Article

Self-Fulfilling Prophecies

Stephanie Rennick

Department of Philosophy, University of Glasgow, Glasgow G12 8QQ, UK; stephanie.rennick@glasgow.ac.uk

Abstract: Causal loops are a recurring feature in the philosophy of time travel, where it is generally agreed that they are logically possible but may come with a theoretical cost. This paper introduces an unfamiliar set of causal loop cases involving knowledge or beliefs about the future: self-fulfilling prophecy loops (SFP loops). I show how and when such loops arise and consider their relationship to more familiar causal loops.

Keywords: causal loops; self-fulfilling prophecies; time travel; foreknowledge; coincidence; inexplicability

> *Joakim stands in front of his wardrobe, indecisively. In a flash of inspiration, he selects the black kimono. After all, he ponders, he already knows that's what he'll wear: the Delphic Oracle texted him as much that morning. Saved the mental labour of choosing an outfit, his mind wanders: 'she knew, because I'm wearing it; but I'm wearing it, because she knew that I would . . . '*

Causal loops are a recurring feature in the philosophy of time travel literature, where they are commonly divided into two varieties: object loops and information loops [1–3]. Here I discuss a different set of causal loop cases that involve knowledge or beliefs about the future: loops that arise in a subset of what I call self-fulfilling prophecies (SFPs)[1]. Although SFPs are a popular trope in the folk canon, they have yet to be the focus of detailed philosophical treatment. This paper serves as a first map of the conceptual terrain of SFPs, identifies their structure and features, and considers their relationship to more familiar causal loops.

I start by introducing causal loops more generally (§1) and then turn to introducing SFPs (§2). The rest of the paper deals with potential worries, including the apparent inexplicability (§3) and improbability (§4) of SFP loops. I will argue that neither worry is insurmountable, and that we should consider at least some SFPs even less problematic than the causal loops we are more accustomed to.

1. Causal Loops

A causal loop is a chain of events where each "is one of the causes of the next event, and whose last event . . . is one of the causes of the first event" [2] (p. 259). Events in a loop need not be complete causes or effects of one another, but may be:

> In a causal loop, the arrows of causation go around in a circle, but there might be additional arrows that lead into the circle, or arrows that lead out of it. If there are no such branches then the loop is said to be *causally isolated* [2] (p. 259).

Causal loops crop up frequently in the time travel literature, most famously described by Lewis as,

> Closed causal chains in which some of the causal links are normal in direction and others are reversed . . . Each event on the loop has a causal explanation, being caused by events elsewhere on the loop. That is not to say that the loop as a whole is caused or explicable. It may not be [1] (pp. 148–149).

A particularly memorable example comes from French and Brown: an archaeologist travels several millennia into the past in an attempt to discover the origins of a recently unearthed human skeleton, only to die and become that skeleton [4] (p. 208). This scenario involves a causal loop—discovering the skeleton results in the archaeologist's time travel, which results in his death, decomposition, and eventual discovery by the archaeologist. An important element of causal loops is the (weak) predestination they entail: the fact that the archaeologist discovers the skeleton-that-is-himself seems to entail that he will travel in time (and die, and decompose etc.). It is true[2] that the archaeologist, upon finding the skeleton, will go on to travel in time. Note that this isn't a fatalistic conclusion—the archaeologist did not *have* to time travel, but the fact he does results in the skeleton's existence (and importantly, identity: if he was not going to travel in time, he still might have found a skeleton, just one belonging to somebody else.)[3]

While some have disputed the possibility of causal loops depending on their accounts of time and causation[4], most philosophers of time travel think them at most weird [2,3,13,14]. As Lewis remarks,

> Strange! But not impossible, and not too different from inexplicabilities we are already inured to [1] (p. 149).

However, time travel is not the only context in which causal loops arise. Meyer notes a second:

> The other cases involve models of the general theory of relativity—first discussed by Kurt Gödel (1949)—that possess closed timelike curves in which time itself loops along a particular wordline. In such models, there is no backwards causation and travelling back in time requires no particular effort; one just has to follow an appropriately chosen worldline [2] (p. 259).

Skrzypek adds a third:

> [Loops] can be constructed by granting the existence of some causal agent existing in eternity, something or someone that has equal and simultaneous access to events at several times. In such cases, either the causal efficacy of the later event could "run through" the causal efficacy of the eternal being to the earlier event, or the causal efficacy of the eternal being could "run through" the causal efficacy of the later event and back through eternity to the earlier event [15].

The example he focusses on is the (later) impact of Jesus's life, passion and resurrection on the (earlier) immaculate conception of Mary.

The loops this paper focuses on cross-cut those identified by Meyer and Skrzypek: although many of the cases I discuss are time travel stories involving backwards causation, not all need be: foreknowledge-generating loops could also arise via backwards causation without time travel, or with forwards-causation via closed time-like curves (CTCs). Some philosophers of religion have worried that causal loops might result from a particular kind of divine foreknowledge called *simple foreknowledge* [16–18]. Under this account, God's witnessing the future is akin to us peering out a window: he has immediate, direct access to truths about the future[5]. Suppose that God today knows one of his faithful followers (FF) will die on Wednesday, so decides to warn him on Monday that he should tidy up his affairs. As a result of the warning, the follower gets very panicky, culminating in a heart attack on Wednesday (see Figure 1).
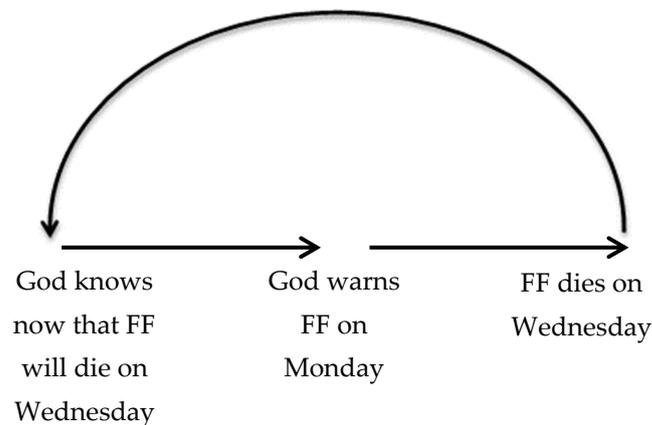
Figure 1. Divine foreknowledge in a causal loop (arrows denote causation).

Interestingly, considerable philosophical labour has been expended trying to prevent a connection between divine foreknowledge and causal loops, by—for instance—limiting God's knowledge to certain facts or certain times ('bracketing' God's knowledge) or making it the case that God only has knowledge after certain choices have been made. The worry centres around what Hunt calls 'the metaphysical principle' (MP):

MP: It is impossible that a decision depend on a belief which depends on a future event which depends on the original decision [16] (p. 486)[6].

MP-violating scenarios are not only causal loops, they are SFPs, and so they are loops of the particular variety we are interested in here. But as will become clear, you do not need God to get them.

2. Self-Fulfilling Prophecies (SFPs)

SFPs occur when knowledge or awareness of the future—perhaps by prediction, testimony, revelation or observation—is a crucial factor in future events occurring as the 'prophecy' describes. In a literal example of 'prophecy', Oedipus kills his father and marries his mother as a result of actions to avoid fulfilling a prophecy in which it is foretold that he will kill his father and marry his mother [19][7]. He knows the content of the prophecy and by trying to thwart it inadvertently ensures its veracity. SFPs are particularly vivid examples of how foreknowledge can impact the causal chain of events the knowledge describes: had he not heard the prophecy, it seems unlikely that Oedipus would have behaved as he did. The conundrum resulting from the feedback of foreknowledge on present action is vividly depicted in the following interaction from *The Matrix:*

The Oracle: [. . .] And don't worry about the vase.

Neo: What vase?

[Neo knocks a vase to the floor]

The Oracle: That vase.

Neo: I'm sorry.

The Oracle: I said don't worry about it. I'll get one of my kids to fix it.

Neo: How did you know?

The Oracle: What's really going to bake your noodle later on is, would you still have broken it if I hadn't said anything [21].

(Note that SFPs need not involve agents actively trying to *thwart* the events in question, although in examples from popular culture they often do; more on this below.)

Some SFPs are causal loops, and some are not. Suppose I say to you, "I know you will wear purple tomorrow", and then you wear purple because I said you would. By the definition given above, this is an SFP. If, however, I have just pulled 'purple' from the aether, then even if it turns out to be true that you wear purple, there is not a loop here[8]. There are

lots of examples like this in "The Fortuneteller": the town's fortune teller, Aunt Wu, makes a series of predictions that turn out to be correct, although the villagers fail to recognise that her accuracy is the result of their actions, rather than her having actual foreknowledge. For instance, Wu tells an old man that on the day he meets his true love he will be wearing red shoes, and as a result, he dons red shoes every day [22]. This is a standard causal chain: the man's actions result from what he perceives to be foreknowledge, but they do not influence Wu's prophecy (Figure 2):

Wu predicts red shoes $\longrightarrow$ Man dons red shoes

Figure 2. A Simple Causal Chain.

(Note that had he not believed the prophecy—had he not given it sufficient credence—it would not have been self-fulfilling; more on this below).

We can also have non-loopy scenarios with the causal chain running in the opposite direction if agents fail to act on the basis of their foreknowledge (these aren't self-fulfilling). Recall the case of the faithful follower (FF) from §1, where God knew the FF would die, warned him in advance, and thereby caused his heart attack and ultimate demise. Now suppose that God decides not to intervene: the death of FF causes God to know that FF dies, but the causal connection is one-way (as shown in Figure 3):

God knows today that FF will die on Wednesday $\longleftarrow$ FF dies on Wednesday
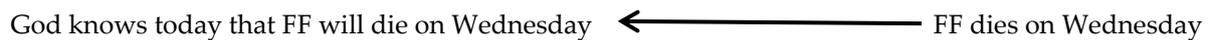
Figure 3. Divine foreknowledge without loop.

Sometimes, though, SFPs form causal loops. For instance, in Garth Nix's *Old Kingdom* Trilogy, the prophetic Clayr, in the present, have visions of inducting new members; then in the future, they induct those members based on the previous vision [23] (p. 15f). But they only have the vision because the new members will be inducted (as depicted in Figure 4):


Clayr have vision $\longrightarrow$ New members inducted
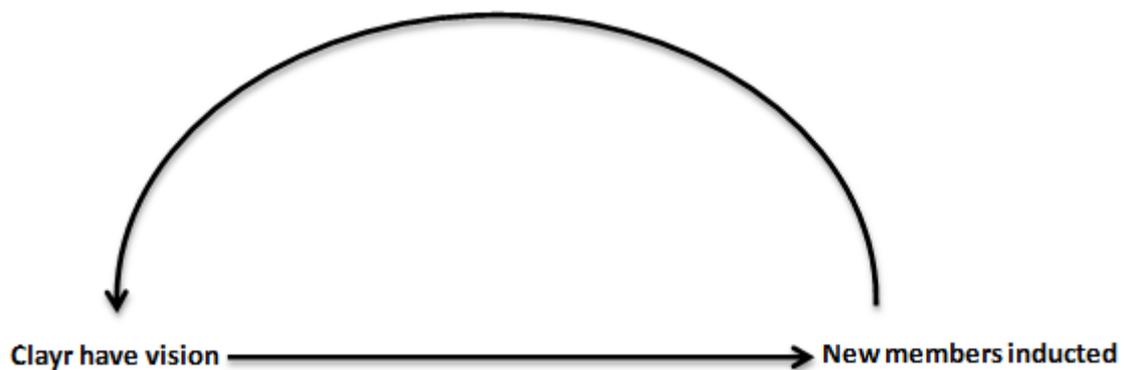
Figure 4. A Causal Loop[9].

Despite the familiarity of these scenarios from popular culture, mythology and religion, foreknowledge-generating causal loops (i.e., SFP loops) have not received detailed treatment in the philosophical literature. This might be surprising; indeed, three names might come to mind as potential counterexamples—Lewis, Goldman and Hanley. However, while Lewis briefly discusses both loops and fatalism at different points in "The Paradoxes of Time Travel" [1], that foreknowledge might result in such loops is never suggested[10]. Goldman on the other hand does provide an excellent example of a foreknowledge-generating loop, as described below, but his agenda is entirely different: to rebut what he calls 'anti-predictionism', i.e., an objection to determinism based on the impossibility of predicting voluntary actions[11]. Hanley comes the closest: he discusses loops involving intentional ac-

tion and the coincidences they might involve, and I will refer to his arguments throughout. However, it is worth noting that the role of belief in his cases goes unmentioned.

Goldman's Book of Life thought experiment goes as follows:

> While browsing through the library one day, I noticed an old dusty tome, quite large, entitled "Alvin I. Goldman". I take it from the shelf and start reading. In great detail, it describes my life as a little boy. It always gibes with my memory and sometimes even revives my memory of forgotten events. I realise that this purports to be a book of my life . . . I look at the clock and see that it is 3:03 . . . I turn now to the entry for 3:03. It reads: *"he is reading me. He is reading me. He is reading me* [26] (p. 144).

Like the Clayr case, this scenario involves a causal loop: at t_0, Goldman reads a page in the book of life that describes events taking place at t_1. At t_1, after the reading, Goldman performs an action φ. Suppose that the author of the book gains knowledge of Goldman's φ–ing (there are various ways to conceive of this, from time traveller to fortune teller) as a result of Goldman's φ–ing. That is, the book records that Goldman φs because he φs. But Goldman may, at least in part, φ as a result of this knowledge (or in defiance of the knowledge if it is a thwarting case[12]), and thus the knowledge is a contributing factor in the causal chain of events. To make this clearer, consider the following:

> I now turn to the entry for 3:28. It reads, "He is leaving the library, on his way to the President's office." Good heavens, I say to myself, I had completely forgotten about my appointment with the President of the University at 3:30 . . . Since I do have a few minutes, however, I turn back to the entry for 3:22. Sure enough, it says that my reading the 3:28 entry has reminded me about the appointment [26] (p. 144).

Why did he leave the library at that time? Because it was written he would. Why was it written? Because he left at that time. Like the archaeologist case, there is (weak) predestination at play here: the fact that Goldman reads a book which truthfully describes his future actions entails that he performs such actions. It is true at 3:22 pm that Goldman will leave the library at 3:28 pm. He did not have to, but the fact that he does makes the book true (and report that detail, instead of a variation).

However, not all of the events mentioned in the Book of Life case are loopy: neither those that occurred before Goldman started reading, for instance, nor any that might have occurred due to Goldman's misreading (or misremembering) of the text. In those cases, Goldman's actions are either causally unrelated to the prediction, or themselves the cause of the predictions (via backwards causation, a CTC or similar). And the claim above that 'Goldman's action makes the book report that detail' is predicated on the assumption that the writing in the book came about as a result of Goldman's actions: as if a fortune teller, God, or time traveller witnessed Goldman performing said actions and recorded them accordingly. But if this is not the case, then the book causes Goldman to remember his appointment, but his remembering his appointment has no bearing on the book's contents[13].

As Goldman's Book of Life and the Clayr examples demonstrate, in some cases knowledge of the future can influence actions in the present: the knowledge bears on the causal chain of events it describes. Sometimes (but, as demonstrated, not always) this leads to a causal loop[14].

However, as I have alluded to throughout, while many SFP scenarios do involve foreknowledge (at least under standard accounts of knowledge), we need not meet the bar of 'knowledge' for such a loop to arise [28] (p. 63). What SFP cases have in common is that awareness of (and belief in) the content of the prophecy leads to its coming true. In the story of Oedipus, for example, we are led to believe that the protagonist would never have killed his father and married his mother if he had not heard and believed the prophecy in the first place: it is the prophecy that provides the impetus, and serves as a catalyst, for the events that follow[15].

In most cases, at least one character in an SFP scenario (either the prophet or the person to whom the prophecy pertains[16] has a true belief about the future. For instance, the Clayr have a true belief that they will induct certain recruits into their ranks. But even this is not required; one could imagine a scenario where the predictor has access to and reports a truth about the future without believing it, where the subject of the prediction has a false belief about the future, and an SFP (even a loopy SFP) still arises. For instance (I call this case "A Comedy of Cellars"):

> Julia peers into her crystal ball and witnesses a future so surprising that she can't believe it to be the case; nonetheless she's a diligent sort and records her vision in her diary. Her assistant, Sue, glances in the diary at the end of the day, but isn't wearing her glasses so misreads Julia's handwriting, forming the false belief that her twin sister Prue will end up trapped in the cellar. Sue avoids the cellar as she is deathly afraid of the dark, but ventures down to save her more adventurous sibling. In her haste she forgets the keys, thereby trapping herself in the cellar and proving true the vision that Sue—not Prue—would be stuck in a place Julia would never expect her to tread.

In this case and other SFP cases, the belief (albeit false here) plays a significant role in events transpiring as they do. Nonetheless, we might think that cases such as this one are so improbable as to occur only very rarely. I consider the coincidences involved in SFP loops in §4. First though, I address another worry. One might wonder, in this case and the others discussed, *why* the sequence of events comes about: whether we can adequately explain where the knowledge, belief or decision comes from. This is a common worry about certain causal loops, which I turn to now.

3. Inexplicable Loops

There are several (interrelated) ways in which the charge of inexplicability is levied against causal loops. For instance, one of the main objections in the literature on the simple foreknowledge case outlined in §1—where God knows his faithful follower will have a heart attack and, in telling him, brings about the events and thus his knowledge—is that such loops serve as a bad explanation for why events occur [18]. In the time travel literature, it has been observed that while each event in a loop has a causal explanation, as it is caused by other events in the series, the loop as a whole may have no cause and thus be inexplicable. Finally, the content of at least some causal loops is argued to arise *ex nihilo*, and this is also claimed to be inexplicable. I find myself unsure as to whether SFP loops fall into this latter category, so instead will here make a disjunctive claim: if they do not, then SFP loops are even less troublesome than some other kinds of causal loops. However, even if they do, the time travel literature gives us good reason to resist that inexplicability worry, along with those levied against causal loops more generally.

An important feature of some, but not all, causal loops is the lack of origin of their content. Take, for instance, the main character in Robert Heinlein's *"-All You Zombies–"*, who is—thanks to time travel and a mid-life sex change—his own mother, father, daughter and son [31][17]. Their genetic information is caught up in a closed causal loop, with no apparent origin: it comes 'from nowhere'. Not just information can be loopy in this way, objects can too; imagine a time traveller who goes back in time and gives his younger self plans to build a time machine. The younger self grows up, builds the time machine, and goes back to give himself the plans[18]. As with Heinlein's character, the quandary lies in the blueprints' origin: where did they come from in the first place? So as not to confuse these with other types of causal loop, I call them CEN loops: loops that involve information or objects created (or appearing) *ex nihilo*[19]. In time travel scenarios, all sorts of objects, information, and even people, can have a loopy causal origin[20].

It is generally agreed in the literature that CEN loops are at least logically possible, although for other reasons they might be unpalatable; as Meyer notes, "loops are widely thought to constitute a theoretical cost of any view that permits them" [2] (p. 260).

At least some SFPs look to be CEN loops: the content of the Clayr's knowledge might seem to come 'from nowhere', and likewise for Goldman's decision to leave the library. Here's another example to consider:

> Billy is a contestant on a game show with very similar mechanics to the Newcomb Problem: he has a choice between one box and two boxes, and a highly-accurate predictor will predict his choice in advance. However, this particular predictor has access to the future (by time travel, crystal ball or a Gödelian telescope), and this is why she is so accurate—she witnesses the future choice, and thereby knows what Billy will choose. Billy is a stalwart two-boxer: in every conversation with friends and colleagues prior to the game he has insisted he will pick both boxes; he dreams at night of picking both boxes etc. However, the tables are turned when the predictor reveals her prediction to him prior to his choice: she says he will pick one box. As a result of this revelation, Billy decides to pick one box—after all, he reasons to himself, he now knows he will (see Figure 5)[21].

Where does the content of the knowledge come from? Well, the predictor knows that Billy will pick one box, because Billy picked one box. But Billy picked one box, because the predictor knew he would pick one box. If you do not like the word 'know' here, replace with 'believe': the predictor believed Billy would pick one box, as she witnessed Billy pick one box. But Billy picked one box, because she believed he would (or more long-windedly, Billy picked one box because he believed that she believed he would pick one box). Each link on the chain is explicable in terms of the previous link, but the loop as a whole, and the information contained in the foreknowledge (i.e., the outcome of Billy's decision), appears not to be.

In his "Paradoxes of Time Travel", Lewis concedes that CEN loops are inexplicable but remains unperturbed, describing their possibility as "remarkable" and arguing that information arising from nothing is no different than the many other inexplicable phenomena we manage to accept, such as the "decay of the tritium atom" [1] (p. 149). Nonetheless, he considers loops with information for content—like our SFPs might be—especially remarkable, asking

> Where did the information come from in the first place? Why did the whole affair happen?

And concluding,

> There is simply no answer [1] (p. 149).

Hanley disagrees, arguing that the question, 'Where did it come from in the first place?' is malformed; it is unanswerable, but only because there is no first place to talk about on a loop. Instead, he suggests, "the well-formed question 'Where did the information come from?' has a straightforward answer: from itself, by completely ordinary causal means" [3] (p. 137).

Perhaps the best articulation of why the bizarreness of CEN loops is merely apparent is found in Levin, who argues that questions about the origin of objects or information caught in a loop are

> No different from questions about where anything originally came from. We can ask about the origin of the atoms that make up [the time traveller]; their timeline is not neatly presented to us. The atoms either go back endlessly, or if the universe is finite, they just start. In either case the question of ultimate origin is as unanswerable as the question of the [loop contents'] origin. What makes us think that when such questions are asked about the loop they are different and ought to be answerable is that the entire loop is open to inspection. Sub specie aeternitatis this difference disappears [37] (p. 70)[22].

That is, we do not expect to be able to explain the causal history of the atoms that make up the objects surrounding us as they stretch back so far in time (and perhaps endlessly).

By contrast, in Billy's case, we have the entire causal history of the information (or decision) open to us. Thus we expect to be able to determine how or why it came about[23].

In terms of explicability, we can say the same things about loopy SFPs as has been said in the time travel literature. Each event in the loop comes about (at least partly) as a result of its predecessor. Each event is as explicable as events in a linear causal chain, with the added benefit that there are no uncaused first causes. There may not be an explanation for why the loop as a whole came about, but that is not unique to causal loops: if it is a problem, it is one the universe, God, and—as Levin notes—atoms, face as well.
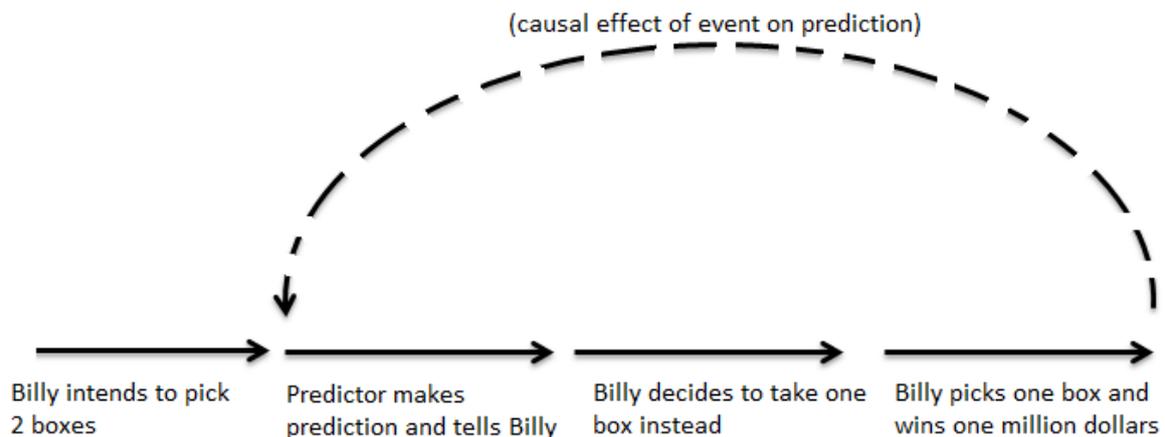

Figure 5. A candidate CEN loop.

4. Improbable Loops

> The really puzzling question is, 'Why does the information work?' . . . It's no
> mystery how the information works . . . The interesting question is why this is so.
> And the only possible answer is: coincidence [3] (p. 138).

The logical possibility of causal loops has been established decisively elsewhere, but one might still be concerned that cases like those I have introduced are so improbable that they nonetheless impose a significant theoretical cost. The kinds of coincidences involved in causal loops—and CEN loops in particular—have been discussed in various places in the time travel literature, and it is generally agreed that how improbable they are depends at least in part on the content of a given loop [3,14,24]. I will not replicate this discussion here, but I do want to say a little bit about why SFP loops, even if they involve content arising *ex nihilo*, might be less improbable or only require more ordinary coincidences than some other loops.

There are three reasons for this. Firstly, SFP loops are not object loops (but rather a subset of, or at least more akin to, information loops, to follow Lewis and Hanley's dichotomy). Objects age, they experience wear-and-tear, and if caught in a loop they must reverse age at some point so as to end up—atom-for-atom—as they started (Hanley calls this the 'restoration problem') [3] (p. 131f). SFPs are not subject to this problem, and thus do not require coincidences to surmount it. Secondly, the kind of *ex nihilo* content in the SFP cases presented above is in general simpler than that featured in the more extreme time travel loops; there are exceptions to this, and nothing relies on it, but it seems at least plausible that the kind of information we might come to know (e.g., that I will wear a red dress next Friday) is less complex than—for instance—the complete genetic information that makes up a person[24].

Finally, and most significantly, recall that one of the ways SFPs can be self-fulfilling is due to complying agents. Take the case of the Clayr (depicted in §2 Figure 4): the Clayr have a vision of the future in which they induct new members into their ranks. When the foretold day arrives, they induct these members. Given their visions are usually reliable, there is nothing coincidental about them seeing the new members being inducted. Likewise,

there is nothing coincidental about them inducting those members: they did so on the basis of the vision. At most, as Hanley notes of an analogous time travel case, "it is coincidental that they happen to know what they need to know, in order to do what they do", but "the sorts of coincidence involved are completely ordinary" [3] (p. 136).

Certainly there are SFP cases in fiction, or that we could imagine, that involve a great number of coincidences: those, for instance, where the agent hears a prophecy and wishes to foil it, only for circumstance to conspire such that they end up bringing about what they wanted to avoid[25]. "A Comedy of Cellars" in §2 might be one such example. Just as when we fail in more mundane tasks—like trying to get to work on time—some of the coincidences that scupper us will be ordinary (a run of red traffic lights, an unscheduled call from an overly talkative relative) and some may be more improbable (a monkey escaping from the local zoo and running off with our keys). Even if this type of case—with an agent trying to avert a true prophecy—does involve highly improbable coincidences, we have no reason to believe that it would be representative of SFP scenarios[26]. Indeed, if the coincidences required in foiling-but-self-fulfilling-anyway cases are especially unlikely or more numerous as opposed to the compliance cases, it is likely that the latter would comprise the bulk of SFP loops[27]. If the Clayr hadn't (a) willingly complied with the vision, then either (b) wacky hijinks would have ensued such that they complied inadvertently (as in so many fictional cases) or (c) the vision would have proven false. Neither (a) nor (c) requires any surprising or improbable coincidences, and of those only (a) results in a loop. So, even if (b)-type cases are improbable and thus unlikely to occur, this has little bearing on the probability of SFP loops more generally. Hanley observes that intentionality could reduce the improbability of causal loops: "the existence of agency may be the very thing that permits causal loops to obtain" [3] (p. 148). Cases where agents aim to bring about what they know will occur are prime examples of this.

5. Final Thoughts

Scenarios where agents have future-directed knowledge or beliefs that come to affect their present action are interesting and until now, underexplored, sources of causal loops. At worst, SFP loops are as inexplicable and improbable as the more familiar causal loops appearing in the time travel literature, but I suspect that in fact they require less to get off the ground.

I want to finish with one last question: do SFPs only arise in cases where agents know (or have beliefs about) their own futures? The short answer is no. What follows is the longer answer.

It is helpful to differentiate between what we might call 'first-person' and 'third-person' foreknowledge. (For ease I'll say 'foreknowledge', but throughout 'forebelief' can be substituted). By first-person foreknowledge I mean cases where someone knows their own future, rather than someone else's[28]. Most instances of third-person foreknowledge will not result in loops, as the foreknowledge will be merely the result of events occurring as they do, rather than a cause of the events—that is, the foreknowledge describes the events rather than (contributing to) bringing them about. For example, if a time traveller knows I will wear a red dress next Friday because she saw me wearing it in the future, then my wearing the dress causes her foreknowledge (in addition, presumably, to other causes). However, if I remain unaware of her knowledge, it seems unlikely that it would influence my choice of attire, thus there is no loop (see Figure 3 for an analogous case).

In all of the examples mentioned thus far, it is specifically the object of the prophecy coming to learn of it (or mistakenly thinking they have learned of it) and putting some stock in it that leads to its coming true. Given this, it is reasonable to wonder if SFPs are by their nature restricted to the first-person: if the foreknowledge must be possessed by the focus of the foretelling, rather than (or in addition to) a third party.

Although many SFPs do take this form—and indeed, they seem to be the most common and vivid in the folk canon—the foreknowledge need not *strictly* be first-person. For instance, in *Kung Fu Panda*, Grand Master Oogway has a vision in which the villain,

Tai Lung, escapes his prison. Tai Lung is not privy to the content of the prophecy. In an attempt to thwart the villain, Master Shifu sends a bird to the prison to increase security, thereby providing Tai Lung with the means of escape: a feather for a lock pick [39]. This is nonetheless an SFP as Shifu's foreknowledge plays a crucial role in bringing about the events that were foretold, even though the stated subject of the prophecy—Tai Lung—is unaware of this.

However, one might argue that all SFPs are implicitly first-person. Shifu is not aware of the role he plays in Tai Lung's escape because he does not have complete information. If a complete prophecy had existed describing the full set of events, then his actions would be contained therein.

In this paper I have focused on knowledge and belief, but it might well be that other future-oriented mental states can generate causal loops. Historically, the kinds of causal loop scenarios that have typically been discussed—especially those containing objects or information arising *ex nihilo*—have involved the kinds of things we do not ordinarily expect to come from nowhere: time machine blueprints, a working set of human DNA, and so on. But SFPs suggest a different direction: less unusual stuff—like beliefs—can generate causal loops[29]. Hanley points to intentionality as another source. It remains to be seen what else among our mental furniture might be suitably loopy.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

[1] It may be that these are best understood as a subset of information loops, but I won't assume as much. Either way, they have some features that aren't shared by information loops more generally, such as the role that belief plays in their occurrence.

[2] Or will become true, depending on your theory of time. For ease I will assume four-dimensionalism as per Lewis [1], but much of what I say could be adapted to other theories of time.

[3] Whether there are additional free will limitations with this kind of case, or backwards time travel more generally, is a bigger question than I have scope for in this paper. See for example [5].

[4] For example, [6] (pp. 123, 175–177). Cf. [7] (Chapter 12); [8] (Chapter 17); [9] (p. 4 fn. 1). Mellor's arguments (and similar) have been challenged in various places, including [10] (pp. 131–134); [11,12].

[5] As opposed to conditional or counterfactual knowledge, for example.

[6] It isn't logically impossible, but may well be theologically impossible (if, as has been suggested, it undermines the providential usefulness of divine foreknowledge). More interesting for my purposes is the claim that MP-violating circumstances are a bad explanation for why events occur as they do, cf. [18]. This is discussed further in §3.

[7] For a particularly explicit example see the conversation between Harry and Dumbledore on Voldemort's actions in [20] (pp. 740–741)

[8] Whether or not this would count as a case of fore*knowledge* as opposed to just forebelief (i.e., the belief that you will wear purple, formed as a result of my testimony) will depend on the connection between truth and belief that your epistemology requires. Either way, it isn't a loop, as the causal chain runs one-way from the belief today to your getting dressed tomorrow. If, however, I said you would wear purple because I'm a time traveller from the future where I saw you wearing purple, then it'd be a loop.

[9] This is simplified: we might expect intermediate steps such as 'Clayr decide who to induct' between their vision and the induction.

[10] Of course, time travel resulting in knowledge or beliefs about the future and the consequences of that in terms of free will, intentionality and resulting coincidences are discussed in several places, including [3,24,25]. However, that knowledge of (or beliefs about) future facts can generate causal loops (whether in tandem with or independently of time travel) was not the focus of these analyses.

[11] That the thought experiment he provides happens to be such a loop is grist to my mill, but undermines neither the novelty of this paper's conceptual mapping nor its conclusions.

[12] NB. These are directly analogous to bilking cases in the time travel literature (Cf. [25]).

[13] Likewise, if the author of the book was a supercomputer that could calculate the future based on deterministic laws and a complete understanding of the present (i.e., could have foreknowledge without backwards causation), this would not result in a causal loop.

14 To think that the relationship between foreknowledge and foreknown events always results in such a loop is what Craig calls "a misunderstanding in which the causal relation between an event or thing and its effect is conflated with the semantic relation between a true proposition and its corresponding state of affairs" [27] (p. 337).

15 Oedipus does not believe the prophecy is infallible, however, but rather that it will prove true *unless he acts to prevent it.* Another classic SFP case can be found in W. Somerset Maugham, "An Appointment in Samarra" from *Sheppey* (1933), as cited in [29] (p. 57): a merchant in Baghdad encounters Death and, based on what he perceives to be a threatening gesture, flees to Samarra to avoid his fate. Death notes that it wasn't a threatening gesture but "only a start of surprise. I was astonished to see him in Baghdad, for I had an appointment with him tonight in Samarra." If the merchant hadn't believed Death was out to get him, he'd have had no reason to go to Samarra. A nice parody occurs in [30] (pp. 77–78): Rincewind the wizard runs into Death, who comments that they have an appointment elsewhere soon and asks if Rincewind would mind going there. Rincewind declines.

16 These might be the same person, e.g., the Clayr.

17 Similar cases can be found in [14] and *Futurama.* Jane grows up in an orphanage; as a teenager she is seduced by a young man, falls pregnant and gives birth to a baby. Jane suffers trauma to her reproductive organs during labour, but doctors discover she is intersex and she undergoes sexual reassignment surgery. Now identifying as a man, Jane is taken back in time by a Bartender, where he meets and impregnates a young woman called Jane. The Bartender then recruits the young man to serve in the Temporal Bureau. The Bartender—revealed to be an older timeslice of the main character Jane—takes the baby back in time to an orphanage; he returns to the Bureau to contemplate his caesarean scar and the creation and recruitment of himself.

18 This is a version of an example in [1] (p. 149).

19 They are rarely given a name, but are mostly (including by Lewis) bundled together under the generic name 'causal loops'. Occasionally they are called 'ontological loops', a specific type of 'closed causal loop', or the 'ontological paradox'. Cf. [1] (p. 149).

20 See, for instance, [1,3,14]. There is an ongoing debate about whether aesthetic value comes from nothing with regards to an artwork caught in a (time travel) loop: see [32–34]. CEN loops are not limited to time travel scenarios, see for instance [35] (p. 58).

21 A similar line of thinking occurs in [36] (p. 301), with Harry's confidence in casting the Patronus: Harry finds the confidence to cast the difficult spell because he'd 'already done it', but thanks to time travel, the casting he (his earlier self) remembered and the casting he (his later self) performed were one and the same.

22 Alasdair Richmond raises the possibility that a Kantian noumenal self seems just as capable of 'willing' (atemporally) a closed causal chain as a linear one, which lends some credence to the idea that causal loops are no harder to explain sub specie aeternitatis than linear causal chains [38] (p. 102).

23 More recently, Meyer [2] (p. 260) considered the explicability of causal loops in light of two interpretations of the principle of sufficient reason.

24 As is the case in [31]. Divine foreknowledge may be a counterexample to this, if one thinks God is omniscient.

25 This is a common trope in fiction, with vague, misleading or deceptive prophecies. Examples can be found in *Macbeth* ("need fear none of woman borne"); *The Return of the King* (Eowyn and the Witch-king of Angmar); *Buffy the Vampire Slayer* ("Prophecy Girl"); *Mostly Harmless* (Arthur Dent's end); *Pirates of the Caribbean: On Stranger Tides* (Blackbeard's death).

26 Indeed, if Rennick [25] is right about the impact of foreknowledge on intention formation, these cases would be even rarer. See also [24] (§§6.1–6.3) on why agents might try to 'bilk' (change) the past.

27 There is an argument similar in spirit in [24] (see especially p. 377 re dates on objects).

28 The first-person/third-person distinction does not map on to the grammar employed. For instance, Oedipus, upon learning of the prophecy, has first-person foreknowledge—although when referring to 'Oedipus' we speak of him in the third person.

29 And it seems less uncanny that beliefs should arise from nowhere, if indeed that's what's happening in SFP cases.

References

1. Lewis, D. The Paradoxes of Time Travel. *Am. Philos. Q.* **1976**, *13*, 145–152.
2. Meyer, U. Explaining causal loops. *Analysis* **2012**, *72*, 259–264. [CrossRef]
3. Hanley, R. No End in Sight: Causal Loops in Philosophy, Physics and Fiction. *Synthese* **2004**, *141*, 123–152. [CrossRef]
4. French, P.A.; Brown, C. Time Travel. In *Puzzles, Paradoxes and Problems*; St Martin's Press: New York, NY, USA, 1987.
5. Wasserman, R. *Paradoxes of Time Travel*; OUP: Oxford, UK, 2017.
6. Mellor, D.H. *Real Time*; Cambridge University Press: Cambridge, UK, 1981.
7. Mellor, D.H. *Real Time II*; Routledge: London, UK; New York, NY, USA, 1998.
8. Mellor, D.H. *The Facts of Causation*; Routledge: London, UK; New York, NY, USA, 1995.
9. Berkovitz, J. On Chance in Causal Loops. *Mind* **2001**, *110*, 1–23. [CrossRef]
10. Bourne, C. *A Future for Presentism*; OUP: Oxford, UK, 2006.
11. Dowe, P. Causal Loops and the Independence of Causal Facts. *Philos. Sci.* **2001**, *68*, S89–S97. [CrossRef]
12. Riggs, P.J. A Critique of Mellor's Argument against 'Backwards' Causation. *BJPS* **1991**, *42*, 75–86. [CrossRef]
13. Ismael, J. Closed Causal Loops and the Bilking Argument. *Synthese* **2003**, *136*, 305–320. [CrossRef]
14. Macbeath, M. Who Was Dr. Who's Father? *Synthese* **1982**, *51*, 397–430. [CrossRef]

15. Skzrypek, J. Causal Time Loops and the Immaculate Conception. *J. Anal. Theol.* **2020**, *8*, 321–343.
16. Hunt, D.P. Providence, Foreknowledge, and Explanatory Loops: A Reply to Robinson. *Relig. Stud.* **2004**, *40*, 485–491. [CrossRef]
17. Robinson, M.D. Divine providence, simple foreknowledge, and the 'Metaphysical Principle'. *Relig. Stud.* **2004**, *40*, 471–483. [CrossRef]
18. Zimmerman, D. *The Providential Usefulness of 'Simple Foreknowledge'*; Rutgers University: New Brunswick, NJ, USA, 2010; Available online: http://fas-philosophy.rutgers.edu/zimmerman/Providence.Simple.Forek.6.pdf (accessed on 8 September 2021).
19. Sophocles. King Oedipus. In *The Theban Plays*; Watling, E.F., Ed.; Penguin Classics: London, UK, 1947; pp. 25–70.
20. Rowling, J.K. *Harry Potter and the Order of the Phoenix*; Bloomsbury: London, UK, 2003.
21. Wachowski, L.; Wachowski, A. The Matrix, Script. 1996. Available online: http://www.imsdb.com/scripts/Matrix,-The.html (accessed on 8 September 2021).
22. Ehasz, A.; O'Bryan, J. The Fortuneteller. In *Avatar: The Last Airbender*; Nickelodeon: New York, NY, USA, 2005.
23. Nix, G. *Lirael: Daughter of the Clayr*; HarperCollins: Sydney, Australia, 2001.
24. Smith, N.J.J. Bananas enough for time travel? *BJPS* **1997**, *48*, 363–389. [CrossRef]
25. Rennick, S. Things mere mortals *can* do, but philosophers can't. *Analysis* **2015**, *75*, 22–26. [CrossRef]
26. Goldman, A. Actions, Predictions and Books of Life. *Am. Philos. Q.* **1968**, *5*, 135–151.
27. Craig, W.L. Divine Foreknowledge and Newcomb's Paradox. *Philosophia* **1987**, *17*, 331–350. [CrossRef]
28. Effingham, N. *Time Travel: Probability and Impossibility*; OUP: Oxford, UK, 2020.
29. Dennett, D. True Believers: The Intentional Strategy and why it works. In *Mind Design II: Philosophy, Psychology, Artificial Intelligence*; Haugeland, J., Ed.; MIT: Cambridge, MA, USA, 1997.
30. Pratchett, T. *The Colour of Magic*; Transworld: London, UK, 1983.
31. Heinlein, R.A. —All You Zombies—. *The Magazine of Fantasy and Science Fiction*, March 1959.
32. McCall, S. An Insoluble Problem. *Analysis* **2010**, *70*, 647–648. [CrossRef]
33. Caddick Bourne, E.; Bourne, C. The Art of Time Travel: A Bigger Picture. *Manuscrito* **2017**, *40*, 281–287. [CrossRef]
34. McAllister, J.W. Does Artistic Value Pose a Special Problem for Time Travel Theories? *Br. J. Aesthet.* **2020**, *60*, 61–69. [CrossRef]
35. Gott, J.R., III; Li, L.-X. Can the universe create itself? *Phys. Rev.* **1998**, 58. [CrossRef]
36. Rowling, J.K. *Harry Potter and the Prisoner of Azkaban*; Bloomsbury: London, UK, 1999.
37. Levin, M.R. Swords' Points. *Analysis* **1980**, *40*, 69–70. [CrossRef]
38. Richmond, A. On behalf of spore gods. *Analysis* **2017**, *77*, 98–104. [CrossRef]
39. Aibel, J.; Berger, G. Kung Fu Panda, Script. Available online: http://www.imsdb.com/scripts/Kung-Fu-Panda.html (accessed on 14 September 2015).