Oct., 1960), pp. 497-504.
- [2] Kamp, H. (1971) "Formal properties of 'now'". *Theoria*, 37: 237-273.
- [3] McTaggart, J. M. E., (1908) "The Unreality of Time." In *Mind* 17: 457-73.
- [4] Markosian, N., (2008) Time. In *Stanford Encyclopedia of Philosophy*.
- [5] Prior, A. N., (1967) *Past, Present, and Future*.
- [6] Shoemaker, S. (1969) "Time Without Change," *Journal of Philosophy*, 66, pp. 363-381.