International Journal for Philosophy of Religion
https://doi.org/10.1007/s11153-023-09876-z

ARTICLE

An inexplicably good argument for causal finitism

Ibrahim Dagher[1]

Received: 11 April 2023 / Accepted: 20 May 2023
© The Author(s) 2023

Abstract
Causal finitism, the view that the causal history of any event must be finite, has garnered much philosophical interest recently—especially because of its applicability to the Kalām cosmological argument. The most prominent argument for causal finitism is the Grim Reaper argument, which attempts to show that, if infinite causal histories are possible, then other paradoxical states of affairs must also be possible. However, this style of argument has been criticized on the grounds of (i) relying on controversial modal principles, and (ii) providing a false diagnosis of the paradoxes involved. In this paper, I develop a new kind of Grim Reaper argument immune to these criticisms. I show that, by using insights from the literature on time travel, causal finitists should instead argue that infinite causal histories are *problematically inexplicable*, as they entail the possibility of unexplained foiling mechanisms. The fruits of this paper are that (i) a novel supporting argument for the Kalām is developed, and (ii) along the way of building this argument, it is shown that the literatures on time travel and causal finitism are deeply and intimately connected.

Keywords Causal finitism · Benardete · Kalām · Time travel · Grim reaper argument

Introduction

As of late, there has been a growing interest in many philosophical circles—both scholarly and public—with the Grim Reaper argument for causal finitism, the view that there cannot be infinitely many causes behind an effect. While the argument has received a variety of treatments, most notably by Koons (2014), Pruss (2018), and Erasmus and Luna (2020), a common line of reasoning undergirds all of these varia-

✉ Ibrahim Dagher
ibrahim.dagher12@gmail.com

[1] Philosophy Department, University of California, Davis, CA, US

Published online: 08 June 2023

tions. The argument starts with a logically inconsistent story: infinitely many Grim Reapers are ordered in such a way that (i) one of them must ϕ, and yet also (ii) it is not possible for a single one of them to ϕ. The argument then proceeds by advocating for some modal principle—usually a recombination or Lewis-style patchwork principle—which entails that 'if an infinite causal chain is possible, the Grim Reaper story is possible.' However, because this story *isn't* possible, we arrive at the conclusion that infinite causal chains must also not be possible (i.e., causal finitism is true).

Interestingly, this argument bears much structural similarity to a well-known argument against backward time travel. This is the Grandfather argument, which attempts to establish the metaphysical impossibility of backward time travel by contending that if backward time travel is possible, it is possible for the Grandfather story to obtain. And, since that story is logically inconsistent, backward time travel must be metaphysically impossible. Just as with causal finitism, the idea here is that some concept *c* (infinite causal chains, backward time travel) must be impossible because its possibility would, by some modal principle or other, entail that some inconsistent story is genuinely possible.

Whilst a few authors have mentioned—often in passing remarks—the parallel structure of these arguments, the work in each of these spaces has hitherto progressed in almost complete independence.[1] In this paper, I seek to put an end to this isolation. My chief aim is to advance the causal finitism dialectic by developing a new variation of the Grim Reaper argument. To do this, I will draw on new developments recently made in the literature on the Grandfather argument. Specifically, I will apply (and modify) the style of reasoning developed by Yael Loewenstein (2022) in constructing a new variation of the Grandfather argument. Rather than relying on controversial (if not problematic) modal principles to commit advocates of backward time travel to the possibility of an inconsistent story, Loewenstein argues that the possibility of backward time travel leads to a problematic kind of *inexplicability*—the kind that ought to lower one's credence in the possibility of time travel. I hope to show that causal finitists should make a similar kind of move.

I will proceed as follows. In § 2, I explicate in more detail the Grim Reaper argument for causal finitism, focusing especially on the version developed by Pruss (2018). In § 3, I shift my attention to outlining Loewenstein (2022)'s new Grandfather argument. In § 4, after showing that Loewenstein's proposal, in its present form, cannot yet be applied to the Grim Reaper argument, I explore how we might modify the reasoning to expand it to cases like the Grim Reaper argument. I then apply the reasoning in § 5, before providing a brief concluding remark in § 6.

[1] I should note that Pruss (2018, p. 61) remarks that 'there is a similarity between the paradox-based arguments for causal finitism and some arguments based on the Grandfather paradox against time travel,' and briefly spells out how his Grim Reaper arguments can be paralleled as arguments against time travel. Similarly, Erasmus and Luna (2020, p. 168) note that 'it is helpful to note that the preceding argument is akin to the usual intuitive argument against the possibility of time travel.' On the other side, Effingham (2020) briefly suggests how his results concerning the Grandfather paradox might be applied to resolve Benardete paradoxes.

The grim reaper argument

The version of the Grim Reaper argument that I will focus my attention on is the one developed by Alexander Pruss (2018). I'll start by first spelling out the argument, and then showing how Pruss refines it in response to an important objection.

Imagine that there is a lamp with a switch, such that whenever the switch is hit the lamp switches its state from "on" to "off", or else "off" to "on". A grim reaper is a machine which, when set for a time *t*, checks the state of the lamp at *t*. If the lamp is off, the reaper turns the lamp on. If the lamp is on, the reaper does nothing. We can imagine that each reaper is equipped with a dial on it that indicates the time it is set to check at. Now, imagine further that there are infinitely many reapers, one of which is set to check the lamp at 11:00, another at 10:30, another at 10:15, another at 10:7.5, another at 10:3.75, etc. Let us stipulate further that the present time is 9:00, that the lamp is off, and that nothing besides a reaper can change the state of the lamp.

Prima facie, nothing seems to have gone awry just yet. If infinite causal chains are possible, this seems like a perfectly innocuous setup that we could construct. However, appearances can be deceiving. Consider the state of the lamp at 12:00: is it on or off? Surely, it must be on. For either it was on or off at 11:00, and if it was off, it would have been turned on by the 11:00 reaper. So, at 12:00 the lamp is on. And, by stipulation, the lamp is off at 9:00. So, a reaper flipped the lamp. But which one? It can't be the 11:00 reaper, since the lamp can't be off by the time it gets to 11:00—the 10:30 reaper would have already turned it on. Note, though, that it can't be the 10:30 reaper that turned on the lamp either, for the same exact reasoning applies—the 10:15 reaper would have already turned it on. Interestingly, this exact reasoning applies for *every* reaper. Because there are infinitely many reapers between 10:00 and 11:00, each reaper is such that there is some reaper that precedes it. So, the lamp is not turned on by *any* reaper, and so it is off by 12:00: a contradiction.

For Pruss, the lesson we ought to draw from this surprising story is that infinite causal chains are metaphysically impossible. After all, it seems as though this would be a perfectly cogent scenario that we could construct if infinite causal chains were possible. Thus, Pruss (2018, pp. 47–48) gives the following argument:

(1) If causal finitism is false, then the Grim Reaper story is possible.
(2) The Grim Reaper story is impossible.
(3) Therefore, causal finitism is true.

The key premise here, of course, is (1). While the modal connection between the possibility of an infinite causal chain and the Grim Reaper story seems *prima facie* plausible, it is exactly this connection that has been challenged by others in the literature. Most notably, Shackel (2005) argues that there is no positive reason to think (1) is true.[2] The thrust of the objection is that the Grim Reaper story is really just a conjunction of two propositions: (i) there is a (backward) infinite causal chain, and (ii) the chain is such that, for each of its constituents *c*, a property P is true for *c* just in case

[2] Or, at the very least, something *like* (1), since Shackel's diagnosis of the paradox precedes Pruss' version of the argument.

P is true nowhere before *c*. Note that these two propositions are jointly inconsistent, and for the same reason as the Grim Reaper story: P must be true for some *c*, yet it cannot be true for any *c*. The idea is that those who deny causal finitism can simply claim that all the Grim Reaper story teaches us is that it is the *conjunction* of these two propositions that is impossible. All kinds of infinite causal chains can obtain, just not the ones so ordered. In fact, it is precisely because the conjunction of (i) and (ii) is inconsistent that opponents of causal finitism have good reason to resist (1). Absent further motivation for (1), the Grim Reaper argument lacks any dialectical force.

In response, Pruss (2018, p. 49) attempts to provide exactly this. He attempts to motivate (1) by asking us to entertain an additional story. Imagine another Grim Reaper setup, except that now there is one extra reaper set to check the lamp at 9:59. This variant of the story, called the Prefixed Reaper story, is consistent: come 9:59 the lamp will be switched on and the rest of the reapers do nothing. By adding a first reaper, we remove the paradox. So, surely, if infinite causal chains are possible, a world with a Prefixed Reaper story is possible. After all, this setup is perfectly consistent!

However, now consider what would happen if, at 9:58, a tinkerer, Rachel, put her hand on the dial of the 9:59 reaper and attempted to turn it forward, such that the time it would check the lamp at would become later than 10:00. Note that, upon doing this, we return back to the problematic Grim Reaper setup—which exists at no possible world. But how odd is this! Are we really to believe that it is possible that Rachel have her hand on a dial in a Prefixed scenario, and yet it is *impossible* that she move her hand? The Prefixed Reaper story seems so similar to the Grim Reaper story that, if *it* is possible, surely the Grim Reaper story is too. But this gives way for Pruss (2018, p. 49) to make the following minor modification to the original Grim Reaper argument:

(4) If causal finitism is false, then the Prefixed Reaper story is possible.
(5) If the Prefixed Reaper story is possible, then the Grim Reaper story is possible.
(6) The Grim Reaper story is impossible.
(7) Therefore, causal finitism is true.

This argument seems, *prima facie*, to provide strong support for causal finitism. (4) seems like it must follow from the possibility of an infinite causal chain. Additionally, (5) enjoys considerable intuitive support. It is further supported by appeal to modal principles of recombination, which entail that if the Prefixed Reaper scenario is possible, the simple recombination of a single dial is surely possible too. However, opponents of causal finitism can make a number of moves at this point. First, they might reject the truth of such recombination principles and thus find no reason to accept (5). More interestingly, the opponent of causal finitism might reject (4) in the same way they rejected the original (1). The Prefixed Reaper story is impossible, even if causal finitism is false, precisely because certain modal recombination principles are true. After all, mere consistency is not sufficient for metaphysical possibility. The opponent of causal finitism is by no means committed to the metaphysical possibility of the Prefixed Reaper story *just* because it is consistent.

I will not take a definitive stand on whether (4)-(7) is a successful argument or not. My point is merely to show that many of the same dialectical issues plaguing the original Grim Reaper argument plague this one. Thus, it would be helpful to consider new ways of constructing this argument—ways that needn't rely on premises like (5).

In what follows, I will attempt to do exactly that. To do this, I will draw on a new variant of the Grandfather argument developed by Yael Loewenstein (2022). My hope is that a similar variant can be constructed for the Grim Reaper argument.

Loewenstein's new grandfather argument

Despite the structural similarities, there is one key difference between arguments based on the paradoxes of time travel and arguments based on the paradoxes of infinite causal chains. Many arguments against time travel tend to try to establish their conclusion by claiming that a particular fact, entailed by the possibility of time travel, is problematically *inexplicable*.[3] Comparatively, arguments against infinite causal chains tend to try to establish their conclusion by claiming that, if an infinite causal chain is possible, some other inconsistent state of affairs must too be possible.

With this in mind, let's now examine how this inexplicability-style argument is developed by Loewenstein (2022). Loewenstein is concerned with the following standard autoinfanticide story:

> **Autoinfanticide**: An adult woman, Timna_O, travels back in time to t_0, where she has her eyes set on killing her baby self, Timna_Y. Despite the fact that Timna_O is a trained warrior, she will fail at t_1 in her attempt to kill Timna_Y. She will slip on a banana peel, be toppled by a sudden gust of wind, or be spontaneously stung by a wasp. Whatever it is, Timna_O will be foiled in her murder attempt at t_1.

At this point, the standard move made by proponents of the impossibility of time travel is to argue that the fact that Timna will be foiled in *each* of her attempts is a problematically inexplicable fact—and that this should count as evidence against the claim that time travel is possible. It seems as though Timna *can* succeed in such a simple task, and yet she undoubtedly cannot.

However, there is a now-orthodox response to this argument. Following David Lewis (1976), the standard response is that any inexplicability in **Autoinfanticide** is merely illusory. For there are two notions of 'can' at play here. If we consider only facts that are true (temporally) prior to the attempt, then, of course, Timna_O *can* kill her younger self: there is nothing to stop her, she is a trained warrior, etc.[4] However, if we also consider facts that are true *after* the attempt—facts like 'Timna_Y will survive and grow up to become a time traveler'—then, of course, Timna_O *cannot* kill her

[3] However, for discussion of a style of time travel paradox based on the (false) possibility of successful autoinfanticide, see Effingham (2020).

[4] Pre-attempt facts include facts like 'the sky was blue in 1787,' 'Timna_Y is very fragile and incapable of stopping attackers,' etc. The point of emphasis is that Timna_O's killing Timna_Y is metaphysically compossible with this set of facts.

younger self: something or other will foil any attempt she makes. These two notions of 'can' are denoted by $can_{(<t1)}$ and $can_{(\infty)}$. The idea of the standard solution is that the inexplicability of the situation vanishes upon realizing that $Timna_O$ $can_{(<t1)}$ succeed, but that she $cannot_{(\infty)}$ succeed.

Loewenstein argues against this solution to the Grandfather paradox: she maintains that, *pace* the philosophical orthodoxy, a problematic inexplicability remains in cases like **Autoinfanticide** that Lewis' analysis does not account for. More specifically, Loewenstein argues that it is false that $Timna_O$ $can_{(<t1)}$ succeed in killing $Timna_Y$. $Timna_O$ $cannot_{(<t1)}$ succeed, and it is this fact that makes the case problematically inexplicable.[5]

Loewenstein's proposal

To get a better grasp on Loewenstein's argument, it will help to look at her proposal a bit more closely. What Loewenstein seeks to do is provide a general framework for determining, in any given case where an agent will be foiled in attempting to ϕ (where ϕ-ing is some common, easily performed action), whether the agent's failure to ϕ is explicable or problematically inexplicable. To do this, she proposes a sufficient condition for explicable failures. This is Loewenstein (2022, p. 5)'s proposal:

> **Contingent failure**: Although the attempt in question failed, the agent $could_{(<t1)}$ have succeeded.

The idea behind this proposal is that, when one is confronted with a case of an agent that will be foiled in any attempt to ϕ, one can diffuse the seeming inexplicability of this situation by testing whether the agent $could_{(<t1)}$ have succeeded. The motivation for **Contingent Failure** can best be seen by applying it directly to an easy case. Consider the following situation:

> **Standing up**: At this moment in time t_1, it is true that, at t_2, I remain seated. I attempt to stand up at t_1. I will fail in my attempt: I will slip on a banana peel, be toppled by a gust of wind, or stung by a wasp. Whatever it is, something or other will foil any attempt I make to stand up. In circumstances where I will remain seated, any attempt I make to stand up will be foiled.

This case, just like **Autoinfanticide**, is a case in which an agent will assuredly be foiled in any attempt to ϕ. Presumably, though, we don't think that the foiling involved in **Standing Up** is anything problematic. It is perfectly explicable that, in circumstances where I will remain seated, any attempt I make to stand up will be foiled. What is it, though, that makes this case non-problematic? Loewenstein's suggestion is that the reason this case isn't problematic is because my failure to stand up

[5] Importantly, here $Timna_O$ and $Timna_Y$ are not identical in the sense of numerical identity. They are what Loewenstein (2022, p. 8) calls *origination identical*: "the kind of identity at issue here is the kind that would make it such that if $Timna_Y$ is identical to $Timna_O$, then the former's death is inconsistent with the latter's existence."

at t_1 is still, at some intuitive level, a *contingent* failure. Albeit that I will assuredly fail to stand up, there is an important sense in which, at t_1, I merely *happened* to fail. There is an important sense in which my failure to stand up, just like my failure to not trip over a rock or to do the dishes, merely happened to occur: it was just the result of contingent causal happenstance, and there is nothing inexplicable about that kind of failure.

How can we capture the intuitive contingency of my failures in **Standing Up**? Loewenstein's proposal tells us that we can test for the contingency of a failure by testing whether the agent could$_{(<t1)}$ have succeeded. In **Standing Up**, when we hold fixed the future fact that I will remain seated, of course I will be foiled in my attempt to stand up. However, consider merely whether, in **Standing Up**, I could$_{(<t1)}$ have succeeded in standing up. I can! For in considering whether I can$_{(<t1)}$ succeed in standing up, we test merely for whether my standing up is metaphysically compossible with the set of facts true prior to t_1. This set of facts does *not* include, among other things, the fact that I will remain seated, and so I can$_{(<t1)}$ stand up.[6] It is *this* sense in which my failures in **Standing Up** are all merely contingent: for while I failed, I could$_{(<t1)}$ have succeeded, and so at t_1 I merely happened to fail. Holding fixed the contingent fact that I will remain seated makes explicable why it is that I will be foiled in any attempt I make to stand up.

The broad lesson here is this: when faced with a case in which an agent will be foiled in any attempt to ϕ, where ϕ-ing is some appropriately simple action, we must check whether the agent could$_{(<t1)}$ have succeeded. If the agent can$_{(<t1)}$ succeed, then the failures are all explicable, as they are, at bottom, just contingent failures: they are only required in virtue of our holding fixed how the future unfolds.

Turn now to **Autoinfanticide**. We might be inclined to apply **Contingent Failure** and diffuse the seeming inexplicability of Timna$_O$ having to fail to kill Timna$_Y$ in the same way. Simply consider whether Timna$_O$ can$_{(<t1)}$ kill Timna$_Y$. It appears as though she can—we're no longer holding fixed the future fact that Timna$_Y$ happens to live on past her toddler life. Here, however, Loewenstein claims that **Autoinfanticide** is importantly different from **Standing Up**. For, even when we don't hold fixed that Timna$_Y$ happens to survive past childhood, it is *still* the case that Timna$_O$ can't succeed in killing her. That is to say, Timna$_O$ can't$_{(<t1)}$ succeed. Why? The details of the argument are beyond the purposes of this paper—but the basic argument is that there are enough pre-t_1 facts such that they ground, or entail, the fact that Timna$_O$ is identical to Timna$_Y$. This makes it impossible for Timna$_O$ to kill Timna$_Y$, even from a pre-t_1 perspective.

We needn't be concerned with whether this is a sound argument for the proposition that Timna$_O$ can't$_{(<t1)}$ succeed. All that matters for our purposes is seeing that, *if* Timna$_O$ can't$_{(<t1)}$ succeed in killing Timna$_Y$, then the fact that Timna$_O$ is foiled in each of her attempts to kill Timna$_Y$ is problematically inexplicable. Why is this the case? It is because Timna$_O$'s failure, unlike my failure to stand up, is not at all a contingent failure. If Timna$_O$ can't succeed from any temporal perspective, then it's not the case that she merely *happens* to be foiled in her attempts to kill her younger

[6] Presuming, of course, that the past times do not also have the future-tensed fact that I will remain seated at t_2.

self: she *will assuredly* be foiled. But this is strange: how can it be that there is always some causal mechanism foiling Timna_O's attempts, *when we aren't holding fixed that she fails*? As Loewenstein (2022, p. 10) says: "if we can't attribute the presence of the foiling mechanism to mere chance, then, that it will be there when needed, i.e., in most cases when Timna_O decides to try to kill Timna_Y, remains inexplicable."

Generalizing Loewenstein's proposal

It might seem that, having spelled out the details of Loewenstein's proposal for determining what failures are explicable and inexplicable, we are now at the point where we can develop a new Grim Reaper argument. After all, can't the Prefixed Reaper story be cast as a story of agential failure, just like **Standing Up** and **Autoinfanticide**? Using **Contingent Failure**, we should be able to get some insight into whether causal finitists can argue that Rachel's failure to turn the 9:59 dial is problematically inexplicable, and thus counts against accepting the possibility of infinite causal chains.

Unfortunately, a few more details need to be filled out. Consider if we attempted to do this by examining Rachel's failure to turn the dial:

> **Prefixed reaper**: Rachel, at 9:58, is standing in front of a Benardete-sequence of reapers, with an additional reaper set to check the lamp at 9:59. Rachel has her hand on the 9:59 reaper's dial and attempts to turn it so that it is set for a later time. Rachel will fail in attempting to turn this dial: she'll slip on a banana, be toppled by a gust of wind, etc.

How might we apply **Contingent Failure**? We have to test for whether Rachel's failure to turn the dial at 9:58 is a contingent failure. To do this, we consider whether Rachel $\text{can}_{(<9:58)}$ turn the dial. Prior to 9:58, is it true that Rachel is standing next to a Benardete-sequence of reapers? If it is, then Rachel $\text{can't}_{(<9:58)}$ turn the dial. If it isn't the case that this fact is true prior to 9:58, then Rachel $\text{can}_{(<9:58)}$ turn the dial. Either way, though, this heuristic is unhelpful. All **Contingent Failure** tells us is that Rachel's turning the dial is inconsistent with the existence of an adjacent Benardete-sequence of reapers. This, though, we already knew. What we are interested in is unveiling whether Rachel's failure merely *happens* to occur, like in **Standing Up**, or whether it (inexplicably) will assuredly occur despite not holding fixed contingent facts, like Timna_O's. We might be inclined to just say that, because we aren't explicitly holding fixed any future fact about what Rachel does, the fact that she fails is inexplicable. This, though, will be largely unpersuasive without some further explanation as to why this is the case.

Consider another example that **Contingent Failure** is unable to handle:

> **Machine**: There is a machine such that, for any proposition uttered into it, it actualizes that proposition. I go up to the machine, attempting to utter 'my shirt is red and not-red.' I will assuredly fail in attempting to utter this proposition: I'll slip on a banana, be toppled by a gust of wind, etc.

Presumably, the existence of a foiling mechanism whenever I attempt to utter a proposition that I am generally capable of uttering, merely because I am standing in the vicinity of a special machine, is problematically inexplicable. But is it because I could$_{(<t1)}$ have succeeded in uttering this proposition? This cannot be the right explanation: it's not clear what roles the pre-t_1 and post-t_1 sets of facts are even playing in these cases. What **Machine** and **Prefixed Reaper** show is that Loewenstein's proposal is too "temporally loaded." That is to say, the proposal relies too much on temporally divided sets of facts. It can thus explain why an agent's assured failure is explicable or inexplicable only in cases whose salient features are divided across different times (e.g., cases involving future facts or backward time travel). What we need is a generalization of **Contingent Failure**: a broader framework that can deal with cases like **Machine** and **Prefixed Reaper**, where an agent will be foiled in her attempts to ϕ merely in virtue of being near some concrete state of affairs.

The new proposal

To generalize **Contingent Failure** successfully, we need to develop some mechanism for identifying when an agent's failure to ϕ is merely a contingent failure—without relying on whether the agent could$_{(<t1)}$ succeed.

To do this, I'm going to make use of the notion of a "circumstance." A circumstance C is a set of propositions. In particular, circumstances are subsets of the propositions that are true at a world *w* and time *t*. The propositions in a circumstance are sufficiently few and general that many worlds share one and the same circumstance. For example, consider the circumstance {'There is a boy', 'He is licking an ice cream cone'}. This circumstance is true in all kinds of worlds—worlds just like this one, worlds where everything is the same as this one except the sky is yellow, worlds where dragons exist, etc. In this sense, a circumstance resembles the familiar notion of a "world segment": it is a set of facts shared by many different worlds.

Let's now walk through the story of **Standing Up** and see if we can explain the failure in terms of circumstances, rather than in terms of whether I can$_{(<t1)}$ succeed in standing up.

Consider the circumstance {'I attempt to stand up', 'It will be that I remain seated'}. At all the worlds in which this circumstance is true, I will be foiled in my attempt to stand up. *How* I am foiled in my attempt might be a different story from world to world: in some worlds, I'll slip on a banana peel, in others, I'll be toppled by a gust of wind or stung by a wasp. But, at all of them, I will be foiled. We can now ask: is this collective fact—that my attempts to stand up are foiled in *all* these worlds—problematically inexplicable?

The answer, of course, is that this collective fact isn't problematically inexplicable. But why is that? My suggestion is that this is because of the nature of the circumstances we're holding fixed. The circumstance {'I attempt to stand up', 'It will be that I remain seated'} *builds in* a causal mechanism that explains why each failure is foiled. The proposition 'it will be that ϕ' comes with various causal mechanisms (banana peels, powerful winds, scary wasps) in place. 'It will be that ϕ' entails, or at

least builds in, "if an agent attempts ~ϕ, some causal mechanism CM intervenes to foil the attempt."[7]

Here is another way to see it: when I claim "it will be that I remain seated," I am claiming (fundamentally) something about *the world*. I'm claiming that the world is such that it (causally) ensures I remain seated. It's thus not surprising, or problematic, that every world at which it will be that I remain seated, and I attempt to stand up, there is always a foiling mechanism in place to stop me. The circumstances build in this foiling mechanism. It's not as though we held fixed some circumstance wholly unrelated to my standing up, and then it was true that at all the worlds at which I attempt to stand up, I was foiled in my attempt. *That* would be problematically inexplicable.[8]

So, more broadly, when a circumstance C is held fixed, and an agent is being foiled in their attempts to ϕ at all the worlds with C, this is explicable if (and only if) C builds in a causal mechanism which explains why each token attempt is foiled. Whenever an agent is being foiled in their attempts to ϕ at all C-worlds, C ought to provide an explanation of the satisfaction of the description that this class of worlds all have a foiling mechanism.[9] If C provides no such explanation, then we have problematic inexplicability.

Now, let's turn to **Machine**. Using the notion of a circumstance explaining the existence of foiling mechanisms, or "building into" a world the relevant causal mechanisms, we can test whether my failure to utter 'my shirt is red and not-red' is problematically inexplicable. Consider the circumstance '"I attempt to utter that my shirt is red and not-red', 'I am standing next to the machine'}. Again, at all the worlds at which this circumstance is true, I will be foiled in my attempt to make that utterance. In some worlds, I'll slip on a banana peel, in others, I'll be toppled by a gust of wind or stung by a wasp. We now ask: is this collective fact—that my attempts to speak are foiled in all these worlds—problematically inexplicable?

The answer is that this *is* inexplicable. Why? Again, it is because of the nature of the circumstances we're holding fixed. These circumstances do not build in any

[7] It is important to note that we are (also) assuming something in the background here: that no events are uncaused. Of course, this has been an assumption all along: when we claim that, in any world where I attempt to stand up in circumstances such that I will remain seated I will be foiled in my attempt, we of course assume the worlds we're considering do not allow for uncaused phenomena. This is an additional proposition in our circumstance.

[8] While I did not include an analysis of **Autoinfanticide** in terms of circumstances, I believe Loewenstein's argument that Timna_O can't$_{(<t1)}$ kill Timna_Y can be translated equally well (if not better) in these terms. Loewenstein wants to argue that *even when we don't hold fixed that Timna$_Y$happens to survive*, there will still be a foiling mechanism that prevents Timna_O from successfully killing Timna_Y. We can represent this by considering the circumstance {'Timna_O appears at t_0', 'Timna_O attempts to kill Timna_Y'}. Loewenstein contends that (i) at all the worlds with these circumstances, Timna_O will be foiled in her attempt, and that (ii) this is problematically inexplicable. The argument for (i) here will be the same as Loewenstein's original argument: there are enough facts in C to ground, or entail, the proposition that $\text{Timna}_O = \text{Timna}_Y$. However, we are now in a better position to explain why, if we accept (i), (ii) is true. For, *merely* given the fact that Timna_O appears at a time, and that she atttempts to kill Timna_Y, we arguably have not said enough about *the world* to explain why there is always a foiling mechanism in place to foil Timna_O.

[9] Note here that when I say C ought to provide an *explanation* of the satisfcation of the description that this class of worlds all have a foiling mechanism, I now have in mind a *non-causal* explanation. C ought to make sense of, or make clearer why, this description is satisfied.

causal mechanism which can explain the universal foiling. That a special machine is in the vicinity does not come with it a causal explanation for why there is always an intervention in my attempt to utter particular propositions. In saying that there is such-and-such machine by my side, I have said nothing about the world—besides the fact that it has such-and-such spatial arrangement. The universal existence of a foiling mechanism for all of my attempts to speak cannot be explained by the common factor that all of these worlds have such-and-such machines near my utterance. Thus, we get a problematically inexplicable state of affairs, and so we have some reason to be skeptical of the possibility of machines that actualize any proposition uttered into them.

Of course, it should be noted that it's still true that I will always be foiled in my attempts to make my intended utterance: for otherwise a dialetheia would obtain. But it is exactly *this*—that a dialetheia would obtain—that cannot explain the universal existence of a foiling mechanism which causally prevents my attempts to speak. For logic has no causal force: it does not stay the hands, or voices, of anyone. On the other hand, facts like 'I will remain seated' *do* have causal force: they come imbued in their content with facts about how the (causal) course of the world proceeds.

Given all of this, we now have an inexplicability-based argument against the possibility of these kinds of machines. The possibility of a machine which actualizes any proposition uttered into it entails that, at all the worlds at which this machine exists and I attempt to utter 'my shirt is red and not-red,' I am foiled in my attempt. This foiling, though, is problematically inexplicable, and so we should be skeptical of the possibility of these sorts of machines.

Applications to the grim reaper argument

Having explored Loewenstein's style of inexplicability-based argument, and expanded the framework to test effectively for inexplicability in a broader range of cases, we're ready now to apply these fruits to the original dialectic of concern: the Grim Reaper argument for causal finitism. Recall that Pruss' argument centers around:

> **Prefixed Reaper**: Rachel, at 9:58, is standing in front of a Benardete-sequence of reapers, with an additional reaper set to check the lamp at 9:59. Rachel has her hand on the 9:59 reaper's dial and attempts to turn it so that it is set for a later time. Rachel will fail in attempting to turn this dial: she'll slip on a banana, be toppled by a gust of wind, etc.

To support causal finitism, we can make one of two moves at this point. First, we can proceed as Pruss does, and use **Prefixed Reaper** along with a modal principle of recombination to claim that the denial of causal finitism commits one to the possibility of Rachel succeeding in turning the dial. This approach, as we noted previously, is mired with difficulty. The second option is to say that while Rachel will always be foiled in her attempts to turn the dial in **Prefixed Reaper**, it is *this* which is problematically inexplicable, and we should thus decrease our credence in the possibility of an infinite causal chain (i.e., increase our credence in causal finitism).

The only work left for us to do is to establish that the foiling in **Prefixed Reaper** is indeed problematically inexplicable. We established the following framework for determining inexplicability: when a circumstance C is held fixed, and an agent is being foiled in their attempts to ϕ at all the worlds with C, this is explicable if (and only if) C builds in a causal mechanism which explains why each attempt is foiled. Let's now apply this insight to **Prefixed Reaper**. Consider the circumstance {'Rachel attempts to turn the 9:58 dial', 'Rachel is standing next to a Benardete-sequence of reapers'}. At all the worlds in which these circumstances are true, Rachel will be foiled in her attempt to turn the dial. In some worlds, she'll slip on a banana peel, in others she'll be toppled by a gust of wind or stung by a wasp. Again, we ask: is this collective fact—that Rachel's attempts to turn a dial are foiled in all of these worlds—problematically inexplicable?

For almost the exact same reasons as **Machine**, this fact is indeed problematically inexplicable. Simply consider the nature of the circumstances we're holding fixed. The fact that there are many *other* reapers, set for later times and awaiting far away from Rachel, is not the kind of fact that builds in a causal mechanism. It says nothing about the world—beyond that we have a certain spatial arrangement of odd machines. For all the same reasons iterated for **Machine**, the circumstances in **Prefixed Reaper** simply do not offer an explanation for why there is a universal foiling mechanism. Thus, the fact that Rachel's attempts are always foiled is problematically inexplicable, and so *prima facie* evidence in favor of causal finitism.

I'd like to make one last note here. The reasoning for why **Prefixed Reaper** is inexplicable is extremely similar to the reasoning for why **Machine** is inexplicable. It is so similar that I suspect there is a strong argument for why the treatment of **Prefixed Reaper** provided herein is the correct treatment: any objection to **Prefixed Reaper** being inexplicable is, *prima facie*, an equally compelling objection to **Machine** being inexplicable. However, it seems as though the explanation I've provided for why these machines are inexplicable is the correct explanation. If so, it follows that **Prefixed Reaper** is also inexplicable. Alternatively, even if one is hesitant to accept the explanation I've developed here, it seems as though **Machine** and **Prefixed Reaper** are symmetrical cases, such that if one thinks one of these cases is inexplicable, for whatever set of reasons, the other must also be inexplicable, for very similar reasons. And, because **Machine** seems to obviously be inexplicable, so too is **Prefixed Reaper**.

Conclusion

In total, this paper develops a new variant of the Grim Reaper argument. We did this by focusing on inexplicability, rather than controversial modal principles. By modifying and applying insights from the literature on time travel, I argued that Rachel's failures in any Grim Reaper setup are problematically inexplicable. I also hope to have, by example, shown that the literatures on causal finitism and time travel are more intimately connected than one might initially think.

Declarations

Competing interests All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

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/.

References

Effingham, N. (2020). *Time travel: Probability and impossibility*. Oxford University Press. https://doi.org/10.1093/oso/9780198842507.001.0001.

Erasmus, J., & Luna, L. (2020). A philosophical argument for the beginning of time. *Prolegomena*, *19*(2), 161–176. https://doi.org/10.26362/20200201.

Koons, R. (2014). A new kalam argument: Revenge of the grim reaper. *Nous*, *48*(2), 256–267. https://doi.org/10.1111/j.1468-0068.2012.00858.x.

Lewis, D. (1976). The paradoxes of time travel. *American Philosophical Quarterly*, *13*(2), 145–152.

Loewenstein, Y. (2022). Against the standard solution to the grandfather paradox. *Synthese*, *200*(2), 172. https://doi.org/10.1007/s11229-022-03543-y.

Pruss, A. R. (2018). *Infinity, causation, and paradox*. Oxford University Press.

Shackel, N. (2005). The form of the benardete dichotomy. *British Journal for the Philosophy of Science*, *56*(1), 397–417. https://doi.org/10.1093/bjps/axi121.

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