

Don't Look Back in Anger. Freedom,
Fatalism and the Future

Giacomo Andreoletti

Contents

1	A Journey into Fatalism	5
1.1	What Fatalism amounts to	5
1.1.1	Three claims	5
1.1.2	Definitions	10
1.1.3	Why should we be worried?	16
1.2	Two classic Arguments for Modal Fatalism	19
1.2.1	A bit of Modal Logic	19
1.2.2	The Main Argument	21
	Prior Truths	26
	Ockhamism and the freedom to do otherwise	30
	Introducing a Third Truth Value	33
	The Mutable Future	41
1.2.3	Taylor's Argument	44
	The Argument	44
	Taylor's Argument Modal Structure	51
	Standard responses	52
1.3	Again on Definitions	57
1.3.1	Another bit of Modal Logic	57
1.3.2	Modal Fatalism Principles	59
1.3.3	What is wrong with the Main Argument and Taylor's	62
1.4	Why I am not worried	65
1.4.1	Lazy Argument	65
1.4.2	Deliberation	68
1.4.3	Moral Responsibility	70
2	A New Argument for Modal Fatalism	77
2.1	Two Principles which imply Modal Fatalism	77
2.2	The Disagreement on (SUF)	79

<i>CONTENTS</i>	2
2.3 What are we left with	84
3 Time Travel and Fatalism	86
3.1 Time Travel in One-Dimensional Time	86
3.2 Another Odd Story	95
3.3 The Specter of Fatalism with respect to Mit and Time Travel	98
3.4 The Specter fades away	100
3.5 Powers, Fatalism and Abilities	102
3.6 Protecting the Past	105
4 Changing the Past	110
4.1 An Attempt to Change the Past	111
4.2 Two Problems	114
4.3 Times and T-slices	116
4.4 The Past has not changed	121
5 Defending A Mutable Future	124
5.1 Free Will and the Future	125
5.2 Geach and the Changing Future	127
5.3 A case in favor of the Changing Future	130
5.4 Conclusion	132

Acknowledgements

On this occasion there is no one I wish to acknowledge.
Thank you¹.

¹Actually, I am profoundly in debt with several people and friends who had the patience to discuss and read the material of this work. I am sure those people — let me just mention my advisors Giuliano Torrengo and Achille Varzi — know how much I am thankful to them. Also, I am even more sure they have enough sense of humor to appreciate my acknowledgment.

Introduction

“Oh, if only I’d bought that stock! If only I, if only I purchased THAT house years ago! If only I’d made a move on THAT woman. If this, if that.” You know what? Give me a break with your could have’s and should have’s. Like my mother used to say, “If my grandmother had wheels, she’d be a trolley car.” My mother didn’t have wheels. She had varicose veins.

Whatever Works — Woody
Allen

We sometimes regret things we have done in the past. We may even despise ourselves for deeds we have done in the past. If I had not done that, I would not have ended up like this. All of us have plenty of examples which fall into this category. This kind of attitude toward the past might be overwhelming and lead to total despair. As a remedy, people are told that it’s no use crying over spilt milk. What is done is done, and cannot be undone. However, the disappointment remains. The thought that we could have done otherwise and act better than what we have done is still there. Fatalism, the topic I will be mostly dealing with in this work, offers a better medicine. There’s no point thinking that things could have gone otherwise. A fatalist would say that the way they went was the only way they could have gone, and this should be consoling. So, we are done with the past. But what about the present and the future? A fatalist,

very straightforwardly, claims that there is no difference between past and future. According to him or her, we should have the same attitude toward the future. There is only one way things could go. And here, if the fatalist is correct, we are again overwhelmed. How can we claim to be free to want or do anything? Are we just living a life which is already pre-determined? What is the point of living such a life? In what follows I will try to carefully address these questions. My answer will be that we should not be too much worried if the kind of fatalism I favor turns out to be true.

In the novel "Jacques the Fatalist" by Diderot, the protagonist claims all the time that what happens to him was already written on high. What this 'high' might consist of will be investigated throughout this work. In fact, we shall see that there are different kinds of fatalism and different sources that seem to point to fatalism.

Fatalism is often quickly dismissed as an untenable doctrine, which at best can count as sophistry. I don't agree with this judgment. One of my goals is to convince the reader that fatalism, at least one kind of it, has to be taken seriously and that it deserves respect and attention within our metaphysical investigations.

In chapter 1 we shall see what different kinds of fatalism there are and what is their nature. Then, we will focus on two classic arguments to the conclusion that fatalism holds. In chapter 2 I will give my own argument in favor of fatalism. In chapter 3 I will take time travel in one dimensional time as a case study. I will argue that issues which arise within the debate of the metaphysical possibility of time travel can be resolved by fatalism. In chapter 4 I try to refute a recent attempt which has been made to show that the past could change, even if time is one-dimensional. Ultimately, in the final chapter I will talk about a theory according to which the future might literally change. The theory provides two interesting ways to refute fatalism. I will provide an example which is supposed to favor the changing future.

I've been told that a dissertation should have a single main thesis. Mine has it, but it has the form of a disjunction. The received view is that the future cannot change and yet fatalism is false. I want to negate this conjunction. This denial gives us a disjunction. Either we should be fatalists — and it's not so bad after all — or we should believe that we can change the future. Hence, don't look back in anger. Go and change your future, it might work.

Chapter 1

A Journey into Fatalism

In this first chapter I introduce fatalism and its different understandings. I try to make it clear in the first section that there are at least two different kinds of fatalism, here labeled lazy and modal fatalism. The following section on modal logic is needed to understand the nature of modal fatalism. I then address two classic arguments to the conclusion that modal fatalism holds, and the standard responses which are found in the literature on this topic. I then provide a new reason that should incline us to think that those two arguments do not get through. Ultimately, I will try to show that although fatalism has always been seen as an untenable doctrine because of its consequences, this is not necessarily correct. Even though lazy fatalism does have unpleasant and seemingly fallacious implications, modal fatalism can perfectly address the problems which are usually associated to fatalism.

1.1 What Fatalism amounts to

I start this section by discussing three claims. These claims are going to help us with defining lazy and modal fatalism. Ultimately, I will show why fatalism has to be taken seriously and what consequences it might seem to have.

1.1.1 Three claims

We are about to start the discussion on fatalism by putting on the table three claims. The first two are admittedly vague, whereas the third is often employed as a definition of fatalism. These three claims will help us to

provide a distinction between different kinds of fatalism. Here are the three claims.

The Symmetry Thesis. There is a symmetry between the past and the future.

Causal Connectedness. What happens at a time causally depends on what happens at other times.

Powerlessness. Agents are powerless to do anything other than what they actually do.

Let's start with the first claim. What does it mean to say that the future is like the past? It probably depends on what we are talking about. One way to argue for the symmetry (or the asymmetry) of past and future is to put it in ontological terms. Suppose I now have to provide a full inventory of what exists. To start, it seems that I need to include all entities that exist at this very moment in time I am making the list. Should I include past entities, say Socrates, and future entities like the event of my grandson taking his first college class? Here there is disagreement. An eternalist would be willing to claim that it is reasonable to put in the inventory all past, present and future entities. He or she would probably argue for this in virtue of the fact that our best scientific theories, like special theory of relativity, treats the time-dimension in the same way the three space-dimensions are treated. It would be a non-sense to not include in the inventory an object spatially located in a country far away from where I am doing my list, just because it is spatially away from me. The same could be said for those entities located at times distinct from the time I am making my list. Another possible response is to claim that past entities do exist, whereas no entity which exists is such that it is also located in the future. This is the so called growing-block view, according to which everything which exists is either past or on the edge of the block. Or, we could say, as the presentist does, that all the entities which exist are located in the present time, the time when I am doing my list. Back to the symmetry/asymmetry distinction, presentism and eternalism treat past and future symmetrically with respect to ontology. According to eternalism we have both past and future entities in our inventory of the world. According to the presentist, neither of them are in our list. The growing-block theorist instead posits an ontological asymmetry between past and future. We have past entities in our inventory, and no entity in the list is such that it is future.

The symmetry/asymmetry issue can be addressed with respect to other aspects. One might say for instance that while the past is closed, the future is open. Take the first American Presidential election. It is an event which

now lies in the past. Almost everyone agrees that there is nothing we can do to change, alter or influence past events. For instance, we cannot now prevent that George Washington was elected, nor we can bring about that John Adams was elected. What is done cannot now be undone. Things might seem different with respect to future events. For instance, we can now influence, at least partially, the result of the next presidential election.

Another symmetry between past and future might come from the fact that certain Laws of Nature do not distinguish between past and future. Take Newtonian dynamics. All there is in a Newtonian picture of the universe is just particles. Particles have intrinsic properties like mass and charge, and the only dynamical property they have, i.e. the only property which changes with respect to time, is position in space. Newtonian dynamics dictates how the universe evolves, in the sense that it provides an algorithm to predict with certainty how an isolated system, or the entire universe, is going to be at another time, given how it is at the time we take to make our prediction. It turns out that the algorithm for inferring toward the future and the one for inferring toward the past are one and the same. That is, if you have a time T and another time T_1 at a certain temporal distance, there is just *one* way to predict with certainty the state of the system at T_1 given how the system at T is, together with the assumption that the time flows from T toward T_1 . The point is that it does not matter at all whether T_1 lies in the past or in the future with respect to T . This in turn implies that whatever can happen forward, can also happen backward. That is, if you have a movie of a Newtonian process, made up of a sequence of instantaneous frames¹, and you put all the frames in reverse order, the result is another movie which is as well in perfect accord with Newtonian Laws. Whether other systems of Laws of Nature fail to make a distinction between the past and the future is the subject of a wide debate we cannot address here.

The symmetry/asymmetry issue has a fundamental role in debates about metaphysics of time. We have strong intuitions that the future is different from the past, in the sense that the past is fixed, whereas the future is the realm of different alternative possibilities. But, are we entitled to trust this intuition? As we have seen, one may have reasons to treat past and future ontologically on a par. Or, one may posit the symmetry in virtue of the nature of Laws of Nature governing our world. Furthermore, once a decision is made on this issue in one sense or the other, what

¹Note that velocity can't be part of an instantaneous state, because velocity is the rate of change in position with respect to time. Hence, you need more than one instantaneous state to make sense of velocity.

consequences should we have? We live our lives toward the future. What if the past *is* like the future. Would this be a threat to our Free Will? What other consequences would we have?

Let us now turn to the Causal Connectedness thesis, according to which what happens at a time causally depends on what happens at other times. What causation amounts to, if there is such thing, again, is the subject of an enormous debate. For the purposes of what we are going to discuss, we can rely on Lewis's counterfactual account of causation. The idea there is that a cause is something that makes a difference in the course of events. A difference with respect to what? The counterfactual analysis responds by saying to what would have happened, had the cause not occurred. As Lewis (1973) puts ² it

Whether *e* occurs or not depends on whether *c* occurs or not.

The dependence consists in the truth of two counterfactuals.

$O(c) \Box \rightarrow O(e), \neg O(c) \Box \rightarrow \neg O(e)$.(p.563)

The truth conditions for counterfactuals used by Lewis (1973) are given as follows. Where *A* and *B* are propositions, $A \Box \rightarrow B$ — if *A* were true, *B* would also be true — is true at a possible world α if and only if either a) there isn't a possible world where *A* is true or b) there is a possible world β where *A* and *B* are true and such that it is more similar to α than any other possible world where *A* is true and *B* false. A major role is played by the notion of comparative similarity. This notion allows us to say things like the world β is more similar to α than the world γ is. Lewis takes this notion to be primitive. Judgments of similarity between worlds are based on a trade off between similarities of Laws of Nature governing the worlds, and similarities of facts within regions of spacetime of those worlds. A precise account of similarity can't be given, Lewis argues, and this is why he takes the relation to be a primitive notion. However, two constraints are imposed by Lewis on the relation of similarity among possible worlds. First, for any world α , α is more similar to itself than any other world. Second, the relation of similarity is such that it gives rise to a weak ordering where ties are permitted. The first constraint yields the result that if *A* and *B* are true at a world α , then the counterfactual $A \Box \rightarrow B$ is true at α . This holds in virtue of the clause b) of the truth conditions for counterfactuals adopted by Lewis. If *A* and *B* are true at α , necessarily there is a possible world in which *A* and *B*, α itself, which is more similar to itself than any world in which *A* is true and *B* false.

To give an example, I throw a stone toward a window at a time, and slightly later in time the window breaks. The event of me throwing the

²where $O(x)$ refers to the proposition that *x* occurs.

stone brought it about that the window breaks. Had I not thrown the stone, the window would have been unbroken. The previous counterfactual sentence appears to be true. A world where I do not throw the stone and the window breaks nonetheless must contain more differences, with respect to the actual world, than any world where I don't throw the stone and the window doesn't break. Therefore, the counterfactual is true at the actual world in virtue of the clause b^3 . Hence, the counterfactual analysis says, this counts as an episode of causation. What I did caused something else, my behavior at a time affected the nature of another time.

What about cases of Backward Causation, namely cases where the effect precedes in time the cause? Suppose time travel toward the past is possible. The time traveler could punch herself in the face before her trip, say in 2015, and then have a black eye after her arrival in the past. If so, the effect would precede in time the cause. Lewis's counterfactual analysis does not rule out cases of backward causation, neither does the Connectedness Thesis as it is formulated.

Our world is a world where it seems that causation constantly takes place. Events continuously cause, or bring about, or prevent other events. But this is not necessarily so, we can imagine scenarios where there are no such things. Suppose Tom is a marionettist playing with his puppets. There are events which happens in the "marionette-world". Arguably, the events of the marionettes are not causally related to one another. They are all causally dependent on what Tom does with the wires. Similarly, suppose there is an omnipotent God⁴. For all we *know*, it could be the case that what happens in the universe is always the result of God's intervention, like what happens in the marionettes 'world' is always the result of Tom's moves. In such case God would be the only entity with causal powers, and all which happens in the universe would be causally dependent on God's acts. Or, we can also imagine a possible world where there are no lawlike and causal regularities. Such a world would be a sequence of unrelated events where constant and abrupt changes would render a non-sense any talk about causation. An example could be a world where the position of all particles are represented by functions of the position with respect to time and such functions are discontinuous at all their points.

³The other relevant counterfactual, i.e. had I thrown the stone, the window would have been broke is true in virtue of the fact that the antecedent and the consequent are true at the actual world.

⁴We will make use of examples involving God just illustratively. This is not a theology dissertation. Even though fatalism has been discussed a lot within theology because of the worries it raises, fatalism is a metaphysical doctrine which can be discussed without bringing God into the picture

The Connectedness Thesis can be rendered in terms which are in accordance with Special Theory of Relativity. It would become something like the claim that what happens at a spacetime point may be causally dependent on what happens at another spacetime point, provided the two spacetime points are separated by a time-like (or light-like) interval, where two spacetime points are separated by a time-like (or light-like) interval if one could get from one point to the other by moving with a speed smaller than (or equal to) the speed of light.

To sum up, we take it as a platitude that in our world there are continuously episodes of causation. This thought is also a guidance in our lives. We act in certain ways because we know, or tend to think, that from certain causes, certain effects will follow. The Connectedness Thesis claims that a network of causation is at play in our world. Events in our world are not isolated from each other. Rather, they are intertwined in a causal network. Lewis's counterfactual account is a powerful way to make sense of this causal network. As we shall see later, there are kinds of fatalism which can rely on counterfactual claims in order to account for causation.

The Powerlessness claim says that no agent has ever the power to do otherwise than what he or she actually does. It is simple and complex at the same time. On the one hand, it is complex because it depends on what we mean by powers, and this may be a subject of great controversy. We'll have a great deal more to say about this later. On the other hand, it can easily be illustrated by examples. Suppose Holly has a microchip installed in her brain. The microchip controls her brain and dictates any action she performs. It seems plausible to say that Holly has never the power to do otherwise than what she actually does. The microchip forces her to raise her right hand at a given time, and then she does. Could have she done otherwise? The answer is negative, given that her conduct was determined by the microchip. What the Powerlessness claim does is to generalize this kind of situation to any agent. According to it, every agent is, *in some sense or another*, similar to Holly in this respect. A person's conduct is always such that the person has never the power to do otherwise than what he or she in fact does.

1.1.2 Definitions

With the help of three claims discussed above, we are ready to address the problem of what fatalism amounts to. It is a metaphysical doctrine, and as such it might be problematic to pin down an unique definition of it. Different authors have meant different things with the same word. We

will go through different understandings of fatalism.

We introduce the first understanding with a quote from Van Rensselaer Wilson (1955). He says:

The typical fatalist contends that human effort, human wisdom, human skill, even human stupidity, have no causal continuity with the future. The same future will occur, according to the fatalist, no matter what we human beings know or don't know, do or don't do, seek or shun. (p.70)

Another passage from Ryerson (2011) seems to attach the same thought to the doctrine of fatalism.

The fatalist contends, quite radically, that human actions and decisions have no influence on the future. (p.5)

These two passages can suggest a first characterization of fatalism.

Lazy Fatalism. What will happen will happen no matter what one does.

Surely, as Ryerson observes, this is a bold claim. Even more surely, this doctrine is clearly false, at least in the world we happen to inhabit. It is not the case that future events will happen no matter what we do. Suppose that I am poisoned and as a matter of fact I will die in twenty minutes. Suppose also that there is an antidote capable of healing me. The antidote is such that with certainty it will make me recover. Clearly, my future death in twenty minutes is not going to happen no matter what I do. On the contrary, my future death in twenty minutes depends on things I do now. For instance, my future death depends on me abstaining from taking the antidote now. I said that this doctrine is clearly false, at least in our world. It could be true in conceivable scenarios. Tom the marionettist is again playing with his puppets. This time he is following a script he wrote years ago. The marionettes play is well written. The marionettes act *as if* there were causal connection among them. On the script, marionette A punches marionette B, and then B falls down on the floor. Being in a bad shape that day, Tom is not playing very well with his wires. He makes a mistake when he is playing that part of the script. A doesn't actually punch B, because A misses B. B falls down nonetheless, cause Tom is more careful with that move of his wires. In such scenario it is reasonable to say that the event of B falling was going to happen no matter what. All actions are fated to happen in the sense given by Lazy Fatalism. For instance,

A's actions are causally irrelevant to future happenings, because the only things causally relevant are Tom's moves⁵.

So, regarding the three claims discussed above, Lazy Fatalism denies the Causal Connectedness thesis. It accepts the symmetry claim, where the symmetry amounts to the idea that present actions are causally irrelevant for both the past and the future. Actions, in this view, do not affect the past, and neither affect the future. Lazy Fatalism is silent about the third claim, in the sense that it is compatible both with the acceptance of the Powerlessness claim and with the denial of it. A case like the marionettes case is a case where the marionettes are incapable of doing other than what they actually do, because they are overwhelmed by a power they have no control over, namely Tom's moves. But, one might imagine a situation where there are agents capable of acting differently from how they actually act, and yet future events happen no matter what they do. An universe without any regularity outside agents would arguably be such a case. Suppose you ordinarily persist in a world where things outside you continuously change in an abrupt and totally unpredictable manner. In such case, you might have the power to do otherwise, yet it seems there would be no way to affect future happenings outside you.

Lazy Fatalism is clearly wrong because we take it to be the case that there is a causal network in our world. We do affect with our acts future events: I take a medicine and as a result I recover. I throw a stone and then the windows breaks. It is thus no wonder that no philosopher endorsed such a view. But why could Lazy Fatalism be, maybe just for a little while, considered possibly true in our world? Probably because it may sound similar to the widely accepted idea that the future cannot change. One might think that future events will happen no matter one does because the future cannot change. But this is clearly not the case. The fact that the future cannot change does not imply that there aren't causal connections among events. If the future effects are fixed and part of the future, so are their causes, provided the Connectedness thesis is true.

We shall move now to a more interesting account of fatalism. We start with a quote from Mackie (2003), where she says:

Generalizing, the fatalist concludes that one never has the ability to avoid doing anything than one actually does (p. 129)

Van Inwagen (1983) says something similar.

⁵And notice that here the counterfactual analysis of causation delivers the correct answer. In the described scenario, A does not punch B and B falls down as if it were punched by B. A missing B does not cause B fall, because it is not true that had A punched B, B would have not fallen.

Fatalism, as I shall use the term, is the thesis that it is a logical or conceptual truth that no one is able to act otherwise than he in fact does.(p.23)

The two quotes are very similar. We can express the content they share ⁶ with this second account of fatalism.

Modal Fatalism. For any agent, for any act, for any time, the agent has not within his or her power to act otherwise than he or she actually does.

This is a kind of fatalism which is shaped in terms of the concept of the power to do otherwise. I call it Modal Fatalism because it employs the notion of power, which as we shall see better later, can be treated as a modal expression. According to this account, if you acted in a certain way at a certain time, you couldn't have done otherwise, and this of course generalizes to all acts, agents and times. It should be noted that Modal Fatalism relies on modal discourse in terms of powers, whereas Lazy Fatalism focuses on distinct times and denies causal connections among events happening at those times.

Iacona (2007) offers this definition of fatalism

Fatalism is the doctrine according to which if something happens, it is necessary that it happens, and if something does not happen, it is impossible that it happens. (p. 45)

This definition boils down to Modal Fatalism, provided you restrict it to agents' acts. Saying that whatever happens is necessary means that when someone acted, he or she couldn't have done otherwise. Similarly, saying that whatever does not happen is impossible means that the acts not performed were not within the agent's power.

Some authors define fatalism in terms of Free Will. For instance, Merricks (2009) says:

In my opinion, the Main Argument is the strongest argument for fatalism; that is, it is the strongest argument that moves from truths in the past to a present or future lack of freedom. (p.88)

⁶Van Inwagen puts in the definition of fatalism that the lack of the power to do otherwise is a logical or conceptual truth. I don't understand what this specification means. Fatalism is a metaphysical doctrine. If it's true, reasonable, debatable and the like, it's just because it does or might follow from other concepts through logical reasoning, as any other metaphysical doctrine.

The Main Argument is a standard argument for fatalism; we will talk about it to a great extent later. What Merricks has in mind here, though, is the understanding of Free Will in terms of the concept of being able to do otherwise. What Free Will amounts to of course is the subject of a wide debate. It is true that traditionally the power to do otherwise has been considered constitutive of the notion of Free Will. But there are alternatives. One might cash out Free Will in terms of spontaneity. In this sense, one act would be free, if nothing external coerced it. Or one might, as Frankfurt does, say that one acts freely if the desires on which the act is based are those which are in turn desired. And of course, Free Will can be understood in many other different ways. For the rest of this work, we will stick to the notion of having the power to do otherwise. This is the notion modal fatalism targets, and it is the one which is denied in the Powerlessness claim.

How does modal fatalism relate to the three claims we have seen above? modal fatalism amounts to a claim that Powerlessness holds. It does then accept the Symmetry Thesis, in the sense that according to modal fatalism no one has ever had, has or will have the power to do otherwise. Past and future are alike in this respect. It is then silent with respect to the Causal Connectedness Thesis.

There seems to be another account of fatalism, cashed out in terms of unpreventability of future events. This account would claim that the doctrine of fatalism amounts to the idea that all future events are now unavoidable or unpreventable, like the past ones clearly are. Taylor (1963a) seems to suggest so in this passage:

We all, at certain moments of pain, threat, or bereavement, are apt to entertain the idea of fatalism, the thought that what is happening at a particular moment is unavoidable, that we are powerless to prevent it.(p.54)

Now, I think that arguably modal fatalism implies this kind of unpreventability. Take the case where I am poisoned and I die ten minutes by now. This clearly implies that no one will prevent the occurrence of my death, i.e. no one will perform an act such that it does prevent my death from happening. Now, according to modal fatalism, this means also that no one will have the power to perform an act such that it ensures my survival. And the way we normally understand something as unpreventable is in terms of powers. That the sun will rise tomorrow is unpreventable, in the sense that no one has the power to stop the sun from rising. Hence, modal fatalism yields this kind of unpreventability. But again, we must stress that this does not necessarily mean that there aren't causal connections.

Had I took the antidote, surely I would have survived. It is important to stress that modal fatalism *is* compatible with the Causal Connectedness thesis. Hence, I would suggest that we should avoid cashing out modal fatalism in terms of unpreventability. It is true that modal fatalism implies unpreventability in the sense I have just shown. Yet, unpreventability sort of intuitively conveys the wrong idea that nothing *could* prevent future events which will happen or cause those that which won't.

To wrap up this part, we have our three claims, which have been useful to introduce different kinds of fatalism. We can illustrate them via the following table.

Table 1.1: Kinds of fatalism

	Symmetry	Causal Connectedness	Powerlessness
Reasonable Fatalism	✓	✓	✓
Modal-Lazy Fatalism	✓	✗	✓
Lazy Fatalism + PTDO	✓	✗	✗

All kinds of fatalism accept a symmetry between past and future. As Taylor (1962) says:

A fatalist, in short, thinks of the future in the manner in which we all think of the past. (p.41)

Then, we have distinguished between two cases of fatalism, lazy and modal fatalism. Lazy fatalism is the thesis according to which things happen no matter what agents do, whereas modal fatalism claims that agents have never the power to do otherwise than what they actually do. The former is a denial of the Causal Connectedness thesis. The latter is the claim that Powerlessness holds. So, we have fatalism if either there is an ✗ under Causal Connectedness or a ✓ under Powerlessness. Hence, the case where there is a ✓ under Causal Connectedness and a ✗ under Powerlessness does not appear in the table because it would not be a case where we have fatalism.

Modal-lazy Fatalism combines the claim that there are not causal connections among events and the claim that agents are powerless to do other than what they in fact do. Take a case where an omnipotent God-like creature governs everything that happens, like Tom does with his marionettes. The creature would be the only entity with causal powers, and such scenario would amount to a case of modal-lazy fatalism.

The third case, lazy fatalism + power to do otherwise (PTDO), denies the causal network among events and allows agent to have the power to

do otherwise. A scenario without any kind of regularity or causal laws outside agents would be such a case. In such case, agents would be free in the sense captured by the notion of having the power to do otherwise. Yet, things outside them would happen no matter what they freely do.

The first case, modal fatalism combined with the existence of a causal network is the most interesting case. It denies to agents the power to do otherwise, yet in such fatalism agents can affect or bring about future events in virtue of the existence of a causal network. I label this 'reasonable fatalism'. Reasonable fatalism is a sub-case of modal fatalism. Modal fatalism by itself simply denies to agents the power to do otherwise. Reasonable fatalism also accepts the Causal Connectedness thesis. For most of the rest of this dissertation I will be focusing on modal fatalism and its sub-case reasonable fatalism.

1.1.3 Why should we be worried?

Why should we be worried about modal fatalism, i.e. the idea that agents have never the power to do otherwise? This question has two senses. The first one has to do with certain consequences we are inclined to think modal fatalism has. The second sense has to do with the thought that many metaphysical and scientific theories *seem* to imply modal fatalism as a consequence. According to the first sense, we should be worried about modal fatalism because it is traditionally considered a threat to our human lives. According to the second one, modal fatalism is an issue that has to be addressed carefully.

As for the first sense, many philosophers think that modal fatalism has undesirable consequences. For instance, it might seem that deliberation would be useless, if modal fatalism were a true doctrine. We think about deliberation in terms of choosing among different open alternatives, and according to fatalism there is always just one path open, the one we will actually embark ourselves in. And then there is moral responsibility. One might think that no one should ever be held responsible and deserve punishments and rewards, if modal fatalism were true, because with any given act we do, it is true that we can't do otherwise than what we do. And intuitively, when someone is blaming me for something I didn't do, it sounds as a really solid excuse to say that I couldn't do that, because it was not within my power. Similarly, if someone is blaming me for something I did, it sounds as a really good excuse to say that I could not have done otherwise. The problem is that modal fatalism seems to justify these two excuses to any act whatsoever. We shall have a great deal more to say

about these worries in section 1.4. I will argue there that even if these *seem* highly problematic consequences for modal fatalism, there are still ways to accommodate them without rejecting modal fatalism.

We turn now to the issue that modal fatalism seem to follow from various metaphysical and scientific theories. Are there reasons which seem to point out that modal fatalism might be true after all? I think that plenty of metaphysical doctrines and scientific theories seem to suggest that modal fatalism is true, in the sense that they cast doubts on our power to do otherwise, which is precisely what is denied by modal fatalism.

Does a block universe view entail modal fatalism? According to a block universe view, all times are ontologically on a par. Past, present and future times are all equally real. Talking of future and past is just a matter of a relative standpoint. But we are all fatalist with respect to the past, in the sense that we know there is nothing we can do about it. Should we also be fatalist with respect to the future?

What about 4-dimensionalism? According to 4-dimensionalism, persons and objects are four dimensional worms, extended through space and time. This means that my future temporal parts constitute, together with all the other ones, what I *am*. And my future temporal parts exist in the same manner my present part does — they are fully characterized and unalterable. How could I have the power to do other than what is done by my (now) future temporal parts?

What about the Bundle Theory about objects? According to the Bundle Theory, objects and persons, *are* just collections of properties. Suppose I am sitting at T and standing at T₁. How can we make sense of this change that occurs to *me*? If at T what defines me are just my properties, then as soon I lose one, it might seem that I cease to exist at T₁. One standard way to avoid this undesirable consequence is to claim that it is not the case that at T I am constituted by the property of being sitting. Rather, what constitutes me are the property of being-sitting-at-T and the property of being-standing-at-T₁. But again, if I *am* something which is constituted by the property of being-standing-at-T₁, at T, how could I refrain from standing at T₁?

What about super-essentialism, the idea that every property we have is essential? This is not a popular view, but it happens that philosophers like Leibniz endorsed such a view. If any property is essential, it seems that things cannot ever be otherwise. I raise my hand at a certain future time. This is part of my essence exactly like the fact that I am an human being. How could I possibly have the power to refrain from doing something which is essential to me?

What if the prevailing laws of nature are deterministic? If so, a single

instantaneous state of the world together with the laws of nature determine an unique past and an unique future. So, say that the state of the world ten seconds ago determines that I will raise my right hand ten seconds from now. How could I have the power to refrain from raising my hand? This event depends on a past state and on laws of nature, two things I don't have control over.

What if the prevailing laws of nature are indeterministic? Take quantum mechanics in the Copenhagen interpretation. Here things evolve in time according to the Schrodinger equation. And this dynamics provides perfectly deterministic and correct predictions about future times, and so we would be at the previous point. Indeterminism comes in just when we have measurements. Here what happens is that two outcomes are equally probable given the same initial conditions. Needless to say, it means that which outcome obtains is just a matter of chance. And again, it seems we can't have control over what is chancy.

What if the future totally depends on how the past was? Everyone believes that our personality, the environment we lived in, and our past experiences affect the way we act. What if they, maybe together with other factors we did not mention, *fully* determine the way we act? Those are things we don't have control over, because they all lie in the past. If so, it may be that we don't have control either over what those things entail, i.e. the way we act.

Some of our best scientific theories might also suggest modal fatalism. Take the special theory of relativity. According to it, an absolute present does not exist. What is the present, i.e. all the things simultaneous with a present time, is a matter of frame of reference. People in different frames of reference will disagree about what events are present. This means also that what is past and what is future is relative. Alice on her spaceship could pass by Bob, and they would disagree about pastness and futureness of the same event. The same event might be past for Alice and future for Bob. And the theory says that they are both correct, because they are in different frames of reference.

Of course, the theories we mentioned do not strictly imply the truth of modal fatalism. There are ways to resist a fatalistic conclusion, even though one accepts the theories we have seen above. But what I want to highlight is that modal fatalism is a doctrine we should seriously deal with, given that lots of different theories or doctrines seem to point to that conclusion. The other thing worth considering carefully is whether modal fatalism does really posit a threat to human agency and our moral lives. I shall argue that modal fatalism together with the Causal Connectedness thesis — reasonable fatalism — has not such bad implications as it is commonly

thought.

1.2 Two classic Arguments for Modal Fatalism

Having shown what modal fatalism amounts to, we now move forward and see how one could argue for a fatalistic conclusion. The two arguments we are about to analyze, the Main Argument and Taylor's argument, are arguments to the conclusion that no agent has ever the power to do otherwise than what he or she actually does. Hence, they are arguments for the Powerlessness Thesis and consequently for modal fatalism. The first one, The Main Argument as it is commonly called, traces back to Stoic philosophy and has been thoroughly discussed throughout all the history of philosophy. Everyone dealing with issues like truth and time is supposed to say something about the Main Argument. As a matter of fact, no one accepts the Main Argument. Yet, it is interesting to see how there are different strategies in the literature to resist it.

The second argument traces back to Aristotle with respect to its structure. I will analyze the, I think, improved version Taylor (1962) gave. Taylor's argument is intriguing and it is not clear how it can be refuted. Before going into the details of the two arguments, we shall deal with a bit of modal logic, because it will be needed to have a better grasp of the two arguments.

1.2.1 A bit of Modal Logic

In order to get a better understanding of the Main Argument, we need to say that the schema K is valid in all models of modal logic. This section is devoted to understanding what the previous sentence means. I will rely on Chellas (1980). First of all, here is the schema K.

$$\mathbf{K}. \Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$$

where A and B are meta-variables which range over sentences of the object language L . The object-language L is composed by all atomic sentences ($p, q, r \dots$), together with the non-atomic ones constructed out of atomic ones, standard connectives and modal operators, according to the usual recursive definitions. K^7 states that if necessarily A implies B , then if necessarily A , then necessarily B . Necessity distributes over a conditional.

⁷which is equivalent to $(\Box(A \rightarrow B) \wedge \Box A) \rightarrow \Box B$ in virtue of propositional logic.

A *schema* is a set of sentences which share the same form. By the schema K we want to refer to all sentences which have the form of a conditional, where in the antecedent of it there is a necessitation in front of a conditional with a first sentence in the language as antecedent and a second (or possibly the same) sentence as consequent, and in the consequent of it we have another conditional where the antecedent is the necessitation of the first sentence and the consequent is the necessitation of the second sentence.

For instance, if we assign to the metavariable A in K the value $\Box p$ and to the metavariable B the value $\neg q$, the following sentence is an instance of the schema K

$$(1.1) \quad \Box(\Box p \rightarrow \neg q) \rightarrow (\Box\Box p \rightarrow \Box\neg q)$$

Sentences may be true or false. In modal logic, truth and falsity are relative to models and possible worlds. That is, a sentence is true or false *at a possible world in a model*. A model is a structure $\langle W, R, P \rangle$, where W is a set of possible worlds, R is a binary relation on W and P is a function from sentences of the object-language to subsets of W . R is the accessibility relation which tells us what worlds are accessible to other worlds. We write $\alpha R \beta$ to mean that the world β is accessible (or possible) with respect to α . The main idea is that it is not necessarily the case that every world in W is accessible with respect to any other world in W . Different kinds of accessibility relations specify different models. P provides an assignment of truth values to atomic non-modal sentences. Technically, P is a function from atomic sentences of our language to a subset of W . Given an arbitrary sentence p , $P(p)$ gives you the set of all worlds in W where p is true.

Truth conditions for non-modal sentences are as usual, whereas truth conditions for modal sentences are the following.

$\Box A$ is true in the model M at the world α iff for every β such that $\alpha R \beta$, A is true at β in M .

$\Diamond A$ is true in the model M at the world α iff for some β such that $\alpha R \beta$, A is true at β in M .

We can now move on to the notion of validity, a concept which is built up in terms of truth at a world in a model. Validity applies to sentences, and it applies to schemes as well, given that we defined them as sets of sentences. There are different degrees of validity. We have the highest degree of validity when a sentence is true at every possible world in every model. For instance, tautological sentences enjoy this degree of validity.

No matter how the model is, the sentence is true in all possible world within the model. We can have other degrees of validity. For instance, a sentence is said to be valid *in a model* M if it is true in all the worlds in the set W of M . Or, a sentence may be valid in a *class of models*. Suppose we build up a set of models, say the set of all models where R is reflexive. A sentence is said to be valid in the class iff for every model in the class it is the case that the sentence is true at all worlds provided by the model. For instance, it is provable that any sentence of the form $A \rightarrow \Diamond A$ is valid in the class of all models where R is reflexive.

We are now in a position to prove that the schema K is valid in all models and enjoys the highest degree of validity, i.e. K is true at every possible world in every model. We give an informal proof of it. To do so, we take an arbitrary world α in an arbitrary model M and we make no assumption whatsoever on the nature of W , P and R , which specify the model M . We then assume that the antecedent of K is true at α in M . Now we need to show that the consequent of K is true at α in M . The consequent of K is a conditional, hence we proceed by assuming its antecedent $\Box A$ as true at α in M . We now reason as follows. We are not making any assumption on the nature of M . So, either α doesn't R any world, or α R s at least one world. In the first case, $\Box B$ is true at α in M , because in such case B is true at any world which is accessible from α . In the second case, there is at least one world which is seen by α . We call this world β and again, we make no assumptions whatsoever on its nature, besides the fact that $\alpha R \beta$. With the assumptions made thus far, necessarily β is a world where $(A \rightarrow B)$ and A are true in the model. It follows that B is true as well at β in M . Hence, we have the result that in any world accessible from α it is the case that B is true. In other words, $\Box B$ is true at α in M . Given that we made no assumptions on the nature of M and α , the result generalizes to all models and all worlds. Thus, we proved what we wished to prove. K is valid in all models of modal logic.

1.2.2 The Main Argument

To my knowledge, nobody accepts the Main Argument⁸. Yet, I think it is an argument worth considering, given its puzzling nature. The argument's strength is proved by the fact that there are several different strategies to resist it. The Main Argument moves from two plausible ideas; the first one is the fixity of the past, while the second one is the existence of prior truths. The fixity of the past may be expressed by the famous saying which claims

⁸Also known in the literature as the Master Argument.

that there's no point crying over spilt milk. If something belongs to the past, there is nothing we can do to change it. We could learn lessons from what happened in the past, but we have no power to prevent what already happened. It is often said that not even an all-powerful God could change what took place in the past.

Prior truths are truths which involves two distinct times. Let us take a case where at T_{-10} an irreversible process is going to cause an explosion in a nuclear plant within ten seconds. It is true at T_{-10} that the process begins. It is also true at T_{-10} that, given the unstoppable nature of the process, there would be an explosion at T_0 . The latter sentence expresses what we mean by a prior truth. It is a truth about a time, T_{-10} , which delivers a content belonging to another time, in this case T_0 . It is a necessary condition for making sense of prior truths that one accepts the notion of propositions having a truth value at times, rather than having it *simpliciter*. That is, we can, or maybe we have to, say that 'p is v at t' is a correct expression when we talk about truth, where 'p' refers to a proposition, 'v' refers to a truth value, and 't' refers to a time. In the case of the nuclear explosion, the prior truth amounts to say that the proposition that *there BE a nuclear explosion at T_0* is true at T_{-10} . I use italic, the infinitive and the reference to a specific time to refer to tenseless propositions, i.e. propositions which are not dependent on the context of evaluation, since they do not have any tensed constituent and come with a specific time in their content. One who accepts the notion of prior truths, has to accept that propositions have a truth value *at a time*, even if we are dealing with the kind of tenseless propositions just described. It is easy to see that there are two times involved in a prior truth, T_0 and T_{-10} in the example given. The former is embedded in the content of the proposition, while the latter is the time at which the proposition has the truth value.

Let us see now one ill-formed version of the Main Argument. We go through this ill-formed version in order to highlight the force of the correct version.

The ill-formed Main Argument

Suppose Jones raises his right hand at T. From this it follows that the proposition that *John RAISE his right hand at T* is true at T. If such proposition is true at T, then it is always true (i.e. true at all times). Then,

- (1) it was true 1,000 years prior to T that *Jones RAISE his right hand at T*
- (2) Necessarily, if it was true 1000 years prior to T that *Jones RAISE his right hand at T*, then Jones raises his right hand at T

- (3) Therefore, it is not the case that it is within Jones power to refrain from raising his right hand at T

The argument starts with an assumption about an arbitrary and apparently contingent fact, here the fact that Jones raises his right hand at T. A prior truth about what Jones does at T then comes in with premise (1). It then takes as a further premise that there is a necessary connection between the prior truth and what the prior truth is about. This second premise seems unproblematic: how could it not be the case that Jones raises his right hand at T, if it was already *true* that he would do so 1,000 years prior to T? The argument ultimately gets to its fatalistic conclusion by transferring the necessity to Jones' action. If Jones raises his right hand at T, this happens out of necessity, in virtue of the necessity involved in (2). Of course, there is nothing special with the act of Jones raising his right hand at T, nor with the time 1,000 years prior to T which constitutes the prior truth. The argument can be easily generalized to any action, any time and any agent.

The argument is ill-formed because its structure amounts to a modal fallacy. Let us say that p is the proposition that *it BE true 1,000 years prior to T that Jones RAISE his right hand at T* and q is the proposition that *Jones RAISE his right hand at T*. Then, the ill-formed Main Argument structure is:

1	q	Hyphotesis
2	p	1 - Premise (1)
3	$\Box(p \rightarrow q)$	Premise (2)
4	$\Box q$	1,2,3

The step from 1 to 2 is informally justified by the notion of prior truths. If q is the case, then it *BE true 1,000 years prior to T that q* , which is p . Step 3 states the necessary connection between p and q , which as we have seen it appears to be unproblematic. And finally step 4 follows from the first three steps.

In order to see why the argument structure is fallacious, it is sufficient to provide a counterexample, i.e. a model where the hypothesis and the two premises are true and the conclusion false. To do so, we take a model where we have two possible worlds, α and β . We then assume that the accessibility relation is such that $\alpha R \beta$, i.e. β is an accessible world with respect to α . Ultimately, the model is such that p and q are both true at α and false at β . We now evaluate the hypothesis and the premises of the argument at α . They turn out to be true in the described model. The

hypothesis and the first premise are true at α simply because p and q are true at α . Premise (2) is also true at α because $(p \rightarrow q)$ is true in every accessible world with respect to α . $(p \rightarrow q)$ is true at β , in virtue of the fact that p is false at β . So, the hypothesis and the two premises are true when evaluated at the world α in the model. However, the conclusion $\Box q$ is false at α . In fact, it is not the case that q is true in all the accessible worlds with respect to α , because q is not true at β . This quick modal reasoning refutes the ill-formed Main Argument in virtue of its fallacious structure and also suggests how to amend it. The obvious move to make is to add a necessity-box in front of the first premise, so that we can employ **K** to reach the fatalistic conclusion. Then, the structure becomes

1	q	Hyphotesis
2	$\Box p$	1 - Premise (1)
3	$\Box(p \rightarrow q)$	Premise (2)
4	$(\Box(p \rightarrow q) \wedge \Box p) \rightarrow \Box q$	K
5	$\Box q$	2,3,4

With respect to the previous structure, here a necessity-box is added in front of p in step 2, and therefore the principle **K**, which we have seen is valid in all models of modal logic, can be summoned.

What reasons could we have to add a necessity-box in front of p ? The reason might come from the fixity of the past, because after all p reports a fact about the past, namely something that was happening 1,000 years prior to the time of Jones raising his right hand, i.e. the fact that at that time the proposition that *Jones RAISE his right hand at T* was already true. Fischer and Todd (2015) make this point, when they are talking about incompatibilism between prior truths and the power to do otherwise:

Clearly, the thought here as something to do with the nature of the past and the nature of what we could plausibly have a choice about. That is, the incompatibilist recommends that we accept premise (1) on the basis that it is an instance of a more general thesis concerning our lack of power over the past. (p.4)

It is not within my power now to affect the event of Caesar crossing the Rubicon, because it is simply too late for me to exercise some power over that event. In this sense, this event and the proposition that states its occurrence are now necessary. Similarly, one might argue that Jones has no

power over what *was* true 1,000 prior to his existence. So, in the same sense the proposition that *Caesar CROSS the Rubicon in 49 BC* is now necessary, p is now necessary because there is nothing Jones could do to change the fact that in the past the proposition q was true. Having said that, this is how the Main Argument becomes.

The (correct) Main Argument

Suppose Jones raises his right hand at T . From this it follows that the proposition that *John RAISE his right hand at T* is true at T . If such proposition is true at T , then it is always true (i.e. true at all times). Then,

- (4) Necessarily, it was true 1000 years prior to T that *Jones RAISE his right hand at T*
- (5) Necessarily, if it was true 1,000 years before T that *Jones RAISE his right hand at T* , then Jones raises his right hand at T
- (6) Therefore, it is not the case that it is within Jones power to refrain from raising his right hand at T

Again, the argument purports to establish that if Jones raises his right hand at T , then he can't refrain from doing it. We have seen that there is nothing special with Jones' act, nor with the time T . Hence the argument generalizes, and if correct it proves the schema $A \rightarrow \Box A$, given its premises. This is exactly the conclusion a fatalist would need.

It is worth noting that it looks as if different modalities are involved in the argument. The modality in the second premise appears to be of the broadest kind. There are not possible worlds such that the antecedent of the conditional embedded in the necessity-box is true, and the consequent false. Or, it seems unconceivable to think about a scenario in which the antecedent is true and the consequent false. The modality in the first premise instead seems to be of another kind. It has to do with what Jones has within his power at T and it is the kind of modality the fatalism debate addresses. The modality in the conclusion is the same of the first premise. So, one might think that the whole argument is equivocal because different modalities are involved. I don't think this is the case. After all, the modality used in the conclusion is narrower or equal to all modalities involved in the two premises. The argument would be equivocal if it drew a conclusion involving a modality broader than the narrower modality used in the premises. But this is not what happens in the (correct) Main Argument.

So, the argument is valid. The conclusion follows from the two premises. Hence, there are three ways to reject it. The first one is to claim that the argument is not sound, i.e. that at least one of its two premises is not true. The second possible way is to claim that the premises are equivocal or even meaningless. The third one is to attack the background assumptions the argument relies on. As I already said, nobody accepts the (correct) Main Argument together with its background assumptions. And this is understandable. The (correct) Main Argument sounds like metaphysical sophistry. Why should our freedom of action be limited by prior truths, when it seems that those truths depend on what we do? However, there is not an universal agreement about what is wrong with the (correct) Main Argument. Several different strategies have been employed to resist the fatalistic conclusion. In the next sections we shall see four different ways to criticize the argument: the rejection of the notion of prior truths, Ockhamism, the introduction of a third truth value and the mutable futurism.

Prior Truths

Van Inwagen (1983) provides a reason for not accepting the Main Argument. His strategy to reject it focuses on the notion of prior truths. He attacks the notion of a proposition being true (or false) at a time. According to Van Inwagen, an expression like “the proposition p is true at T ” is either meaningless, or, if we attach the only plausible meaning to it, it does not grant the Main Argument’s conclusion. We shall focus here on the first horn of the disjunction, i.e. the idea that talking about truth (or falsity) at a time is meaningless. Of course, if things are so, the Main Argument does not get through, because it relies on talks about truth at a time in its two premises and in the background assumptions.

Before showing why talks about propositions being true or false at a time are meaningless, Van Inwagen illustrates the account of propositions he favors. He says he does not need to give a full account of propositions, because few basic concepts about truth and propositions are sufficient to refute the Main Argument and its usage of prior truths. Van Inwagen starts by saying what propositions are. As he says:

I do not mean anything mysterious by “proposition”. I use this word as a general term for the things people assent to, reject, find doubtful, accept for the sake of the argument, attempt to verify, deduce things from, and so on. (p. 62)

He then goes on by saying that propositions have properties.

Propositions...may be empirically verifiable, hard to understand, inexpressible in the tongue of a certain tribe, and so on. (p.64)

Among the properties of propositions, there are truth and falsity. Propositions may *be* true or false.

Van Inwagen then moves towards considerations about sentences. Sentences are typically *used* to express propositions. I can utter the sentence "the Queen Elizabeth I died on 24 March 1603" to unambiguously express the proposition that *the Queen Elizabeth DIE on 24th March 1603*, a proposition which is true if and only if the Queen Elizabeth dies on 24th March 1603. Sentences, like propositions, may be true or false. But, Van Inwagen argues, they are true (or false) just *in virtue* of expressing true (or false) propositions. Thus far, nothing special has been claimed. However, there is a special type of sentences which needs to be analyzed more carefully. These are sentences containing indexical words, i.e. words which can express a different content in different contexts of utterance. Here are the examples Van Inwagen uses to make his point:

- (7) I am Napoleon (uttered by a madman)
- (8) I am Napoleon (uttered by Napoleon)
- (9) I am about to die (written by Ariel in 1973)
- (10) I am about to die (written by Ariel in 1983)

We assume Ariel dies in 1983, three days after having written the sentence (10) in his diary. (7) and (8) are made up with the same words. Yet, the madman said something false and Napoleon said something true. The same holds for (9) and (10). Ariel is wrong when she writes down (9) in her diary, while she is right when she writes down (10). We have seen that sentences are true or false in virtue of expressing true or false proposition. So, it must be the case that (8) and (10) express true propositions, whereas (7) and (9) express false propositions. Given that the proposition expressed by (8) and the one expressed by (7) have different properties, because one is true while the other is false, they must be numerically distinct in virtue of the principle of the Indiscernibility of Identicals. And indeed they are. (8) expresses the true proposition that Napoleon is Napoleon, whereas (7) expresses the false proposition that the madman is Napoleon. What about (9) and (10)? Again, the two sentences differ in their truth value, one is true and the other false. Hence they must express different proposition. What is the source of this indexical phenomenon? Here the indexical "I"

clearly refers to the same individual in both contexts. “About” is a vague term, and we can assume for the sake of simplicity that it unambiguously means within one year. So, we are still looking for a reason why (9) and (10) express two distinct propositions, in spite of the fact that the same sentence is used by the same individual. The only candidate left is the present tense in the sentence, which is used at two different times. In fact, the sentence as it used in (9) expresses the false proposition that *Ariel DIE in 1973*, while the same sentence used in (10) expresses a distinct proposition, namely the true proposition that *Ariel DIE in 1983*. Hence, tenses within sentences behave like indexical terms in some important respect. The same sentence can deliver different contents if uttered at different times.

We now go back to the Main Argument. We have seen that in order to state it, we take advantage of expressions like the proposition q was true 1,000 years prior to T . This usage is attacked by Van Inwagen. As he says

...the above paragraphs contain phrases I do not understand. Among them are: “true at some particular moment”, “true at every moment”, “became true”, “remained true”, “is unchangeably true”, and so on. That is — and we must be very careful about this — I do not see what these phrases mean if they are used as they are used in the above argument for fatalism. (p.66)

Here is why Van Inwagen thinks so. The idea again is that tense in sentences behaves indexically. The result is that the propositions expressed by sentences already contains a reference to a time⁹. If at T I say “I am happy”, I express the proposition that *Giacomo BE happy at T* . If at T I say “I will be happy in two days” I express the proposition that *Giacomo BE happy at $T+2$* . Now, Van Inwagen observes that it seems more appropriate to say that such propositions are true or false *simpliciter* instead that true or false at times. They are true just in case what is said to take place at a certain time takes place at that time, and false otherwise. Talking about truth at a time is misleading and unnecessary.

A supporter of the truth-at-a-time talk may react by appealing to the following discourse, which is discussed by Van Inwagen. One might say in 2010 “that municipal bonds are a good investment”. In case municipal bonds were a good investment in 2000, but they are no longer a good investment in 2010, then a possible reply could be “that used to be true, but it is no longer true nowadays”. This is a piece of discourse which

⁹with the exceptions of timeless propositions like the proposition that the sum of all angles is always the same in any triangle. These exceptions are of course irrelevant with respect to fatalism, which is concerned with human agency, which obviously happens in time.

appears to be natural and transparent. Moreover, it seems to suggest that there is something —to make it clear, one and the same thing — which used to be true and which is no longer true. That something might be a proposition, because we have seen that propositions are the bearers of the properties of truth and falsity. If things are so, the fatalist usage of truth-at-a-time would be legitimate. Suppose that a newspaper used the headline “municipal bonds are a good investment” in the year 2000 and used the same headline in 2010. According to the talk “that used to be true” taken at face value, the two headlines expressed one and the same proposition, namely the proposition that municipal bonds are a good investment. Let’s call this proposition “*s*”. If things were so, the specification of time when talking about the truth of falsity of *s* would be not only legitimate, but even necessary. We would need to say things like — *s* was true in the year 2000, whereas *s* is false in the year 2010. But, Van Inwagen has an easy time to explain away this “used to be true” kind of talk. If such an usage were to be taken literally, we should be allowed to say things like “the 2000 edition of the newspaper used to be reliable, but today, that very same edition is not reliable anymore, because it states something, i.e. *s*, which is nowadays false”. But this last statement is clearly absurd. The 2000 edition is and always will be reliable, because it expressed the true proposition that *municipal bonds Be a good investment in 2000* by means of the headline “municipal bonds are a good investment”. The 2010 edition is and will always be not reliable, because the same headline expresses the false proposition that *municipal bonds BE a good investment in 2010*. Clearly, the “used to be true” usage refers to the fact that if someone had used the sentence “municipal bonds are a good investment” in 2000 he would have expressed a true proposition, whereas if he had used the same sentence in 2010 he would have expressed a different and false proposition.

As such, the argument does not uncontroversially establish that the truth-at-a-time talk is meaningless. It may establish that tenseless propositions do not change truth value over time. However, this does not imply that the truth-at-a-time talk is meaningless. We do say things like it is true today that *Caesar DIE in 44BC* and one may take this kind of talk literally. Or, we could say that even if propositions fundamentally have a truth value *simpliciter*, there is a derivative sense according to which it is correct to say that propositions have truth values at times, in virtue of the principle that if something is true (or false) *simpliciter*, then it is true (or false) at all times. Or, one might favor tensed propositions. In such case, even if there is a temporal index embedded, it might be reasonable to have truth values which change over time, thereby making sense of the truth-at-a-time talk. For instance, Prior (1976) seems to make this point:

One way of sliding out, on which I don't think it is necessary to waste much time, is to make nonsense of most of the argument by denying that you can attach a time to a truth-value. But in fact we all use phrases like "was true yesterday" perfectly happily when we are not philosophizing...

We have seen one strategy Van Inwagen used to refute the Main Argument. It is based on a rejection of the concept of prior truths, which the Main Argument needs in order to get through. It seems uncontroversial that tenseless propositions do not change truth value over time. Yet, this does not rule out the legitimacy of the truth-at-a-time talk which is employed in the Main Argument. We shall see a different strategy in the next section.

Ockhamism and the freedom to do otherwise

In this section I will illustrate the so-called Ockhamist strategy to refute the Main Argument. I will rely on Fischer and Todd (2011) version of Ockhamism. The strategy is based on two main elements. The first one is an account of what an agent can or cannot do, which does justice to the idea that the past is fixed. The second one is the distinction between soft and hard facts.

The account of what an agent can or cannot perform sounds as follows.

Can Account

An agent A can perform (has within his or her power) an act ϕ at a time T and a world w iff there is an accessible possible world where A does perform ϕ at T .

The account as such is incomplete, because it does not say anything about the nature of the accessibility relation. The most obvious way to characterize the accessibility relation is to require that the accessible possible worlds are such that they are historically identical to w up to the time T . In other words, two worlds are historically identical up to T just in case they are qualitatively indistinguishable up to T . This account of "can" does justice to the idea that the past is fixed, i.e. the idea that we have no power over what is past. The following case illustrates why things are so. Suppose that today John wants to buy an expensive car, but he is far from having the necessary money to buy it. Everybody would agree that John cannot purchase the expensive car. His lack of money would stop him from purchasing it. If John has not the money to buy it, it means that

at the moments in time prior to today, there is not enough money on his bank account. We check now what the described “can” account predicts. In order to establish whether John can buy the car today, we have to check all the possible worlds which are identical to the world where John wants to buy a car, from their beginning to today, i.e. the moment in time in question. In all those worlds, John has not the necessary money to buy the car. This implies that in all those worlds John *does not* buy a car. As a result, the account correctly predicts that in the world that John inhabits, he *can’t* purchase the expensive car. If he could purchase the car, John could do something which would change the past, which is absurd.

That being said, a fatalist who accepts the Main Argument could adopt this account of can and argue in favor of fatalism. He or she could say that looking at John’s yesterday bank account was after all unnecessary. In fact, if John doesn’t purchase today the car, it was true 1,000,000 years ago that he would not have done so, and this fact lies in the past. Hence, he or she would conclude that all the accessible worlds are such that there are facts which imply that John would not purchase the car today. This of course generalizes to any fact, any time and any agent, and so it would lead to the fatalistic conclusion that no one has ever the power to do otherwise.

To avoid this fatalistic conclusion, here it comes the Ockhamist distinction between *soft* and *hard* facts. Fischer and Todd (2011) puts the distinction this way.

The distinction between hard and soft facts (although not the terminology) traces back to William of Ockham. Ockham employed this distinction to give a certain sort of response to the Divine Foreknowledge Argument, but the distinction is crucial also in providing a proper interpretation of the argument itself. Hard facts are (in some way that is hard to characterize precisely) temporally non-relational as regards the future (relative to the time they are about). More specifically, a hard fact about some time T is genuinely about T and not also genuinely about some time after T. In contrast, a soft fact is temporally relational as regards the future (relative to the time it is about); that is, a soft fact about some time T is at least in part genuinely about some time after T. (p.115)

To illustrate the distinction, we take the following two facts. It is a fact about 2008 that Obama was elected. It is also a fact about 2008 that Obama was elected 8 years prior to me typing this sentence. The first fact qualifies as hard, because it has to do only with 2008. The latter is soft, because it is

as it is in virtue of things happening in 2008 (Obama's election) and things happening in 2016 (me writing that sentence).

With the Ockhamist machinery in our hands, we can now challenge the Main Argument. The reader probably remembers that the Main Argument appeals to facts like the fact that it was the case 1,000 years prior to T that Jones would raise his right hand at T. Clearly, this sort of fact goes under the category of soft facts, because it says something about 1,000 years prior to T *and* about a time which is future with respect to 1,000 years prior to T. Now, an Ockhamist would say that when we look at the accessible worlds with respect to a time T and a world w to determine what an agent can or cannot do, we should take as accessible all the worlds which are identical to w up to T just with respect to the hard facts. This, in the case of Jones and his right hand, yields the result that it may be the case that there are accessible worlds such that they would differ from Jones's world with respect to the soft past. That is, there might be worlds where it was true in the past that Jones would *not* raise his right hand at T and Jones does *not* raise his right hand at T. This would be the case in circumstances where we would not have "hard" reasons to think that Jones's behavior is forced by things which would undermine his power to do otherwise, like the past presence of a microchip in his head. If so, the "can" account gives us the result that it might be the case, if circumstances allows it, that Jones can refrain from raising his right hand at T, even if he does not do so at T. So, we have a counterexample against modal fatalism.

Going back to the Main Argument, it should be made it clear that the Ockhamist does not deny that

- (11) it was true 1,000 years prior to T that Jones would raise his right hand at T.

provided Jones raises his right hand at T. (11) truly reports a soft fact about what was the case 1,000 years prior to T. What is not true, the Ockhamist believes, is the fact that Jones has no power at T over the fact reported by (11). He or she would argue that whether (11) holds or not depends on what Jones does at T. If he has the power to do otherwise at T, then it is within Jones's power whether it was the case that 1,000 years prior to T that Jones would raise his right hand at T. Hence, the Ockhamist, in virtue of this dependence relation, argues that the necessity-box in front of (11) in the (correct) Main Argument may render (11) false — it would be false if as a matter of fact Jones has such a power at T.

To sum up, if Jones raises his right hand at T, the Ockhamist concedes to the fatalist who believes in the Main Argument that it was true 1,000 years prior to T that Jones would raise his right hand at T. This means that,

with certainty, John will raise his right hand at T. Whether Jones has or not the power do to otherwise, though, does not depend at all on the soft fact employed in the Main Argument. It does depend just on the hard past with respect to T.

On a side note, the distinction between soft and hard facts can be employed to make sense of the so-called Thin Red Line Theory about the future, which is itself inspired by the work of Ockham. According to this view, at *any* moment in time there is a single unique actual future, labeled the Thin Red Line, which corresponds to how things will actually unfold. Yet, it may be the case that at *some* point in time, more than one future is possible, besides the actual one. Take a time T and all the *hard* facts about times prior to T. It may be the case that more than one continuation is possible after T. For instance, an indeterministic scenario where the same initial conditions can give rise to different outcomes would be such a case. However, by taking into account all soft facts about times prior to T, we have the result that only one possible future is singled out, i.e. the Thin Red Line. For instance, there would be a soft fact about the past according to which the outcome at T_1 is such and such. By taking into account only hard facts, distinct futures are possible. By bringing into the picture soft facts too, the Thin Red Line is lighted.

Introducing a Third Truth Value

In this section we shall deal with another classic strategy to resist the fatalistic conclusion provided by the Main Argument. The strategy is based on the introduction of a third truth value between truth and falsehood. We will illustrate it with the help of Lukasiewicz (1967). As he says:

I can assume without contradiction that my presence in Warsaw at a certain moment of next year, e.g. at noon on 21 December, is at the present time determined neither positively or negatively. Hence it is possible, but not necessary, that I shall be present in Warsaw at the given time. On this assumption the proposition I shall be in Warsaw at noon on 21 December of next year can at the present time be neither true nor false. For if it were true now, my future presence in Warsaw would have to be necessary, which is contradictory to the assumption. If it were false now, on the other hand, my future presence in Warsaw would have to be impossible, which is also contradictory to the assumption. Therefore the proposition considered is at the moment neither true nor false ... (p. 53)

Here is the way I understand this famous passage. Lukasiewicz starts his reasoning with the idea that what happens at a time, say the year 1922, is compatible with more than one future. What happens in 1922 is compatible with his presence in Warsaw in December 21st 1923 and, it is also compatible with his absence from Warsaw the same day, i.e. there is no contradiction with assuming all the facts about 1922 together with Lukasiewicz's future presence in Warsaw, and there isn't a contradiction as well if we assume all the facts about 1922 together with his future absence from Warsaw. That is to say, how the world is in 1922 does not settle the issue about Lukasiewicz's whereabouts one year later. The obtained result is that his presence in Warsaw the day in question is not necessary, and so is his absence. We state this first conclusion with (12).

- (12) it is not necessary in 1922 that Lukasiewicz will be in Warsaw one year later and it is not necessary in 1922 that Lukasiewicz will not be in Warsaw one year later.

This first part of Lukasiewicz reasoning does not justify by itself the introduction of a third truth value. For instance, an Ockhamist might say that even if it is true in 1922 that Lukasiewicz would (or would not) find himself in Warsaw one year later, this might not be necessary (or impossible). And so here is the point where the Main Argument kicks in. The reader should remember that the Main Argument purports to prove the validity of conditionals of the form $A \rightarrow \Box A$, and that one of the two sources the Main Argument relies on are the so called prior truths, whereas the other is the fixity of the past. For any future event, The Main Argument states, there are prior truths about that event, therefore the future event will happen of necessity.

Let us say that we are in 1922, "F(n)" is the future operator which means "in n time units it will be the case that", "n" is the temporal index which specifies the time interval, and "w" stands for Lukasiewicz's presence in Warsaw. We can then form the two following tensed propositions. The proposition $F(1)w$, which is the proposition that in one year it will be the case that Lukasiewicz is in Warsaw, and the proposition $F(1)\neg w$, which is the proposition that in one year it will be the case that Lukasiewicz is not in Warsaw.

We might then give this very brief argument, which moves from the future excluded middle (FEM) together with the conclusion of the Main Argument to the idea that one of two mutually exclusive future events is necessary. In 1922, we can reason as follows.

FEM-MA-N(From FEM and Main Argument to Necessity)

- (13) either $F(1)w$ is true or $F(1)\neg w$ is true
- (14) if $F(1)w$ is true, then $F(1)w$ is necessary
- (15) If $F(1)\neg w$ is true, then $F(1)\neg w$ is necessary
- (16) therefore, either $F(1)w$ is necessary or $F(1)\neg w$ is necessary

Premise (13) is not exactly an instance of (FEM). (FEM) states that any claim of the form $F(n)\phi \vee F(n)\neg\phi$ is valid¹⁰. It must not be confused with the standard excluded middle form $\phi \vee \neg\phi$, because the negation in (FEM) is embedded in the future operator. (FEM) is justified by the intuitive thought that among two mutually exclusive events which exhaust the possibilities, necessarily one of them will take place at a given future time. It seems straightforward to derive (13) from $F(1)w \vee F(1)\neg w$. Thus, we have (13). Premises (14) and (15) are true if one accepts the Main Argument. Prior truths about future events, together with the fixity of the past, lead to the necessity of the events the prior truths are about. (16) then follows. But (16) is in contradiction with (12), the conclusion of the considerations made in the first step of Lukasiewicz's reasoning — that his presence in Warsaw is not necessary, nor is his absence.

We need to stop for a while at this point. The careful reader could be worried about the fact that thus far we discussed the Main Argument in terms of tenseless propositions. The argument FEM-MA-N instead employs tensed propositions, where the tense is fundamental and part of the content of the proposition. Hence, we should show how the Main Argument can be given in terms of tensed propositions. This will justify the steps (14) and (15) employed in FEM-MA-N. To do so I borrow Øhrstrøm (2009) tempo-modal formalism and the argument he illustrates there, with some passages made more explicit.

The tempo-modal formalism takes the following five principles as axioms, where “p” and “q” stand for any well-formed formula in the language, and the operator $P(n)$ is the past operator which has to be read as “it was the case n units of time ago that..”, in analogy with the metric future operator introduced above.

Axioms

- A1 $F(y)\phi \rightarrow P(x)F(x)F(y)\phi$
- A2 $\Box(P(x)F(x)\phi \rightarrow \phi)$
- A3 $P(x)\phi \rightarrow \Box P(x)\phi$

¹⁰I will be using the greek letters ϕ and ψ as metavariables for propositions which come with tensed operators, as it is customary in the literature.

$$A4 \quad (\Box(\phi \rightarrow \psi) \wedge \Box\phi) \rightarrow \Box\psi$$

$$A5 \quad F(x)\phi \vee F(x)\neg\phi$$

These axioms parallel the assumptions which the (correct) Main Argument needs in order to get through. A1 corresponds to the move the Main Argument makes from truth of a proposition at a time to truth of the same proposition at an earlier time. The Main Argument moves from the idea that if the relevant tenseless proposition is true at a time, it is also true at earlier times. Here A1 tells us that we are allowed to 'go back' in time as much as we want, provided we make the necessary adjustments with the tensed operators. A2 corresponds to the second premise of the Main Argument. It seems to be a matter of necessity that if it was the case x units ago that it would be the case x units of time later that p , then now p . Similarly, the Main Argument claims that it is a matter of necessity that if it was true 1,000 years ago that *Jones RAISE his right hand at T*, then Jones raises his right hand at T. A3 delivers the idea of the fixity of the past. If something was the case in the past, it is now unpreventable, i.e. there is nothing we can do now about it. As in the Main Argument, it seems that two kinds of necessity are at play. A3 employs the kind of necessity fatalism is interested in, i.e. necessity in terms of what an agent has or hasn't within his or her power. A2 (and A4 as well) refers to a kind of logic or metaphysical necessity. A4 is just the modal principle K, which as we have seen it is needed for the Main Argument to be formally valid. A5 is the Future Excluded Middle. Similarly, The Main Argument employs the standard excluded middle to conclude that all actions are either necessary (if performed) or impossible (if not performed).

We are now in a position to give the (correct) Main Argument in tensed terms. Again, we reason about Lukasiewicz's future whereabouts as if we were in 1922.

The Tensed Main Argument

1	$F(1)w \rightarrow P(1000)F(1000)F(1)w$	A1
2	$P(1000)F(1000)F(1)w \rightarrow \Box P(1000)F(1000)F(1)w$	A3
3	$F(1)w \rightarrow \Box P(1000)F(1000)F(1)w$	1,2
4	$\Box(P(1000)F(1000)F(1)w \rightarrow F(1)w)$	A2
5	$\Box(P(1000)F(1000)F(1)w \rightarrow F(1)w) \rightarrow (\Box P(1000)F(1000)F(1)w \rightarrow \Box F(1)w)$	A4
6	$\Box P(1000)F(1000)F(1)w \rightarrow \Box F(1)w$	4,5
7	$F(1)w \rightarrow \Box F(1)w$	3,6

Step 1 is an instance of A1. Step 2 is an instance of A3. Step 3 follows from 1 and 2 and the transitivity of conditionals. Step 4 is an instance of A2. Step 5 is an instance of A4. Step 6 follows from 4 and 5 in virtue of propositional logic. The conclusion 7 follows from 3 and 6 in virtue of propositional logic. We can obtain the result $F(1)\neg w \rightarrow \Box F(1)\neg w$ just by replacing w with $\neg w$ in the tensed Main Argument. These two results are the formal analogues of what is employed in (14) and (15) of FEM-MA-N.

Time to get back to FEM-MA-N. The conclusion of it is in contradiction with (12), which intuitively seem to be true according to Lukasiewicz. Hence, something must be given up. Lukasiewicz clearly accepts steps (14) and (15) of FEM-MA-N when he says “for if it were true now, my future presence in Warsaw would have to be necessary. . .”. This way the only possibility left is to reject premise (13) of FEM-MA-N. What happens if we claim that (13) is false?

(13) either $F(1)w$ is true or $F(1)\neg w$ is true

If (13) is false, we have equivalently that it is not the case that $F(1)w$ is true and it is not the case that $F(1)\neg w$ is true. That is, we have (\neg 13)

(\neg 13) $\neg(F(1)w \text{ is true}) \wedge \neg(F(1)\neg w \text{ is true})$

(\neg 13) is a conjunction, so by classical logic we are allowed to draw conclusions from both conjuncts and then put the results together. We start from the first conjunct.

(i) $\neg(F(1)w \text{ is true})$. (ii) Hence, $F(1)w$ is false. (iii) Therefore, $\neg F(1)w$

The step (ii) is justified if Bivalence holds and we use standard truth tables compatible with Bivalence. If truth and falsity are the only possible truth

values as Bivalence dictates, from the fact that it is not the case that $F(1)w$ is true it must follow that $F(1)\neg w$ is false. The step (iii) is justified by the thought that the negation of a false claim is true, which holds independently from Bivalence.

The second conjunct leads us to $F(1)w$ in a similar but more complicated way.

(iv) $\neg(F(1)\neg w$ is true). (v) Hence, $F(1)\neg w$ is false. (vi) Hence, $\neg F(1)\neg w$. (vii) Hence, $\neg\neg F(1)w$. (viii) Therefore, $F(1)w$

Step (v) is justified if Bivalence holds. Again, if truth and falsity are the only possible truth values, from the fact that it is not the case that $F(1)\neg w$ is true, it follows that $F(1)\neg w$ is false. Step (vi) is justified by the idea that a negation of a false statement is true, which, as just seen, is independent from Bivalence. Step (vii) moves the negation sign which follows the future operator in (vi) before the future operator. Its intuitive ground is the fact that saying that at a given future time something will not take place is tantamount to saying that it is not the case that at the given future time that something will take place. We shall see later that the claim just made can be disputed. Step (viii) is just the result of canceling out two negation signs from (vii). (iii) and (viii) form a contradiction.

To sum up, FEM-MA-N lead us to (16), which is in contradiction with (12). Given that Lukasiewicz believes in the intuitive truth of (12) and accepts the premises (14) and (15) of FEM-MA-N, Lukasiewicz considers the premise (13) to be the culprit and he thus rejects it. But in turn, we have shown that the denial of (13) together with Bivalence and other plausible principles lead us to the two contradictory statements (iii) and (viii). Hence, something else must be given up. Lukasiewicz thus rejects Bivalence. Statements about future events which are not presently settled have a third truth value which goes beyond the list of truth values allowed by Bivalence. This way, (13) is not true. If the future Lukasiewicz's whereabouts is presently unsettled, both $F(1)w$ and $F(1)\neg w$ have the third truth value, and thus neither of them is true. It doesn't matter how we call this third truth value. It is usually referred to as indeterminate or neither true nor false. What is important is that it represents a middle ground between truth and falsity and that the introduction of it amounts to a denial of Bivalence. By doing this, we no longer have the contradiction of the claims (iii) and (viii). We derived both of them by employing Bivalence. If we reject Bivalence, the steps from (i) to (ii) and from (iv) to (v) made above are no longer justified.

It turns out that this is not the only way to make sense of future con-

tingent events, if there are such things. It is possible to argue that all future contingent propositions are false, instead that assigning to them the third truth value. In Chapter 7 of his 1967 *Past, Present, and Future*, Prior discusses such a view.

The reasoning employed to defend the all-false view on future contingents is similar to the one it is used to justify the introduction of a third truth value. Most of the steps are the same. An all-false theorist starts from the idea that (12) is in contradiction with the conclusion of FEM-MA-N. He or she then denies the truth of premise (13) of FEM-MA-N.

(13) either $F(1)w$ is true or $F(1)\neg w$ is true

which leads us to

(\neg 13) $\neg(F(1)w \text{ is true}) \wedge \neg(F(1)\neg w \text{ is true})$

We then have the two inferences from the two conjuncts.

(i) $\neg(F(1)w \text{ is true})$. (ii) Hence, $F(1)w$ is false. (iii) Therefore, $\neg F(1)w$

(iv) $\neg(F(1)\neg w \text{ is true})$. (v) Hence, $F(1)\neg w$ is false. (vi) Hence, $\neg F(1)\neg w$. (vii) Hence, $\neg\neg F(1)w$. (viii) Therefore, $F(1)w$

Thus far, we are reasoning following the same track used by a third-value theorist. But, and here comes the difference, an all-false theorist accepts the steps from (i) to (ii) and from (iv) to (v), and he or she concludes the reasoning by saying that both $F(1)w$ and $F(1)\neg w$ are false. Bivalence is thus preserved, because a third truth value is not introduced and (13) is denied in virtue of the fact that both $F(1)w$ and $F(1)\neg w$ are false. Of course, at this point something must be said to avoid to have both (iii) and (viii), otherwise we would have a contradiction. At least one of the paths which lead us there must be stopped. The strategy goes as follows. The main idea is that the operator $F(n)$ means something like "it is presently determined that in n time units it will be the case that...". This way $F(n)\psi$ is true iff ψ is already settled, and false otherwise. This means that $F(n)\psi$ is false in two cases. Either $F(n)\psi$ is false because $\neg\psi$ is determined to happen in the future, or because ψ is not presently settled. This way it is easy to see that the move from (vi) to (vii) is not legitimate, and this is so because in such a view we can't safely move a negation sign from the scope of a future operator to the outside. In other words, formulas like $F(n)\neg\psi$ and $\neg F(n)\psi$ are not equivalent. In fact, if ψ is presently unsettled, $F(n)\neg\psi$ is false, whereas $\neg F(n)\psi$ is true because it is an external negation of a false statement.

The all-false view of future contingent propositions makes sense of contingency, necessity and impossibility of future events. If $F(n)\psi$ is true, then ψ is necessary. If $F(n)\neg\psi$ is true, then ψ is impossible (and so far we are agreeing with a third truth value theorist). If $F(n)\psi$ and $F(n)\neg\psi$ are both false, then ψ is contingent.

Of course, these two views, the third truth value view and the all-false view, do not come without costs. One of the main problems the two views share is the so-called prediction problem. Suppose someone said in 1922 that Lukasiewicz will be in Warsaw one year hence. Assume then that when 1923 comes, Lukasiewicz as a matter of fact finds himself in Warsaw. It seems correct to say that what has been said in 1922 was true. But this is not what the two theories say. According to the third truth value theorist, what was said wasn't true because it was neither true nor false. According to the the all-false theorist things are even worse, what was said was literally false.

Another problem worth mentioning has to do with the Law of Excluded Middle. It seems a logical truth that any sentence of the form $\psi \vee \neg\psi$ is true. But things get problematic for the third truth value theorist. Suppose that ψ is the sentence "it will be the case in one year that Lukasiewicz is in Warsaw", pronounced in 1922. This sentence and its negation are neither true nor false according to the third truth value theorist. What shall we say of the whole disjunction when the two disjuncts have the third truth value? We may be tempted, upon considering this case, that the whole disjunction is true when the two disjuncts have the third truth value. After all, Lukasiewicz will or will not be in Warsaw the day in question. But what about sentences like "either it will be the case in one year that Lukasiewicz is in Warsaw or it will be the case in one day that I have eggs for lunch" where the meal I am going to have tomorrow is also presently unsettled. If the disjunction of two indeterminate disjuncts yields a truth, then we have to say that this disjunction is true. But this seems intuitively wrong, because it could turn out that Lukasiewicz will not be in Warsaw and I will not have eggs. One might then say the disjunction operator is not truth-functional. It yields a truth when the two disjuncts are both indeterminate and are contradictories, and it yields another truth value when the two disjuncts are both indeterminate but are not contradictories. This could probably do, but besides seeming an *ad hoc* move, it seems also that the disjunction operator is truth-functional. The all-false theorist has a similar problem. He or she can claim that the excluded middle holds unrestrictedly, because there isn't a third truth value in the picture and the external negation works in the ordinary way. But the future excluded middle, $F(x)\psi \vee F(x)\neg\psi$ is not guaranteed to be true, because both disjuncts

may be false. And again, this seems intuitively wrong. And there are also other problems we don't have time to discuss. Iacona (2007) contains a thorough discussion of these problems.

To conclude, the introduction of a third truth value (and the all-false view as well) is justified by, among other things, the acceptance of the idea coming from the Main Argument that prior truths imply necessity. So, one might say that this move is not properly a rejection of fatalism. But, I think these two strategies go legitimately in the category of the theoretical moves against fatalism for at least two reasons. The first reason is that they deny a crucial background assumption used in the (correct) Main Argument. The Main Argument claims that if Jones raises his right hand at T, it is true at T that *Jones RAISE his right hand at T*. The Main Argument then goes from the truth at a time, to truth at all times. A third truth value theorist and all-false theorist deny this latter move. Even if it's true at T that Jones raises his right hand at T, it doesn't follow that it was true before T that Jones would raise his right hand at T. If Jones' act was unsettled prior to T, the proposition in question was either neither true nor false or simply false. The second reason is that the two views are motivated by the idea that the past together with the present may fail to settle a unique future, which is precisely the opposite of what a fatalist would be willing to say.

The Mutable Future

In this section we shall deal with a rather unknown view, i.e. the idea that the future is mutable. Todd (2016) recently presented the picture, which is originally due to Peter Thomas Geach (1977). I will first explain what this view amounts to and we'll see then how it provides at least two ways to refute the Main Argument.

The Mutable Futurist claims that the future might literally change. What does this could possibly mean? We will see first what it does *not* mean. The Mutable Futurist doesn't claim that there is change in the future. To say that there is change in the future is to say that at least two future times are qualitatively different or that at least two future times contain qualitatively different entities. For instance, probably the year 2097 will contain more advanced technology than the year 2083. This kind of change in the future is an uninteresting platitude everyone agrees on. Nor the Mutable Futurist claims that the future changes because future times will become present, as a standard A-theorist would say. Rather, the Mutable Futurist claims that a given content or constituent of a given future time might be replaced by something else as time goes by. This way the future changes, from containing something to containing something else

or from being constituted by something to being constituted by something else.

Todd puts it this way in presenting the view.

what is distinctive of Geach's view is that it is possible for something X to be such that, at T_1 , it will happen at T_{10} , but at T_3 such that it will not happen at T_{10} . (p.2079)

certain events sometimes can make it so that what will happen ends up not happening, or that what previously was such that it will not happen does happen...we get the following unique result: it could very well be that at T_1 it was true that X would happen at T_{10} , but that, at some time later than T_{10} , it is not true that X happened at T_{10} . (p.2080)

In such a picture future oriented propositions behave accordingly. We may have cases where future contingent propositions change their truth value over time. They can go from being true to being false, and from being false to being true. Suppose that at T_1 X is going to happen at T_{10} , and then at T_3 the future changes, Y instead of X is going to happen at T_{10} . If things are so, the proposition that X will happen at T_{10} is true at T_1 and it becomes false at T_3 , while the proposition that X will not happen at T_{10} is false at T_1 and it becomes true at T_3 .

I bet that the reader at this point has already a lot of objections to this view about the future. We will have a great deal more to say about it in the last chapter. As for now, I just need to say that I agree with Todd when he claims that this view should deserve the role of a metaphysical contender within the debate on time. We are about to take a look at two motivations for this view. The first one comes from Geach himself, the first proponent of this picture, whereas the second comes from Todd.

Geach says that what motivates the changing future view is what he calls the "logic" of prevention. As Geach says:

But what then is prevented? Not what did happen, but assuredly what was going to happen. The aeroplane was going to crash into the sea and 100 men were going to be drowned; the pilots' prompt action prevented this. For not everything that does not happen is prevented: only what was going to happen. (p.47)

Geach argues for the idea that when an event is prevented from happening, that event is such that it was going to happen, but then it didn't, thanks to

the preventative act which ensured the non-occurrence of it. Suppose that a preventative act occurs at the time T_3 and it prevents the occurrence of an event X at T_{10} . Geach would say that before T_3 X was going to happen, and after T_3 X is no longer going to happen. At T_3 , the future time T_{10} changes from containing X to not containing X .

The second motivation, given by Todd, has to do with the utility of foreknowledge. It seems uncontroversial that foreknowledge of the future is useful in case I have to make a decision in the present. Suppose I am about to buy stocks and I am wondering which investment will make the more gain. Surely, the foreknowledge of how the stock market will behave would make me incredibly rich in few days. It seems that the more knowledge I have, the better for me. But, if I had a *total* knowledge of what will happen in the future, and the future cannot change, then this knowledge would seem to be totally useless, because I would foreknow also my future actions, stocks purchases included. Suppose A is going to be a good investment and B a bad one. And suppose I'll buy B . If I knew everything about the future, I would also know that I will buy B , despite knowing it's a bad investment. So, partial knowledge of the future is an advantage when I have to deliberate. And, the more knowledge of the future I have, the better for me. But, as this knowledge approaches completeness, it becomes useless. This is true if the future is immutable. So, here is the argument in favor of the mutable future. The immutability of the future together with full foreknowledge of the future seem to imply that full foreknowledge is useless. But foreknowledge seems to be extremely useful. Hence, one might conclude by modus tollens that the future *must* be mutable. These two motivations are not conclusive. One might have an alternative account of prevention which does not posit a mutable future, or one might have a different account of foreknowledge. We will discuss this view in more details later.

To conclude, I will show how a Mutable Futurist can reply to the Main Argument. First of all, like a third truth value theorist, the Mutable Futurist can claim that from the fact that it is true at T that Jones raises his right hand at T it does not follow that it has always been true in the past. Maybe it was the case before T that Jones would not raise his right hand at T , but then the future changed.

The second reply has to do with the second premise of the Main Argument.

- (8) Necessarily, if it was true 1000 years before T that *Jones RAISE his right hand at t* , then Jones raises his right hand at T

According to a Mutable Futurist the conditional is not guaranteed to

hold. It is not guaranteed that if it was true 1000 years prior to T that Jones would raise his right hand at T, then Jones does so at T. Some event happening from 1000 years prior to T and up to T might prevent Jones' raising his right hand at T. If the future changes from containing the event of Jones raising his right hand at T to not containing it, the antecedent is literally true while the consequent false. This way the second premise of the Main Argument is falsified.

1.2.3 Taylor's Argument

We now turn to another classic argument in favor of modal fatalism. In (1962) Richard Taylor published a paper where he provides an argument in favor of fatalism. His point is that six commonly accepted and reasonable assumptions lead to fatalism. The argument has an Aristotelian flavor, because it employs the principle of Bivalence and it has the form of a disjunctive dilemma. However, I think Taylor's argument is an improvement of the Aristotelian one. The further step Taylor does seem to strengthen the fatalistic conclusion. Like Aristotle, Taylor himself wasn't a fatalist. His ultimate position was to reject at least one of the assumptions which grounds his argument. Anyway, here I am not concerned with what view Taylor held, rather, I'll try to illustrate and analyze his argument and the responses to it. I will argue that most of the standard responses found in the literature do not refute the argument. Later on, I will show what I think is wrong with it. Taylor's argument proves too much, because it conflates deterministic and indeterministic cases.

The Argument

Taylor's argument rests on six assumptions, which he calls "presuppositions". So, before going through the argument, we shall see what these six assumptions are. As he says:

Presuppositions. The only presuppositions we shall need are the six following.

First, we presuppose that any proposition whatever is either true, or, if not true, then false. This is simply the standard interpretation, *tertium non datur*, of the law of excluded middle, usually symbolized ($p \vee \neg p$), which is generally admitted to be a necessary truth.

Second, we presuppose that, if any state of affair is sufficient for, though logically unrelated to, the occurrence of some

further condition at the same or any other time, then the former cannot occur the latter occurring also. This is simply the standard manner in which the concept of sufficiency is explicated. Another and perhaps better way of saying the same thing is that, if one state of affairs ensures without logically entailing the occurrence of another, then the former cannot occur without the latter occurring. Ingestion of cyanide, for instance, ensures death under certain familiar circumstances, though the two states are not logically related.

Third, we presuppose that, if the occurrence of any condition is necessary for, but logically unrelated to, the occurrence of some other condition at the same or any other time, then the latter cannot occur without the former occurring also. This is simply the standard manner in which the concept of a necessary condition is explicated. Another and perhaps better way of saying the same thing is that, if one state of affair is essential for another, then the latter cannot occur without it. Oxygen, for instance, is essential to (though it does not by itself ensure) the maintenance of human life, though it is not logically impossible that we should live without it.

Fourth, we presupposes that, if one condition or set of conditions is sufficient for (ensures) another, then that other is necessary (essential) for it, and conversely, if one condition or set of conditions is necessary (essential) for another, then that other is sufficient (ensures) it. This is but a logical consequence of the second and third presuppositions.

Fifth, we presuppose that no agent can perform any given act if there is lacking, at the same or any other time, some condition necessary for the occurrence of that act. This follows, simply from the idea of anything being essential for the accomplishment of something else. I cannot, for example, live without oxygen, or swim five miles without ever having been in water, or read a given page of print without having learned Russian, or win a certain election without having been nominated, and so on.

And sixth, we presuppose that time is not by itself efficacious; that is, that the mere passage of time does not augment or diminish the capacities of anything and, in particular, that it does not enhance or decrease an agents powers or abilities. This means that if any substance or agent gains or loses powers or abilities over the course of timesuch as, for instance, the power

of a substance to corrode, or a man to do thirty push-ups, and so on then such gain or loss is always the result of something other than the mere passage of time. (p.42-43)

We shall say something about these six assumptions. The first one, as it is stated, appears to be the principle of Bivalence, even if Taylor labels it the Law of Excluded Middle (LEM). The Principle of Bivalence states that truth and falsity are the only possible truth values for declarative sentences expressing propositions, those truth values being exhaustive and mutually exclusive. So, a semantic abides Bivalence if any sentence in the language has exactly one truth value, truth or falsity. The Law of Excluded Middle (LEM) is validated if any sentence of the form $(A \vee \neg A)$ is true. The two principles are not one and the same. For instance, we could have semantics like supervaluationism where LEM holds, whereas Bivalence does not. In any case, as we shall see, the Taylor Argument just needs the standard LEM.

The second, third and fourth assumptions deal with the relations of necessary and sufficient conditions. Taylor uses as relata of the relations states of affairs. If a state of affairs s is sufficient for another state of affairs n , it means that s cannot occur without n also occurring. If things are so, the occurrence of n is said to be necessary for the occurrence of s . We use italic letters to refer to propositions which claim that a state of affairs occurs, i.e. s is the proposition that s occurs. We then express the notions of sufficiency and necessity through the material implication. $s \rightarrow n$ expresses the fact that s is sufficient for n and n is necessary for s . It is important to note that this kind of relation might hold between states of affairs no matter what is the time of their occurrence — that is, we can have a sufficient or necessary relation among any two states of affairs either when the two states of affairs are simultaneous or when one precedes or follows the other in time. This property of indifference with respect to time for the relations of sufficiency and necessity among states of affairs is crucial in the Taylor's argument. Taylor himself gives the following examples to illustrate the relation.

- (17) John is alive from 10 to 11 pm \rightarrow there's oxygen around from 10 to 11 pm.
- (18) John ingests cyanides at 10 am. \rightarrow John dies at 11 am
- (19) John wins the elections in November \rightarrow John has been nominated in May.

In the first case, the two states of affairs are simultaneous. In the second one, the state of affairs which ensures the other one occurs earlier in time. In the last one, the state of affairs which is sufficient for the other one

occurs later in time. Taylor underlines that these relations don't have a logical nature. Ingesting cyanides does not logically imply death. Things are so in the actual world, i.e. given the prevailing circumstances and laws of nature occurring in our world. For instance, there are possible worlds where cyanides does not lead to death.

The sixth presupposition, the one about the passage of time, might be illustrated as follows. Suppose it is not within my power to speak Finnish at T, because I have never studied it and suppose also that I then learn it within two years. So, I can't speak Finnish at T, but I can at a later time. Sure enough, learning Finnish takes time. But it's not the fact that time passed that explains why I acquired this power. I acquired this power or ability because of the studies I have done, not because time has passed. I have to confess I don't understand why Taylor needs his sixth assumption. It seems to me that his argument does not rest upon this innocuous and plausible assumption.

The fifth assumption is the most important. It appears to be innocuous. In fact, it is mostly because of this assumption that we are able to reach the fatalistic conclusion. The assumption states that no agent can perform an act at a time if there is lacking at least one condition necessary, at the same or any other time, for the occurrence of that act. For instance, I *can't* buy a car if I don't have enough money, because having enough money is a necessary condition for the purchase. Or, I *can't* survive if there's no oxygen around, because the presence of oxygen is necessary for life. The list goes on. I *can't* see if I am blind. I *can't* have a drink with Socrates if Socrates is dead. These seem to be all straightforward and uncontroversial claims. Taylor argues that these claims are true in virtue of the principle captured by his fifth presupposition.

We are now in a position to analyze the argument. First of all, Taylor invites us to consider the following situation. Suppose that if today I read on a completely reliable newspaper an headline which states that there has been a sea battle yesterday, then yesterday there has been a sea battle, whereas if I read another kind of headline on the same newspaper, then yesterday no sea battle occurred. That is, the state of affairs of me reading today an headline which states that there has been a sea battle is *sufficient* for the occurrence of a sea battle yesterday, and, reading another kind of headline is sufficient for the sea battle non-occurrence. We name the states of affairs of me reading the first kind of headline and me reading the second kind h and h' , whereas b and b' stand for the sea battle occurrence and non-occurrence. We then have

$$(20) \quad h \rightarrow b$$

$$(21) \quad h' \rightarrow b'$$

Imagine then that I am about to perform the act of reading the newspaper. If someone were to claim that it is within my power to read the first kind of headline *and* it is within my power to read the second kind, which we may express modally like this

$$(22) \quad \diamond h \wedge \diamond h'$$

then, everybody would agree that he or she has said something false. The battle either occurred yesterday or did not, and the newspaper headline reliably depends on what occurred yesterday. Furthermore, what I am about to read depends on something which is now beyond my control, namely what happened yesterday. If it were within my power to bring about the state of affairs h , *and* it were within my power to bring about h' , necessarily it would be the case that it is within my power to bring about a state of affairs sufficient for the past to change, given that we have (20) and (21). Hence, (22) must be false. Taylor argues that the reason why (22) is false can be explained by an argument which adopts the six presuppositions we have been talking about.

Fatalism toward the past

- (23) Either a sea battle occurred yesterday, or no sea battle occurred yesterday.
- (24) If a sea battle occurred yesterday, then there is lacking a necessary condition for the act of reading the second kind of headline. So, it is not within my power to read the second kind of headline.
- (25) If no seabattle occurred yesterday, then there is lacking a necessary condition for the act of reading the first kind of headline. So, it is not within my power to read the first kind of headline.
- (26) Hence, Either it is not within my power to read the second kind of headline or it is not within my power to read the first kind of headline.

The conclusion is a proper denial of (22). In fact, (26) states that at least one of the two conjuncts of (22) is false. It is easy to see that the argument just given employs Taylor's assumptions. Premise (23) is just an instance of LEM. Premise (24) utilizes the assumptions on the nature of the sufficient and necessary relations together with the fifth assumption. By reasoning under the hypothesis that a sea battle occurred yesterday, we have the result that a condition necessary for reading the second kind of headline is lacking. Hence, via the fifth assumption, it is not within my power to

read the second kind of headline. Premise (25) is of course analogous, and then, by assuming classical logic, the conclusion logically follows, because the argument has the form of a disjunctive dilemma.

To sum up, the argument for fatalism toward the past appears to be correct, and everybody should be willing to assent to it. The problem is, and Taylor insists on this, that a parallel argument can be given when there is a lack of a necessary condition in the future rather than in the past like in the headline case.

Here is the further example Taylor provides. Suppose I am an admiral who is about to issue an order to my fleet. If I issue a certain order, then a sea battle will occur tomorrow. If I don't or I give another order, no sea battle will occur tomorrow. If we name o , o' , b and b' respectively the states of affairs of me issuing the battle order, not issuing it, and the sea battle occurrence and non occurrence we have that

$$(27) \quad o \rightarrow b$$

$$(28) \quad o' \rightarrow b'$$

We can now provide the Taylor argument to show that only one between two alternative courses of action can be within my power. The argument is the same as the one gave above, we just need to change some tenses.

Fatalism towards the Future

- (29) Either a sea battle will occur tomorrow, or it won't.
- (30) If a sea battle will not occur tomorrow, then there is lacking a necessary condition for the act of issuing the battle order. So, it is not within my power to issue the battle order.
- (31) If a sea battle will occur tomorrow, then there is lacking a necessary condition for the act refraining from issuing the battle order. So, it is not within my power to not issue the battle order.
- (32) Hence, either it is not within my power to issue the battle order or it is not within my power to refrain from issuing it.

Again, the conclusion states that among two mutually exclusive and exhaustive courses of event, only one of them is within my power and open to me. Needless to say, this kind of argument can be generalized to any agent and any act, every time we have the kind of sufficient and necessary relations Taylor uses.

Does the conclusion stated by Taylor's argument amount to the kind of fatalism we are interested in? We have seen that the core fatalist claim has

a conditional form. It states that if something happens, then it is necessary that it happens. Or, equivalently, if something does not happen, then it is impossible that it does. (32) instead has a disjunctive form. The answer to the previous question is in the affirmative. (32) can be transformed to the conditional fatalistic claims with the help of LEM. That is, either I give the sea battle order or I don't. Suppose I do. Clearly then, it is within my power to give the order. This means that the first disjunct of (32) is false and the second must be true. The second disjunct states that it is not within my power to refrain from issuing the order, which is equivalent to saying that issuing the order is necessary. So, If I issue the sea battle order, I necessarily do so. Suppose instead that I don't issue the order. Then, it is within my power to refrain from issuing the order. Under this hypothesis, the second disjunct of (32) is false and thus the first must be true. And the first just states that it impossible for me to issue the sea battle order. Hence, If I don't issue it, it is impossible that I do so. We could state the same result in a more formal way. That is, we can show in modal formal discourse how the scheme $A \rightarrow \Box A$ and $\neg A \rightarrow \neg \Diamond A$ follows from the schema $\neg \Diamond A \vee \neg \Diamond \neg A$, whose Taylor's argument conclusion (32) is an instance.

1	$\neg \Diamond A \vee \neg \Diamond \neg A$	(32)
2	$\Diamond A \rightarrow \Box A$	1, PL
3	$A \rightarrow \Diamond A$	T \Diamond
4	$A \rightarrow \Box A$	2,3
5	$\Box A \rightarrow A$	T
6	$\Diamond A \rightarrow A$	2,5
7	$\neg A \rightarrow \neg \Diamond A$	6, PL

Step (1) is the schema Taylor's argument is an instance of, which is logically equivalent to step (2). Steps (3) and (5) are instances of the modal schema T and its dual, a schema which a fatalist better accept, in order to avoid bizarre results¹¹. Steps (4) and (7) are what we were looking for. Hence, Taylor's argument conclusion can be reduced to a more standard fatalistic form. That is, what happens is necessary and what does not is impossible.

¹¹It will be clarified in section 1.3 why this must be so.

Taylor's Argument Modal Structure

In order to provide a better understanding of Taylor's argument, I will give its modal structure. I go back to the modal language L already introduced in section 1.2.1. In particular we should bear in mind that all the sentences of L contain a reference to a time and express tenseless propositions with a time index in their content. The notion of necessity, and consequently of possibility, at stake within the fatalist debate is the notion of *historical* necessity. Fatalism is an agent-centered doctrine, in the sense that it is interested in what agents *can* or *cannot* do at a certain time. Hence, it is mainly focused on those sentences of L such that they are about an agent doing something at a certain time — we can name this subset of L, L'. It is natural to interpret historical necessity as applied to atomic sentences of L' with the modal notion of 'being within one's power at time T'. For instance, a formula like $\Diamond p$, where p expresses the proposition that *Jones RAISE his right hand at T*, has to be read as "it is within Jones's power to raise his right hand at T".

That being said, we can modally express Taylor's fifth assumption, that is

Fifth, we presuppose that no agent can perform any given act if there is lacking, at the same or any other time, some condition necessary for the occurrence of that act.

. We modally express it through the following schema

$$(5^{th}) ((A \rightarrow B) \wedge \neg B) \rightarrow \neg \Diamond A$$

which is treated as a theorem by Taylor. Then the argument for fatalism toward the future can be modally given as follows.

1	$o \leftrightarrow \neg o'$																					
2	$o \rightarrow b$																					
3	$\neg o \rightarrow \neg b$																					
4	$b \vee \neg b$		LEM																			
5	<table style="border-left: 1px solid black; border-right: 1px solid black; width: 100%; border-collapse: collapse;"> <tr> <td style="width: 5%; text-align: right; vertical-align: top;">5</td> <td style="padding-left: 5px; vertical-align: top;">b</td> <td style="width: 10%;"></td> <td style="width: 80%;"></td> </tr> <tr> <td style="text-align: right; vertical-align: top;">6</td> <td style="padding-left: 5px; vertical-align: top;">$b \leftrightarrow \neg \neg b$</td> <td></td> <td style="vertical-align: top;">PL</td> </tr> <tr> <td style="text-align: right; vertical-align: top;">7</td> <td style="padding-left: 5px; vertical-align: top;">$((\neg o \rightarrow \neg b) \wedge \neg \neg b) \rightarrow \neg \diamond \neg o$</td> <td></td> <td style="vertical-align: top;">5^{th}</td> </tr> <tr> <td style="text-align: right; vertical-align: top;">8</td> <td style="padding-left: 5px; vertical-align: top;">$\neg \diamond \neg o$</td> <td></td> <td style="vertical-align: top;">3,5,6</td> </tr> <tr> <td style="text-align: right; vertical-align: top;">9</td> <td style="padding-left: 5px; vertical-align: top;">$b \rightarrow \neg \diamond \neg o$</td> <td></td> <td style="vertical-align: top;">5 — 8</td> </tr> </table>	5	b			6	$b \leftrightarrow \neg \neg b$		PL	7	$((\neg o \rightarrow \neg b) \wedge \neg \neg b) \rightarrow \neg \diamond \neg o$		5^{th}	8	$\neg \diamond \neg o$		3,5,6	9	$b \rightarrow \neg \diamond \neg o$		5 — 8	Hyp
5	b																					
6	$b \leftrightarrow \neg \neg b$		PL																			
7	$((\neg o \rightarrow \neg b) \wedge \neg \neg b) \rightarrow \neg \diamond \neg o$		5^{th}																			
8	$\neg \diamond \neg o$		3,5,6																			
9	$b \rightarrow \neg \diamond \neg o$		5 — 8																			
6	$b \leftrightarrow \neg \neg b$		PL																			
7	$((\neg o \rightarrow \neg b) \wedge \neg \neg b) \rightarrow \neg \diamond \neg o$		5^{th}																			
8	$\neg \diamond \neg o$		3,5,6																			
9	$b \rightarrow \neg \diamond \neg o$		5 — 8																			
10	<table style="border-left: 1px solid black; border-right: 1px solid black; width: 100%; border-collapse: collapse;"> <tr> <td style="width: 5%; text-align: right; vertical-align: top;">10</td> <td style="padding-left: 5px; vertical-align: top;">$\neg b$</td> <td style="width: 10%;"></td> <td style="width: 80%;"></td> </tr> <tr> <td style="text-align: right; vertical-align: top;">11</td> <td style="padding-left: 5px; vertical-align: top;">$((o \rightarrow b) \wedge \neg b) \rightarrow \neg \diamond o$</td> <td></td> <td style="vertical-align: top;">5^{th}</td> </tr> <tr> <td style="text-align: right; vertical-align: top;">12</td> <td style="padding-left: 5px; vertical-align: top;">$\neg \diamond o$</td> <td></td> <td style="vertical-align: top;">2, 10</td> </tr> <tr> <td style="text-align: right; vertical-align: top;">13</td> <td style="padding-left: 5px; vertical-align: top;">$\neg b \rightarrow \neg \diamond o$</td> <td></td> <td style="vertical-align: top;">10 — 12</td> </tr> <tr> <td style="text-align: right; vertical-align: top;">14</td> <td style="padding-left: 5px; vertical-align: top;">$\neg \diamond \neg o \vee \neg \diamond o$</td> <td></td> <td style="vertical-align: top;">4,9,13</td> </tr> </table>	10	$\neg b$			11	$((o \rightarrow b) \wedge \neg b) \rightarrow \neg \diamond o$		5^{th}	12	$\neg \diamond o$		2, 10	13	$\neg b \rightarrow \neg \diamond o$		10 — 12	14	$\neg \diamond \neg o \vee \neg \diamond o$		4,9,13	Hyp
10	$\neg b$																					
11	$((o \rightarrow b) \wedge \neg b) \rightarrow \neg \diamond o$		5^{th}																			
12	$\neg \diamond o$		2, 10																			
13	$\neg b \rightarrow \neg \diamond o$		10 — 12																			
14	$\neg \diamond \neg o \vee \neg \diamond o$		4,9,13																			
11	$((o \rightarrow b) \wedge \neg b) \rightarrow \neg \diamond o$		5^{th}																			
12	$\neg \diamond o$		2, 10																			
13	$\neg b \rightarrow \neg \diamond o$		10 — 12																			
14	$\neg \diamond \neg o \vee \neg \diamond o$		4,9,13																			

Step 1-2-3 are justified by how the order-sea battle example is built. Then, the argument starts with an instance of LEM. The argument is valid in virtue of its disjunctive dilemma form. The last step is Taylor's argument conclusion for fatalism toward the future.

Standard responses

Most of objections toward Taylor's argument focus on the schema $\diamond A \rightarrow A$, which as we have seen is derivable via the schema T from (32), i.e. Taylor's argument conclusion in the argument for fatalism towards the future. For instance, Saunders (1962) argues as follows in one of his replies to Taylor.

in order to enjoy whatever situation we might desire, we need not go to the trouble to bring them about; or, rather, to bring them about we have only to acquire the power to bring them about (p.67)

The bottom line of this passage is that it is bizarre to hold such a view. It is bizarre to claim that a mere power to perform an act ϕ is sufficient

for the occurrence of that act. Admittedly, a modal fatalist thinks so and this is indeed a strange result. But, at least two things can be said as a rejoinder. First of all, to point out that a doctrine is strange or counter-intuitive does not by any means harm the consistency or truth of that doctrine. Many metaphysical doctrines, and scientific theories as well, have counter-intuitive bizarre results. The consequence of modal fatalism pointed out by Saunders is strange, just because modal fatalism is strange. Again, this should not count as a good reason to refute modal fatalism. Second of all, we have seen above that the schema $\Diamond A \rightarrow A$ and the schema $\neg\Diamond\neg A \vee \neg\Diamond A$ are intertwined. That is, one can derive the former from the latter via T, and vice versa¹². The schema $\neg\Diamond\neg A \vee \neg\Diamond A$ is just the idea that the future is closed, because it amounts to the thought that among two alternative courses of events only one is open. Rejecting $\Diamond A \rightarrow A$ implies a rejection of $\neg\Diamond\neg A \vee \neg\Diamond A$ as well. That is, if Saunders's objection were conclusive, we would have an argument for the *openness* of the future which lies solely on the intuitive weirdness of a statement which is equivalent to the closeness of the future. Such an argument clearly won't do, and thus neither Saunder's objection does. Metaphysical theories should rest on something more than intuitive appeal.

There is another and more compelling objection which comes again from Saunders. Again, the objection is against the truth of the schema $\Diamond A \rightarrow A$ and Taylor's fifth presupposition. Here is Saunders (1962).

Now to point out that this supposition is, indeed, erroneous. My knocking upon a thin wooden door is a sufficient condition for the door shacking. Hence the door's shaking is a necessary condition for my knocking upon the door. But the door's shaking is not a necessary condition for my *ability* to knock upon the door.... I may decide not to knock upon and the door may not shake, but it does not follow that I did not have it in my power to knock. (p.54)

We must first notice that Saunders here implicitly equates the notion of having an ability and the notion of having it within someone's power to. This is so because Saunders switches between the two expressions in the quoted passage, thereby treating them as they were one and the same. I think this is the main flaw of Saunder's reply to the Taylor's argument. But before arguing for this, let us see how Saunders objection can be analyzed. It attacks the fifth presupposition, which is employed by Taylor in the

¹²We have seen above as $\Diamond A \rightarrow A$ can be derived from $\neg\Diamond\neg A \vee \neg\Diamond A$ via T. How $\neg\Diamond\neg A \vee \neg\Diamond A$ can be derived from $\Diamond A \rightarrow A$ will be shown in section 1.3.2.

premises (30) and (31) of his argument for fatalism toward the future. Let us take (30)

- (30) If a sea battle will not occur tomorrow, then there is lacking a necessary condition for the act of issuing the battle order. So, it is not within my power to issue the battle order.

Saunders would reason as follows about this premise. He would say that the occurrence of the battle tomorrow is a necessary condition for my giving the battle order today. But, he would go on, the battle tomorrow is *not* a necessary condition for my *ability* to issue the battle order. Hence, under the hypothesis that a sea battle does not occur tomorrow, it follows only that I do not issue the order, not that I don't have the ability to do so. It might be the case that I do not issue the battle order, and yet I have the ability to issue it — because, for instance, I know how to issue orders, I know the proper words to use, my mouth muscles are properly functioning and so on. In such case, even if a sea battle does not occur tomorrow and hence I do not issue the battle order, I have the ability to issue it. Thus, by equating the talk about powers and the one about abilities, in such case it would be within my power to issue the battle order. This of course, provides also a counterexample to $\Diamond A \rightarrow A$. If things are as Saunders says, it might be that it is within my power to issue the order and yet I don't issue it. This way, the premise (30) — and (31) as well — turns out to be false, and thus Taylor's argument for fatalism toward the future is blocked.

Taylor has an easy rejoinder against the kind of reasoning employed by Saunders. Taylor observes that if Saunders objection against the argument for fatalism toward the future is sound, then we will also have good reasons to reject the argument for fatalism toward the past. Let us take one of the premises of the argument for fatalism toward the past, (24).

- (24) If a sea battle occurred yesterday, then there is lacking a necessary condition for the act of reading the second kind of headline. So, it is not within my power to read the second kind of headline.

Now, Taylor observes, I certainly have the *ability* to read an headline which states that no sea battle occurred yesterday, even though a sea battle did occur yesterday and the newspaper is completely reliable. This is so because I know how to read, my eyes and my brain are properly working, and so on. So, by following the kind of reasoning Saunders employed and by equating the talk about powers and abilities as he does, it might be the case that when a sea battle did occur yesterday and the newspaper is reliable, it is within my power to read an headline which states that no sea battle occurred. But, performing such an act is sufficient for there being

peace yesterday. At this point, Taylor invokes what is traditionally called the transfer principle.

Transfer Principle. If it is within an agent x power to ϕ , and ϕ implies ψ , then it is within x 's power to bring about that ψ .

The transfer principle is intuitively strong. If it is within my power to swim, and swimming implies water disturbance in the water, then it is within my power to bring about that there is water disturbance. And, in virtue of the transfer principle, if it is within my power to read an headline which says that no sea battle occurred, when in fact it did, and reading such an headline is sufficient for the non-occurrence of it, then it is within my power to bring about that no sea battle occurred yesterday when in fact it did. And this result is of course absurd. To sum up, Taylor reasons as follows with his rejoinder. If we accept Saunders objection against the argument for fatalism toward the future, then we have a parallel objection against the argument for fatalism toward the past. If so, by denying fatalism toward the future, we end up with the absurd result that it may be sometimes in our power to change what happened in the past.

I find Taylor's rejoinder successful against Saunders' objection. And, more can be said about this. We have seen that Saunders implicitly equates the notions of having the ability and have it within one's power. I think those two notions should be clearly distinct. To show why things are so, let us take an example which has been discussed within the literature about Taylor's argument. Take a professional pole-vaulter who often pole vaults 14 feet. Suppose he now finds himself in a room where the ceiling is 10 feet high. We do say that he has the *ability* to pole vault 14 feet, yet it is not *now* within his power to pole vault 14 feet in a room 10 feet high. This should suffice to show that, contrary to what Saunders implicitly assumed, there is a distinction to be made here between the two notions. Powers and abilities behave differently. Abilities generally depend on things which are *internal* with respect to the agent. I can, in the ability sense, walk because my human body allows me to do so. I can read, because I have been trained to do so. I can swim if I have successfully taken swimming classes. Rather, what is and what is not within my power, depend surely on my abilities, but also on *external* circumstances. I can't, it is not within my power to, walk if I am tied up in a chair. I can't read if there is no light. I can't swim if there is no water around. Moreover, agents usually keep the abilities they have throughout their whole life, once they have learned them. Unless something really damaging happens to me, I am likely to keep my ability to read for the rest of my life. On the contrary, powers

are continuously lost and gained through time, because they depend also on external circumstances, which quite obviously constantly change too throughout time. It is clear that what is at stake in the fatalist debate is what is possible and what is not possible for an agent at a specific time under specific external circumstances. Hence, the notion of having within one's power is what we should be interested in.

There is another objection which can be raised against Taylor argument, and again it has to do with his fifth presupposition. We have seen that we can treat it modally as follows.

$$(5^{th}) ((A \rightarrow B) \wedge \neg B) \rightarrow \neg \Diamond A$$

We should note that this schema is valid in few models. It turns out to be valid just in those models where the accessibility relation has the property of vacuity.

Vacuity. If $\alpha R \beta$, then $\alpha = \beta$

And the Taylor's argument conclusion, namely $\neg \Diamond A \vee \neg \Diamond \neg A$ as well is valid in all models where the relation of accessibility has that property of vacuity. So, one might argue that Taylor's argument begs the question, because the fifth assumption already has the fatalistic conclusion in its pockets. However, I don't think this would be a fair objection. Modal logic is just a means to regiment our discourses, and it must be neutral with respect to our metaphysical preferences. And, the fact that fatalism follows from Taylor fifth — together with the other — assumptions is precisely what Taylor wanted to prove. He has independent reasons to think that his fifth presupposition should be accepted, for instance the fact that it explains why the argument for fatalism toward the past is correct.

Another possible move one could make has to do with the notion of a necessary condition for performing an act. One might argue that necessary conditions for performing an act always lie in the past with respect to the performance of that act. This way, the argument for fatalism toward the future would be refuted, whereas the argument for fatalism toward the past would be retained. Indeed, a very good results. Yet, such a move would be highly problematic for at least two reasons. First of all, it seems to be *ad hoc*. Whether the future must be treated like the past with respect to its openness or closure is precisely what is at stake in the fatalism/anti-fatalism debate. Saying that there aren't future necessary conditions for a present performance of an act seems to be justified either only by our pre-theoretical attitudes toward the future or by the fact that we don't *want*

modal fatalism to hold with respect to the future. In the former case, it would be just a matter of intuitions, which a fatalist could after all not share. In the latter case, this response would be just a rejection and not a refutation. Second of all, by saying that there aren't future necessary conditions for a present performance of an act we are equivalently saying that a present performance of an act can never be sufficient for the occurrence of a future state of affairs. Again, a conclusion which seems worse than fatalism.

To sum up, I find Taylor's argument powerful and compelling. Its nature is surely puzzling and as we showed, it is not easy to come up with good reasons to block the argument. All the responses we have been talking about suffer of flaws. I will argue in section 1.3.3 that Taylor's argument proves too much, and thus it cannot be entirely trusted.

1.3 Again on Definitions

In this section I am going to discuss to a greater extent the modal claims which constitute modal fatalism. Before doing so, I introduce another bit of modal logic. Ultimately, I will show what I think is wrong with the Main Argument and Taylor's.

1.3.1 Another bit of Modal Logic

We introduced in section 1.2.1 some modal logic concepts. In this one we are going to say something more about validity and the accessibility relation. I will rely again on Chellas (1980). We have seen that there is an interesting notion of validity. A schema may be valid in a class of models — that is, the schema is true in all possible worlds in all models which belong to that class. Classes of models can be determined via the nature of their accessibility relation. As we have seen, the accessibility relation R in a model tells us what worlds are accessible to other worlds in that model. Relations may have properties. For instance, a relation R in a model M is said to be transitive if and only if for any α , β and γ , it is the case in M that if $\alpha R \beta$ and $\beta R \gamma$, then $\alpha R \gamma$. The class of all models where R is transitive is the set of all models such that their accessibility relation satisfies transitivity. A model where R is, say transitive, is itself said to be transitive. The class of transitive models is the set of all transitive models. What is philosophically interesting is that different accessibility relations validate different scheme.

It turns out that there is a general accessibility relation, k,l,m,n -incestuality, which provides general results in terms of what schemas are validated by

R's properties.

k,l,m,n-incestuality A relation R in a model M is k,l,m,n-incestual if and only if for any α, β and γ in M, if $\alpha R^k \beta$ and $\alpha R^m \gamma$, then there is a δ in M such that $\beta R^l \delta$ and $\gamma R^n \delta$.

By $\alpha R^n \beta$, with $n \geq 0$, we mean that the world β can be reached from α in n steps, via the accessibility relation. For instance, $\alpha R^2 \beta$ means that there is a world γ such that $\alpha R \gamma$ and $\gamma R \beta$. $\alpha R^0 \beta$ means that α is β . R^1 is simply R itself. Similarly, we can write formulas like $\Box^n A$ to mean that the box sign is repeated n times. For instance, $\Box^2 A$ means $\Box \Box A$. Truth conditions for such formulas are as follows. $\Box^n A$ is true in the model M at the world α iff for every β such that $\alpha R^n \beta$, A is true at β in M. $\Diamond^n A$ is true in the model M at the world α iff for some β such that $\alpha R^n \beta$, A is true at β in M.

It can be proven that the schema $G^{k,l,m,n}$ is valid in the class of k,l,m,n-incestual models¹³

$$G^{k,l,m,n}. \Diamond^k \Box^l A \rightarrow \Box^m \Diamond^n A.$$

This gives the general result that if we are given any schema which can be expressed by filling numbers in $G^{k,l,m,n}$, we automatically have a relation of accessibility which is sufficient to validate that schema. For instance, the schema 4

$$4. \Box A \rightarrow \Box \Box A$$

is $G^{k,l,m,n}$ where $l=1, m=2$ and $k=n=0$. This means that a 0,1,2,0-incestual model validates the schema 4.

0,1,2,0-incestuality A relation R in a model M is 0,1,2,0-incestual if and only if for any α, β and γ in M, if $\alpha R^0 \beta$ and $\alpha R^2 \gamma$, then there is a δ in M such

¹³Here is an informal proof. Let α be an arbitrary world in an arbitrary model M which is k,l,m,n-incestual. We then assume the antecedent of $G^{k,l,m,n}$. It follows that there is a β such that $\alpha R^k \beta$ and it is true at β in M that $\Box^l A$. We wish to prove that it is true at α in M that $\Box^m \Diamond^n A$. That is, we wish to prove that for any world such that αR^m it, it is the case in that world that there is at least a world n steps away from it where A is true. So, we pick an arbitrary world, call it γ such that $\alpha R^m \gamma$. Given that $\alpha R^k \beta$ and $\alpha R^m \gamma$, and the model is k,l,m,n-incestual, there is a world, say δ , such that $\beta R^l \delta$ and $\gamma R^n \delta$. Given that $\Box^l A$ is true at β in M, every world l steps away from β is a world where A holds. Hence, it is true at δ that A. We wished to prove that at our arbitrary γ there is at least one world n steps away from γ where A holds. δ can serve such a purpose, given that k,l,m,n incestuality states that $\gamma R^n \delta$ and A holds at δ .

that $\beta R^1 \delta$ and $\gamma R^0 \delta$.

which is for any α, β and γ in M , if $\alpha = \beta$ and $\alpha R^2 \gamma$, then there is a δ in M such that $\beta R \delta$ and $\gamma = \delta$. Which is, if $\alpha R \gamma$ in two steps, then there is a world δ in the model, which is γ itself, which is seen by α in just one step. And this is just a way to express transitivity, i.e. the idea that if $\alpha R \beta$ and $\beta R \gamma$, then $\alpha R \gamma$. Hence, transitive models validate the schema 4.

1.3.2 Modal Fatalism Principles

We have already seen that modal fatalism has to do with what agents *can* or *cannot* do. Its formulation employs modal notions and the specific sense of “can” at stake is what is possible and what is not possible for an agent at a specific time under specific circumstances. We have seen in section 1.2.3 that the notion of having within one’s power seems to capture this idea, and we expressed this notion with the modal symbol \diamond . For instance, a formula like $\diamond p$, where p expresses the proposition that *Jones RAISE his right hand at T*, has to be read as “it is within Jones’s power to raise his right hand at T”. Under this interpretation of the \diamond symbol in terms of powers, it is natural to take the \square symbol as the dual of \diamond . That is, we have as a theorem in L $\square A \leftrightarrow \neg \diamond \neg A$. $\square p$, which is equivalent to $\neg \diamond \neg p$, has to be read as “it is not the case that it is within Jones’s power to refrain from raising his right hand at T”.

We characterized in section 1.1.2 modal fatalism as the idea that for any agent, for any act and for any time, the agent has not within his or her power to act otherwise than he or she actually does. Another definition of modal fatalism is the idea that whatever happens, it is necessary that it happens and what does not, it is impossible that it happens. We can start from this last definition and give it modally as the schema T_c .

$T_c. A \rightarrow \square A$

which is a theorem if and only if its dual $T_c \diamond$ is a theorem¹⁴.

¹⁴If the schema T_c is a theorem, then $\neg A \rightarrow \square \neg A$ is a theorem, which is logically equivalent to $\neg \square \neg A \rightarrow A$, which in turn is equivalent to $\diamond A \rightarrow A$ in virtue of the definition of \diamond . As for the other side, if $T_c \diamond$ is a theorem, then $\diamond \neg A \rightarrow \neg A$ is a theorem, which is logically equivalent to $A \rightarrow \neg \diamond \neg A$, which in turn is equivalent to $A \rightarrow \square A$ in virtue of the meaning of \square .

$$T_c \diamond. \diamond A \rightarrow A$$

T_c , or equivalently $T_c \diamond$, taken as theorems deliver in modal terms the idea that no agent has ever the power to do otherwise than what he or she in fact does. For instance, if A is, say, "Jones raises his right hand at T ", and it is the case that A , it follows via T_c by modus ponens that it is not the case that it is within Jones' power to refrain from raising his right hand at T . That is, it is not within John's power to do otherwise than raising his right hand at T . It must also be noted that T_c is what the Main Argument aims to prove. The reader should remember that the Main Argument starts with claims like "Suppose Jones raises his right hand at T " and it eventually conclude that if so, it is not within Jones' power to refrain from raising his right hand. Moreover, from T_c , or equivalently $T_c \diamond$, taken as theorems, we can derive Taylor's argument conclusion in its form $\neg \diamond A \vee \neg \diamond \neg A$.

1	$A \vee \neg A$	PL
2	$\diamond A \rightarrow A$	$T_c \diamond$
3	$\neg A \rightarrow \neg \diamond A$	2, PL
4	$\diamond \neg A \rightarrow \neg A$	$T_c \diamond$
5	$A \rightarrow \neg \diamond \neg A$	4, PL
6	$\neg \diamond A \vee \neg \diamond \neg A$	1, 3, 5, PL

We are now in a position to exploit the k,l,m,n -incestuality we have been talking about in the previous section. In fact, T_c has the form of $G^{k,l,m,n}$ where $m=1$ and $k=l=n=0$. $0,0,1,0$ -incestual models will validate T_c .

0,0,1,0-incestuality A relation R in a model M is $0,0,1,0$ -incestual if and only if for any α, β and γ in M , if $\alpha R^0 \beta$ and $\alpha R^1 \gamma$, then there is a δ in M such that $\beta R^0 \delta$ and $\gamma R^0 \delta$.

$0,0,1,0$ -incestuality informally means that if a world α R s something, say γ , then there is a world δ , which is α itself in virtue of the identities, which is seen by α . This property is generally named vacuity.

Vacuity. For every α and for every β , If $\alpha R \beta$, then $\alpha = \beta$.

That is, either α is blind or at most it sees itself. Such models validate the fatalistic formulas T_c , $T_c\Diamond$ and Taylor's conclusion¹⁵.

It must be noted that vacuity by itself can produce bizarre results. For instance, at some worlds in vacuous models formulas like $A \wedge \neg\Diamond A$ — the impossible happens — $\neg\Diamond A \wedge \neg\Diamond\neg A$ — some act is neither within one's power, nor is not-doing it — $\Box A \wedge \Box\neg A$ — two mutually exclusive events are both necessary — are true. All these bizarre results are due to the fact that vacuity does not guarantee that a world sees at least one world. It leaves open the possibility of blind worlds, and in those world those three bizarre formulas turn out to be true. A modal fatalist better not have those formulas. The right move to do is to accept also reflexivity.

Reflexivity. For every α , $\alpha R\alpha$.

It can be shown that reflexivity is 0,1,0,0-incestuality. Hence, reflexive models validate T and its dual $T\Diamond$

T. $\Box A \rightarrow A$ **$T\Diamond$.** $A \rightarrow \Diamond A$

These two principles are quite natural to accept for a modal fatalist. The former says that whatever is necessary it happens, while the latter says that whatever happens is possible. Under the interpretation of modal symbols in terms of powers, T says that if it not the case, say, that it is within my power to refrain from ϕ ing at T, then I ϕ at T. $T\Diamond$ says that if I ϕ at T, then it is within my power to ϕ at T. These claims are uncontroversial and not in need of an explanation¹⁶

To sum up, vacuous models captures the modal fatalist principles T_c ,

¹⁵It is interesting to note that, whereas vacuity validates T_c , $T_c\Diamond$, and Taylor's conclusion, there is a relation of accessibility which validate Taylor's conclusion but not T_c and $T_c\Diamond$. In fact, Taylor's conclusion $\neg\Diamond A \vee \neg\Diamond\neg A$ is equivalent to $G^{1,0,1,0}$ in virtue of propositional logic, that is $\Diamond A \rightarrow \Box A$. This means that 1,0,1,0-incestuality validates Taylor's conclusion. The property of 1,0,1,0-incestuality is for any α , for any β , for any γ , If $\alpha R\beta$ and $\alpha R\gamma$ then for some δ , $\beta=\delta$ and $\gamma=\delta$. That is, a world sees at most one world. This shows why if T_c , or equivalently its dual $T_c\Diamond$, is a theorem, then also Taylor's conclusion is a theorem, whereas the converse does not hold. T_c , or equivalently its dual $T_c\Diamond$, cannot be derived from Taylor's conclusion, unless T is also a theorem.

¹⁶We should also note that a modal fatalist who accepts T, $T\Diamond$, T_c and $T_c\Diamond$ has as a consequence a collapse of modalities. From T and $T\Diamond$ we have $\Diamond A \rightarrow \Box A$, in virtue of the transitivity of conditionals. From T_c and $T_c\Diamond$ we have $\Box A \rightarrow \Diamond A$, again in virtue of the transitivity of conditionals. In few steps we realize that then we have that $\Box A \leftrightarrow A \leftrightarrow \Diamond A$. This collapse might seem unpleasant. In fact, it is not and it might count as an interesting modal logic. Again, it appears to be strange just because modal fatalism is strange. A fatalist might not agree on our judgement about what is strange and what is not.

$T_c \diamond$ and Taylor's argument conclusion. A modal fatalist should also have reflexivity in order to capture uncontroversial claims about powers and to avoid the bizarre results we have seen.

1.3.3 What is wrong with the Main Argument and Taylor's

In the previous sections we talked about the Main Argument and Taylor's one. Here I want to say what I think is a common flaw for both arguments. Both of them conflate deterministic and non-deterministic scenarios. We can loosely characterize determinism as the theory which apply or not apply to systems of Laws of Nature. A system of Laws of Nature is deterministic if and only if the system is such that a single instantaneous state of the world together with the Laws of Nature provides an unique way things are at *all* other times. That is, a single instantaneous way the world is has an unique past and future history which is in accordance with Laws of Nature. On the contrary, non-deterministic systems of Laws of Nature allow for more than one history which is compatible with a single instantaneous state of the world.

Here is a simple case of a non-deterministic scenario. I borrow the following example about Quantum Mechanics from Albert (1992). We need to tell a short story about electrons and their color and hardness¹⁷. Hardness and color are two measurable properties of electrons. It happens that those properties can assume only two possible values. Electrons can be either black or white, and can be either hard or soft. We have devices which are able to perfectly measure those properties. We call these devices hardness box and color box. These property measurements are repeatable. An electron which is measured to be, say, hard at a time, will be measured to be hard later on, if no tampering with it occurs between the two measurements. Also, there are no correlations between the two properties. Half of the electrons which are measured to be black and have then their hardness measured are found to be hard, whereas half of them are measured to be soft. Similarly for the white ones, and for the color of hard and soft ones. There is an experiment which can be carried out which has surprising results. We need three measurement boxes, two color boxes and one hardness box. We first feed a set of electrons in the first color box. Half of the electrons will be measured white and half of them black. We

¹⁷Of course electrons do not have the properties of color and hardness. Albert uses those terms to talk about incompatible properties like the angular momentum with which an electron spins around an axis and the angular momentum with which it spins around another axis. Talking about hardness and color just makes it easier to illustrate what is going on.

take the white ones and we feed them in the hardness box. Half of them will be measured to be hard, and half of them soft. We now take those electrons which were measured to be white by the first color box *and* measured to be soft by the second hardness box and we feed them in the third color box. We expect that those electrons should be measured to be white, because that is how their color was measured by the first color box. As a matter of fact, this is not what happens. Half of them will be measured to be white, and half of them will be measured to be black. The hardness box in the middle of the process seemed to randomize the color of white electrons which were fed in. This is an example of the uncertainty principle. Electrons with a definite color don't have a definite hardness property, and viceversa. Thus, we can never say that an electron has a certain color property *and* a certain hardness property. It can be proved that this is not a matter of our ignorance or a lack of proper measurement devices. This is how things fundamentally are. The point that interests us here is that we have a case where the exact same initial conditions give rise to different outcomes. We take the electrons which were measured to be white by the first box and soft by the second one. As we have said, half of them will be measured white and half of them black. It can also be proven that there are no hidden variables here which would explain which electrons will be measured white and which of them will be measured black. If you take an electron which is about to enter the third box, two outcomes are possible, i.e. two possible future continuations. It will either be measured to be white or to be black. Same initial conditions, two possible outcomes.

We now apply the Main Argument and Taylor's one to this electron story. Let us say that at T the electron is about to enter the third box, and at T_1 is measured to be black. Now, we apply to this scenario the Main Argument.

The (correct) Electron Main Argument

Suppose the electron is measured to be black at T_1 . From this it follows that the proposition that *the electron BE black at T_1* is true at T_1 . If such proposition is true at T_1 , then it is always true (i.e. true at all times). Then,

- (33) Necessarily, it was true at T that *the electron BE black at T_1* .
- (34) Necessarily, if it was true at T that *the electron BE black at T_1* , then the electron is black at T_1 .
- (35) Therefore, it is not the case that it is within the electron's power to be measured *white* at T_1 .

Talking about what is within the electron's power may sound odd and a categorical mistake, given that an electron is an inanimate entity. If so, one can replace it with what is or is not within the electron's capability. What I want to highlight is that according to the Main Argument applied to the electron case, if the electron is measured to be black at T_1 , it was *necessary* at T that it would be measured so. But this is clearly a wrong result. We have experimental evidence that at T it was *possible* for the electron to be measured white instead of black. The Main Argument proves too much. It applies necessity where there should not be any necessity. In other words, by applying necessity to every event, it conflates deterministic and non-deterministic cases.

Taylor's argument suffers the same flaw. We consider again the three boxes case to show why. Suppose a reliable observer is observing the measurement of the third box. If the electron is measured to be black, then the observer believes the electron is black. If the electron is measured to be white, then the observer believes the electron is white. This is what it is to be a reliable observer of the measurement of the third color box. Now, this gives us the kind of sufficient and necessary relations among state of affairs which Taylor exploits in his argument. We now apply Taylor's argument.

Fatalism towards the Future

- (36) Either the observer will believe the electron is black, or she will believe the electron is white.
- (37) If she will believe the electron is black, then there is lacking a necessary condition for the electron to be white. So, it is not within the electron's capability to be white.
- (38) If she will believe the electron is white, then there is lacking a necessary condition for the electron to be black. So, it is not within the electron's capability to be black.
- (39) Hence, either it is not within the electron's capability to be white or it is not within the electron's capability to be black.

Again, the delivered result is wrong. It was possible for the electron to be measured white, *and* it was possible for the electron to be measured black. And this contrasts with the Taylor's argument applied to the electron case.

1.4 Why I am not worried

Fatalism is often seen as an untenable doctrine. The fact that everything that happens is necessary and what does not is impossible seem to imply undesirable consequences. It might seem that any action we decide to do would be in vain. It might seem also that the process of deliberation would be futile. And ultimately, it might seem that we are unable to make sense of the moral sphere. I will argue in the next sections that a reasonable fatalist can resort to strategies to resist these dreadful conclusions. The reader should bear in mind that a reasonable fatalist is a kind of modal fatalist. Modal fatalism amounts to the acceptance of the Powerlessness thesis. Reasonable fatalism accepts the Causal Connectedness thesis as well.

1.4.1 Lazy Argument

One of the main critics against fatalism has to do with the fact that if a future event is fated to happen, then it is pointless to make any effort to bring it about its occurrence or prevent it from happening. In other words, it might seem that fatalism implies that goal-directed actions are pointless. This does violence to our common understanding of how our world works. Usually, idle people do not get almost anything out of their lives, whereas people who make efforts to achieve the results they want do. The Lazy Argument is a classic argument against the doctrine of fatalism, because it tries to move from fatalism toward the counter-intuitive idea that goal-directed actions are pointless. It may be stated as follows, where A stands for a sentence about a future event and ϕ stands for a verb expressing an action predicate.

Lazy Argument schema

- (40) either it is fated that A or it is fated that not- A
- (41) If it is fated that A , then, whether or not you ϕ , A
- (42) If it is fated that not- A , then, whether or not you ϕ , not- A
- (43) Therefore, with regard to A , it is futile to ϕ

The steps just given provide us with an argument schema. By filling the schema with the proper ingredients we end up having instances of the Lazy Argument. By an action being futile with respect to a future event A , we mean that that action would be a waste of time with respect to A . Maybe ϕ could be pleasant or enjoyable for different reasons, but if it's futile with

respect to A, it means that it does not affect whether A will obtain or not. So, even though the word “futile” does not explicitly appear in (41)-(42), whereas it does in the conclusion, we can see that by “whether or not you ϕ ” we mean that ϕ is futile¹⁸. This argument has the form of a disjunctive dilemma and it is valid, hence it holds within classical logic. If the schema is also sound, i.e. all its instances are sound, then a fatalist has a problem, because it would follow that every action we take directed to a goal is futile. Hence, a fatalist has to come up with a strategy. Of course, a *modal* fatalist can't deny the first premise. It is part of his or her doctrine that whatever happens, is fated to happen in the relevant sense. Hence, given that A or not-A will obtain, either A is fated to happen or not-A is. A modal fatalist thus has to challenge the truth of at least one of the two other premises, and I want to argue that reasonable fatalism can. Before seeing how this could be done, it is worth making it clear what kind of conclusion the Lazy Argument purports to prove. The conclusion of the Lazy Argument is not full inertia and inactivity as the most rational course of action. Rather, its conclusion is that between different courses of actions, we should *always* prefer the most pleasant *per se*, i.e. the one which by itself brings the most pleasure, regardless of what its consequences might be. Suppose I want to pass an exam, that is, I want the future event of me passing the exam to occur, and I am wondering whether I should study or watch a TV show. By filling the Lazy Argument schema with the desired event and the two alternative courses of actions, we end up having that with respect to the exam, it is futile to study and it is futile to watch the TV show. That is, *both* actions will not affect the outcome of my future exam. In fact, according to the Lazy Argument, any action I take now would *not* affect the outcome of my future exam. Thus, if the Lazy Argument is correct, full inactivity is not necessarily the most rational thing to do. It is the most rational only if it is the one that brings most pleasure. Rather, I should always prefer the action which gives me immediate pleasure, which in some case might be full inactivity. In the case under consideration, according to the Lazy Argument, I better watch the TV show, because at least I would get joy out of it and I would avoid the pain of studying.

These considerations just made give us an hint of why some instances of the Lazy Argument can be considered fallacious by a reasonable fatalist. The Lazy Argument hinges on the idea that there aren't causal connections

¹⁸Dummett(1964) gave a vivid example of an instance of the Lazy Argument, which was popular during the bombing of London during the Second World War. “Either you are going to be killed by a bomb or you are not going to be. If you are, then any precautions you take will be ineffective. If you are not, all precautions you take are superfluous. Therefore it is pointless to take precautions.” (p.345)

among the occurrence of events in our world. That is, the Lazy Argument implicitly denies the Causal Connectedness claim we have seen in section 1.1.1— the idea that what happens at a time may causally depend on what happens at other times. We have already seen that the kind of fatalism I am interested in, reasonable fatalism, accepts that thesis. So, here is how a reasonable fatalist can reject some instances of the Lazy Argument. First of all, we should note that if a future event *E* is fated to happen, so are all the events which constitute past necessary conditions for the occurrence of *E*. Then, we should note that the occurrence of a necessary condition for the future event *E* may be a cause of the occurrence of *E*, and in this sense would not be futile. We use again the case of the exam, and we fill the Lazy Argument schema in the proper way.

Studying for an exam

- (44) either it is fated that I will pass the exam or it is fated that I will not.
- (45) If it is fated that I will pass the exam, then, whether or not I study, I will pass the exam.
- (46) If it is fated that I will not pass the exam, then, whether or not I study, I will not pass it.
- (47) Therefore, with regard to me passing the exam, it is futile to study.

We assume then that I pass the exam only if I study. Then, premise (45) can be disputed by a reasonable fatalist. First of all, if it is fated that I will pass the exam, then I will pass the exam. But, we are assuming that I pass the exam only if I study. Thus, it must be fated also that I study. Second of all, we can see that there is the proper causal dependence between me studying for the exam and me passing it. Suppose that I study and then I pass the exam. The counterfactual which claims that had I not studied, I would not have passed the exam turns out to be true. My studying made a difference, and brought it about my success in the exam. Hence, it is not futile. In this case, premise (45) turns out to be wrong.

We should note that the Lazy Argument sometimes delivers a correct result. Suppose that the result of my exam is not determined by my performance, but rather, the professor grades the assignments in a random and arbitrary way. In this case, the argument given above is correct. If I knew that the professor is going to grade the assignments in such way, the most rational thing to do would be to enjoy myself, because my present behaviour would not affect the future outcome of my exam.

To sum up, the Lazy Argument can be rejected by a modal fatalist who believes in causal connections among events. What we do in the present

can sometimes affect the way the future will be. Hence, in some cases, it is rational to pursue those courses of actions which may be painful in the present but which are going to produce a greater good in the future. The fact that a future event is fated should not affect our decisions, because the means which are going to bring about an end would also be fated and necessary for its occurrence.

1.4.2 Deliberation

Some authors raised worries concerning fatalism and the process of human deliberation. Aristotle made the following point, when he was discussing the fatalistic idea that everything that happens, happens out of necessity.

there would be no need to deliberate or to take trouble (thinking that if we do this, this will happen, but if we do not, it will not).

A reasonable fatalist does think that all events are necessary, and that whatever actually happens is such that it can't be otherwise. But, it hardly follows that it is not the case that if I take a course of action, such and such will then happen, and if I take a different course of action, something else will happen. The fact that different courses of action bring about different effects is simply due to causal connections among events, namely something a reasonable fatalist can embrace without any problem. Hence, this Aristotelian worry about deliberation applies only to lazy fatalism and not to a modal fatalist who embraces the Causal Connectedness Thesis.

Van Inwagen (1983) raised another kind of worry. He starts by saying that

the fatalist believes that if an agent is in fact going to do some particular thing, then that thing is the only thing he can do, the only thing it is open to him to do. (p.61)

This is certainly true for a modal fatalist. He then goes on by saying

And it seems to be a feature of our concept of deliberation that we can deliberate about which of various mutually exclusive courses of action to pursue only if we believe that each of these courses of action is open to us. Therefore, anyone who accepts fatalism... must, on pain of self-contradiction, refrain from deliberating about future courses of action. (p.61)

We often spend time deliberating. And it is true that when we do so, we believe that more than one course of action is open to us. Van Inwagen

thinks that this, together with the belief in the truth of fatalism, gives rise to a contradiction within the set of an agent's beliefs. Say that I am choosing between A and B, two mutually exclusive courses of action. Since I am deliberating whether to do A or B, I believe that both A and B are open to me, Van Inwagen would say. But, if I am a modal fatalist, I also think that only one between A and B is open to me. And this undoubtedly would amount to a contradiction in the set of my beliefs.

At least two responses are possible to this worry. The first one has to do with practical considerations, whereas the second one has to do with our limited knowledge of the future.

The first one is illustrated by Pereboom (2001).

...at the moment of choice, we must indeed make what might well be the false and unjustified assumption that more than one course of action is available to us. It then claims that it is legitimate to assume this cognitive posture because the practical gains of engaging in deliberation are significant enough to outweigh the losses of having false and unjustified beliefs. We are left with the following choice: either deliberate and have a belief that you know might well be false whenever you do, or cease to deliberate. In this case, practical rationality would appear to have the upper hand. (p.136)

I think this way of arguing properly takes care of the worry raised by Van Inwagen. To this, I want to add that the false and unjustified belief (if modal fatalism were true) that more than one course of action is open to me does *not* have any practical loss. False beliefs generally have a negative impact. Suppose I am deliberating about what stocks I should buy from the market. False beliefs about the future market behavior surely affect negatively the choice I am about to make, whereas true beliefs are going to affect positively my choice. Or, suppose that I have the false beliefs that if I do A, then such and such will happen and if I do B, something else will happen. If those beliefs about the future consequences of my actions are false, they are going to have a negative impact on my process of deliberation, in the sense that I am likely to not achieve what I want to achieve with my actions. But, the false and unjustified belief (if modal fatalism were true) that more than one course of action is open, is completely harmless with respect to a practical point of view. I simply can't see how this kind of belief would have any impact at all on my process of deliberation.

The second response has to do with epistemic openness. If I am a coherent modal fatalist, I do believe that only one between A and B is

open to me. But I don't know which one of the two options I am going to choose, i.e. I don't know what will happen, even though I believe that what will happen will necessarily happen. Hence, I don't believe that I will do A and I don't believe that I will do B. I am uncertain about what I will eventually decide to do. And there is absolutely no contradiction in this set of beliefs. In other words, the openness I believe in when I am deliberating might be of an epistemic kind, in the sense that it is grounded in my limited knowledge of the future, not in how things really are. The future might very well be metaphysically closed and epistemically open. How the world is now is compatible with just one possible continuation, only one choice is really open, but the set of my beliefs might be compatible with more than one option about my future choices. By saying that the openness we believe in when we engage in deliberation is just epistemical, the contradiction pointed out by Van Inwagen fades away¹⁹.

1.4.3 Moral Responsibility

Fatalism is often seen as a doctrine which does violence to our intuitions and threatens our ordinary ways of thinking. One of the major worries comes from considerations about moral responsibility. One might argue that fatalism rules out moral responsibility and that this amounts to an unpleasant result. My aim in this section is twofold. First, I wanna show that a reasonable fatalist can endorse a moral principle which has nothing to do with the concept of moral responsibility and can do a lot of explanatory job with regard to our moral attitudes. Second, I wanna show that perhaps there is nothing wrong with abandoning the concept of moral responsibility.

Here is one reason why modal fatalism seems to fail to make sense of

¹⁹I thank Andrea Iacona for pointing out to me a worry with respect to epistemic openness. There seem to be cases where epistemic openness is irrelevant. Suppose I want to have a drink and I don't know whether there is wine or beer in my refrigerator, and just one of the two is there. In such case, it is epistemically open whether I'll have beer or wine. However, it wouldn't be rational to deliberate about what I will drink, because the most rational thing to do would be to check the content of my refrigerator. I think that a reasonable fatalist might respond that this case shows that it is rational to collect information before deliberating, when it is possible to do so. However, I don't think this principle causes harm to the strategy of appealing to epistemic openness to explain deliberation. Most of the cases regarding deliberations about future courses of actions and their possible consequence are such that it is impossible to collect all the relevant information, because what lies in the future is in general unknown to us. Hence, epistemic openness might remain, in spite of the metaphysical closeness due to modal fatalism.

our moral lives. It comes from a very simple argument which combines the so-called Principle of Alternate Possibilities (PAP) and fatalism. Here is the argument.

(PAP) A person is morally responsible for what he or she has done only if he or she could have done otherwise

(48) No one has ever the power to do otherwise

(49) Hence, no one is ever morally responsible.

(48) expresses the fatalistic idea that for any action any person performed, it is never the case that the person could have done otherwise than what he or she has actually done. Combined with (PAP), we have the result that no one is ever morally responsible. If we buy into another principle, we can reach another conclusion. The principle I have in mind is one that relates moral responsibility and the justice of punishments and rewards. The principle, which I shall call the principle of Moral Responsibility (MR), could be stated as follows.

(MR) A punishment (or reward) of a wicked (righteous) act is just only if the person who acted can be held morally responsible for the act.

The acceptance of this principle, together with (49) gives us the corollary that all punishments are unjust. Now, what could a reasonable fatalist say? Of course, a reasonable fatalist can't abandon (48), because it is the main thesis of modal fatalism. And, no one can reasonably accept the conclusion that all punishments we usually assign are morally wrong. For this would amount to say that it is morally wrong to assign a fine to someone who was driving his car so fast that he jeopardized other people's lives or that it is morally wrong to send to jail someone who deliberately killed someone else. Such results are clearly to be avoided and I want to argue that there are ways out for a reasonable fatalist to avoid this conclusion. A fatalist could accept (49), the idea that no one is ever morally responsible, and he or she could have a different account that explains why punishments and rewards are just, even in the absence of moral responsibility. If so, we could save the legitimacy and rationality of punishments and rewards we usually assign.

Before showing how this strategy can be pursued, I want to say something about the quick argument from (PAP) and modal fatalism to (49). The principle (PAP) is controversial. Frankfurt (1969) challenged it and declared it wrong. He came up with counterexamples to that principle. His strategy amounts to show that the power do to otherwise is not a nec-

essary condition for moral responsibility. Fischer (1986) gives an example of a Frankfurt case. The case involves Green, a man who saved a child and has a microchip installed in his brain, which is controlled by a team of scientists in California.

Though Green (the man who rescued the child) has had a device implanted in his brain, the device does not play any role in Green's decision to save the child (and his subsequent action). That is, the device monitors Green's brain activity but does not actually intervene in it. Let us suppose that this is because the scientists can see that Green is about to decide to save the child and to act accordingly. But let's also suppose that the scientists would have intervened to bring about a decision to save the child if Green had shown an inclination to decide to refrain from saving the child. That is, were Green inclined to decide on his own not to save the child, the scientists would ensure electronically that he decide to save the child and also that he act to carry out his decision. (p.41)

This kind of cases can be seen as counterexample to (PAP). This is so because it seems that Green couldn't have done otherwise. It was impossible for him to refrain from saving the child, because had he tried to do so, the device implanted in his brain would have ensured the child save. Yet, our moral intuitions inclines us to say that Green deserves credit for what he did. He was in control of his decisions and behavior when he saved the child, whereas the scientists in the lab who have the power to control his brain didn't play any causal role in this case. Hence, contrary to (PAP), we have a case where it is true that the person is held morally responsible for an act, even though he couldn't have done otherwise. That being said, I don't think that Frankfurt cases would be of any help to the modal fatalist. Sure, by denying (PAP) we would block the argument to (49). But, since Green saved the child without any intervention by the California scientists via the brain device, a reasonable fatalist needs to say that Green could have not refrained from saving the child, and this is so independently of considerations about the microchip in his head. So, I think a reasonable fatalist should resort to another strategy. As I said before, I think the most appropriate way would be to abandon moral responsibility and retain the justness of punishments and rewards we usually assign. The principle (MR) is the one a reasonable fatalist should question.

Can we call into question (MR)? Yes, if we have alternatives. And, there are alternatives. A fatalist may think that punishments and rewards are

just as long as they are effective. That is, we are allowed to assign a punishment to a person only if the punishment is such that we have reasons to think it will be a negative reinforcement which will positively condition the future behavior of the person who receives the punishment. We state the principle of effectiveness of punishments and rewards (EPR) as follows.

(EPR) A punishment (or reward) of a wicked (righteous) act is just only if it has a positive effect with respect to future conduct.

It should be noted that (EPR) needs the idea that there *are* causal connections among events in order to be non-vacuous, in the sense that if there were not any causal connections among events, no punishment or reward would be just, simply because nothing would have an effect. And, the kind of modal fatalism we are interested in — reasonable fatalism — does not deny causal relations among events, because it accepts the Causal Connectedness thesis. Hence, a reasonable fatalist can embrace (EPR). He or she could then argue that it might be the case that a punishment assigned to a person after a wicked deed he or she has committed will bring about that the same person will behave in a better way under similar future circumstances. The same could be said for rewards, they are just as long as they are positive reinforcements with respect to future conduct. It is worth noting that (EPR) does not need at all to posit the power to do otherwise, which is thought to be a necessary condition for moral responsibility by the principle (PAP). This way it could be used by a fatalist to claim that punishments and rewards may be justified, even in the total absence of the freedom to do otherwise. Is the principle (EPR) plausible? For sure we have reasons to think it is. Suppose we find out that a man has killed several people. The police arrests him. Soon, we find out that because of a severe disease the man has, the man is bound to die within few hours. In such a case, my moral intuitions incline me to say that it would be a gratuitous disgraceful act to inflict a punishment to this man. It would be acceptable to make sure that he does not commit other crimes in his last few hours, for instance by confining or controlling him, but other kind of punishments would be unjust and a pure case of vengeance. (EPR) explains why things are so. Given that the man is going to die soon, he will not find himself in circumstances similar to the ones when he killed other people. So, there is no future conduct to influence via the infliction of a punishment. Hence, a punishment in this case would be unjust.

We shall see now how (EPR) deals with some cases traditionally discussed in the literature. Suppose that I find out that a good friend of mine is being electronically manipulated by the same team of scientists

who are controlling Green's brain in the example discussed above. In this case though, the manipulation is constant and we assume it takes place throughout all my friend's life. Any act performed by my friend is decided and brought about by the team of scientists in California. Clearly, my friend should not be held responsible for what he does or fail to do. The (PAP) principle does not consider him a proper subject of moral attitudes because my friend lacks at all the times the power to do otherwise. Things are so for everyone according to a fatalist. Hence, as we have seen, a modal fatalist needs another principle to account for this case. (EPR) gives us the desired results. According to (EPR), it would be unjust to punish or reward my friend for his behavior. In fact, all punishments and rewards will never make a difference in the future conduct of my friend, simply because his future conduct will always be determined by the team of scientists. My constantly manipulated friend is just a mean for the ends the scientists want to pursue. Hence, (EPR) can correctly account for this case without appealing to the concept of moral responsibility.

Let us take the case of Ulysses binding himself to the mast to resist the song of the Sirens. At the moment in time when he resisted the song, he could not have done otherwise because he was tied up. (EPR) again explains why Ulysses would deserve a reward for this. He bound himself when he was in control of himself. He was coerced at the moment in time when he resisted the song, but he himself initiated the coercion. Other cases of coercion are different. Suppose someone put a pill in my drink without my consent. The pill is so strong that it takes full control of my behavior and coerces me to perpetrate evil acts. According to (PAP) I would not be morally responsible for what I do, because the pill blocks my power to do otherwise. (EPR) gives an analogous correct result. It would be unfair to punish me for what I did when I was under control of the pill, because the pill had an invincible force which overwhelmed the control of my behavior. Nothing else could have affected my behavior in the time when I was under the effect of the pill²⁰. Hence, punishments

²⁰It is interesting to note that if the pill has not an invincible force over me, but its force is a matter of grade, then sometimes punishments or rewards might be acceptable. After all, in future similar circumstances I might behave better in virtue of the punishments or rewards I receive now. Say that someone makes me take a pill without my consensus and it *partly* influences my behavior. As a result of the pill and *other* factors coming from my, say, personality, I commit a bad deed. It seems that in this case (EPR) would allow for a partial limited punishment. And this seems correct. If I will find myself in similar circumstances, i.e. under the effect of a not-overwhelming pill, I might behave better as a result of the punishment I received. Hence, it seems that (EPR) might account for mixed cases, whereas (PAP) cannot, because the power to do otherwise seems to be a matter of all or nothing.

and rewards would be unjust in this case, because in such circumstances only the pill exercises a causal role. The cases of the pill, Ulysses and my friend constantly under total control of the scientists highlight a moral phenomenon. External coercion excuses, internal coercion does not. If I am externally coerced against my will, like in the case of the pill and my friend, I am not a proper candidate for punishments and rewards. If I am internally coerced, that is, I initiate the coercion, I am a proper candidate for punishments and rewards. Again, this has nothing to do with the power to do otherwise, and (EPR) provides results in accordance with our moral intuitions.

We have seen that (EPR) can account for punishments and rewards we usually assign without bringing into the picture the notion of moral responsibility. What does it mean for a person to be morally responsible? If being morally responsible just means being the proper candidate for rewards and punishment, then (EPR) suffices. Under this sense, persons are morally responsible, whereas things like particles, tables or trees are not. If so, we don't need (PAP). If being morally responsible also means being the proper candidate of moral blame and praise, I would say that blame and praise are nothing else than a form of punishment and reward. So, again, (EPR) would suffice. But, one could add more to the concept of moral responsibility. For instance, Fischer (1985) illustrates this broad account of moral responsibility.

On this sort of account, a person is a morally responsible agent when he is an appropriate candidate for the reactive attitudes and for such activities as praise and blame and punishment and reward (p.12)

By reactive attitudes Fischer has in mind attitudes like respect, gratitude, love, indignation, and resentment. And, Fischer invites us to consider what it would be for such attitudes if we found out that, say, a good friend of mine that I love and care about is *constantly* manipulated by the team of scientist we have been talking about.

At first it would be hard to know how one would react to such an unusual situation. But, I think, once you had been convinced that direct manipulation exists, a striking thing would occur: many of your most basic attitudes toward your friend would change.

This is certainly true. I would change my attitudes toward my friend in such a case. But what would be the reason of this change? I don't think

my attitudes would change because my friend can never do other than what he actually does. Suppose I am totally convinced that fatalism is true and that the power to do otherwise never exists. This would not strike me as a good reason to change my attitudes toward all the people I know. It would be just a metaphysical belief of mine with a little impact to my feelings of love, friendship and the like. For the same reason, if I became convinced that 4-dimensionalism is true, I would not be offended by the fact that my friend is not wholly present with me while we are hanging out at the bar. The reason why I would change my attitude toward my friend after finding out that a constant manipulation is occurring has to do just with the fact that my friend is externally coerced by the team of scientists. If I admire and respect my friend, after the discovery, I would rather be friend with the scientist than with him. And we have seen that (EPR) accounts correctly for cases of external coercion. Again, we don't need to bring into the picture the power to do otherwise to make sense of our reactive attitudes. (EPR) is enough for a reasonable fatalist.

To sum up, a reasonable fatalist can get rid of the concept of moral responsibility. Moral responsibility comes with the notions of blame and praise, understood as something more than just forms of punishments and rewards. That is, moral responsibility is a judgment about people's conducts. On the contrary, a reasonable fatalist never judges. He thinks things go the only way they can go, and therefore no one is morally responsible in that sense. However, this does not imply that no one has to be punished or rewarded. (EPR) can explain why punishments and rewards may be just. They are just as long as they bring about moral improvements in one person's life. Moreover, the kind of acceptance towards events a reasonable fatalist has, does not imply that we can't hope for a better future. Again, punishments and rewards might be capable of bringing about a better future in our lives.

·
·

Chapter 2

A New Argument for Modal Fatalism

... so it goes ...

Slaughterhouse Five - Kurt
Vonnegut

In this chapter I am going to provide my own argument in favor of modal fatalism.

2.1 Two Principles which imply Modal Fatalism

We have seen in section 1.3.2 that modal fatalism can be stated through the principle (MF).

(MF) For any act ϕ , for any agent A and for any time T , it is within A 's power to ϕ at T if and only if A ϕ s at t .

As we have already seen, the right-left side of the biconditional is uncontroversial. Anyone should agree that if an agent does something at a time, it is within her power to do so. The controversy between a modal fatalist and an anti-modal fatalist arises on the left-right side of (MF).

(lrMF) For any act ϕ , for any agent A and for any time T , if it is within A 's power to ϕ at T , **then** A ϕ s at T .

Here are some instances of the principle (lrMF). If it is in Alice's power to swim at T_1 , then Alice swims at T_1 . If it is in Bob's power to ride a

bicycle at T_2 , then Bob rides a bicycle at T_2 . If it is in Carol's power to give a philosophy lecture at T_3 , then Carol gives a philosophy lecture at T_3 . Here is the upshot of the principle — according to (IrMF), whenever it is within an agent's power to ϕ , then she actually ϕ s and ϕ ing is a necessary condition for having the power to ϕ . That means that if an agent does not ϕ at T , it is not in her power to ϕ because that is simply the contrapositive of (IrMF), thus, it is logically equivalent to it. In other words, (IrMF) captures the idea that there is no room for non-instantiated powers. The only powers we possess are the instantiated ones.

We now turn to two principles which entail (IrMF). The first one is the principle I shall call (NC).

(NC) For any act ϕ , for any agent A and for any time T , if it is within A 's power to ϕ at T , **then** all necessary conditions for A to ϕ at T are met.

(NC) appears to be quite uncontroversial and hardly deniable. That can be seen by what its contrapositive says. That is, if at least one of the necessary conditions for ϕ ing is not met at T , then it is not within the agent's power to ϕ at T . For instance, if a necessary condition for swimming is not being affected by paralysis, then an agent who is paralyzed has not the power to swim; equivalently, an agent who has the power to swim is not paralyzed. Or, if a necessary condition for a violinist to play violin is the presence of a violin around, then the absence of it implies that playing the violin is not within the violinist's power. We should stress that the notion at play is not the notion of know-how, expertise or ability. The professional violinist would keep his ability to play violin even if all the violins on earth magically disappeared forever. But in such case, the violinist would never have again the power to play a violin. And the reason would be the lacking of a necessary condition for the act of playing violin. All of this is captured by (NC).

The other principle I want to appeal to is the following one.

(SUF) For any act ϕ , for any agent A and for any time T , if all necessary conditions for A to ϕ at T are met, **then** A ϕ s at t .

(SUF) is not as uncontroversial as (NC) is. We shall see in the next section how to deal with it within the modal fatalism debate. For the time being, we should note that (NC) together with (SUF) implies (IrMF). This is so just because of the transitivity of conditionals. Thus we have the result that (NC) together with (SUF) and the right-left side of (MF) imply (MF), i.e. modal fatalism. I have argued that (NC) and the right-left side

of (MF) are uncontroversial and should be accepted by any side of the modal fatalism/anti-fatalism debate. Hence, the debate can be cashed out in terms of the principle (SUF). That is, given the chain of implications just shown, an anti-fatalist *has* to deny (SUF) and justify its denial. On the other hand, a modal fatalist might or might not accept (SUF). After all, even if (SUF) together with plausible principles imply modal fatalism, a fatalist may have reasons independent from (SUF) for believing in the truth of modal fatalism. However, for the purposes of this chapter, we take the modal fatalist side to accept (SUF). What I want to highlight is that the anti-fatalist side needs to come up with a justified denial of (SUF). In the next section I shall argue that it is not so easy to do so.

2.2 The Disagreement on (SUF)

Before addressing the disagreement on (SUF), let us see another thing besides (NC) and the right-left side of (MF) on which both the modal fatalist and anti-fatalist sides agree on. A modal fatalist thinks that, given an arbitrary act ϕ , it is not necessarily the case that all necessary conditions for ϕ ing are fulfilled at all times during an agent's life. That is so because, according to the modal fatalist who believes in (SUF), all necessary conditions for ϕ ing are met only when the agent ϕ s, and quite obviously an agent does not always perform the same deeds throughout her life. If an agent, for instance, swims at T_1 , T_3 and not at T_2 , according to the fatalist who believes in (SUF), all necessary condition for swimming are met at T_1 and T_3 , whereas at least one necessary condition for swimming is not met at T_2 . Times where all necessary conditions for ϕ ing are met and the agent does not ϕ are ruled out by (SUF). Hence, in a fatalist account, given an act and an agent, all the necessary conditions for the agent performing the act are fulfilled at some times, whereas they are not fulfilled at other times. But that things are so must be true also in a non-fatalist account. It does not seem plausible to say that all necessary conditions for, say, swimming are always met by a certain agent. Again, if the agent is paralyzed at certain times throughout her life, whereas it is in her power to swim at other times, then there are times where at least one necessary condition for her swimming is not met and other times where all necessary conditions are met. Hence, the anti-fatalist agrees with the fatalist that, given an act ϕ and an agent A, there must be times where all necessary conditions for ϕ ing are not met.

That being said, we are now in a position to state the disagreement. The disagreement between the modal fatalist and the anti-fatalist is, given

an arbitrary act and an arbitrary agent, the partition of times with respect to the fulfillment of necessary conditions for that act and whether or not the agent *does* that act. The modal fatalist, given an arbitrary act ϕ and an arbitrary agent A , very straightforwardly and in accordance with (SUF) partitions times with respect to the fulfillment of necessary conditions for ϕ ing and the agent actually ϕ ing in only two regions. In one region, we have times in which the necessary conditions for ϕ ing are met and the agent ϕ s. In the other region, we have times where at least one necessary condition for ϕ ing is not met and the agent does not ϕ . The anti-fatalist partition is different since she has to claim that there is a further region, namely times in which all necessary conditions for ϕ ing are met and nonetheless the agent does not ϕ . Cases which belong to this third region are ruled out by (SUF), because (SUF) simply claims that if at a time all necessary conditions for ϕ ing are met, that is a time where the agent ϕ s. On the contrary, cases in the third regions are counterexamples to the principle (SUF), and so those are precisely the cases the anti-fatalist needs in order to falsify (SUF). In other words, the fatalist claims that this third region must be empty, while the anti-fatalist has to claim and justify the existence of cases in that third region. A possible fourth region, one in which at least one necessary condition for ϕ ing is not met and the agent ϕ s would not be ruled out by (SUF). Yet it is ruled out by the uncontroversial principles (NC) and the right-left side of (MF). In fact, if in such a region the agent ϕ s, in virtue of the right-left side of (MF), it is the case that ϕ ing is within the agent's power. If so, in virtue of (NC), all necessary conditions for ϕ ing are met.

We now turn to an example to illustrate the disagreement between a modal fatalist and an anti-fatalist. We go back to our naval admiral A , who is in charge of a naval fleet. Throughout his life he issues the order to attack several times. We then take all the times of our admiral's life. The modal fatalist partitions times this way, in accordance with (SUF).

Table 2.1: Modal Fatalist partition

First region	Second region
- All necessary conditions for issuing the order are met	- At least one necessary condition for issuing the order is not met
- A issues the order	- A does not issue the order

In the first region we have times where all necessary conditions for issuing the order are met and A issues the order. In the second region,

we have times where at least one necessary condition for issuing the order is not met and A does not issue the order. As we have already seen, an anti-fatalist has to add a third region, in order to have counterexamples to (SUF). She needs to claim that there are times where all the necessary conditions for issuing the order are met and A does not issue it. So, here is the anti-fatalist partition.

Table 2.2: Anti-fatalist partition

First region	Third region	Second region
- All necessary conditions for issuing the order are met - A issues the order	- All necessary conditions for issuing the order are met - A does not issue the order	- At least one necessary condition for issuing the order is not met - A does not issue the order

Not only an anti-fatalist has to claim the existence of times in the third region. She has also to justify those cases. Both in the first and third region we have times where all necessary conditions for issuing the order are met. Yet in the first region A issues the order, whereas he does not in the third. What does distinguish elements in this third region from the ones in the first region? It will turn out that it is hard to tell a satisfactory story about what is going on in the third region.

The anti-fatalist, in order to deny (SUF), has to claim the following statement, which we shall call 'third region' (TR).

(TR) there is at least a time T such that a) all necessary conditions for issuing the order are met at T and b) A does not issue it at T.

(TR) has to be justified. How can this be possibly done? Let us start with what I call the preventor story. The first-third region distinction may come from the presence of a preventor in the third region. The idea is that at times in the third region, even if all necessary conditions for issuing the order are met, the agent does not issue it because of the occurrence of a preventor, whereas there are no preventors at times in the first region. We might think the preventor $P(\phi)$ of an act ϕ simply as the occurring of a state of affairs which blocks the agent from doing ϕ , even if all the necessary conditions for ϕ are met. To give an example with the act of swimming, a typical preventor could be the absence of water around, but in the sense of preventor here adopted also the agent's unwillingness to swim count as a preventor. In our admiral case, a preventor could be a momentary lack

of voice for our admiral, or a counter-order from our admiral' superior. In such story, the anti-fatalist claim (TR) becomes (TRP)

(TRP) there is at least a time T such that a) all necessary conditions for issuing the order are met at T, b) A does not issue it at T, and c) P(issuing the order).

We might want to ask now whether the *non*-occurrence of the preventor is a necessary condition for issuing the order. The answer must be either affirmative or negative. Suppose the answer is affirmative. We know from (TRP) that all necessary conditions for issuing the order are met. But, then, answering affirmatively to the question above means that the non-occurrence of a preventor is the case, because we are saying that its non-occurrence is a necessary condition for issuing the order and all necessary conditions for issuing it are met. But (TRP) says also that the preventor is (or has been) the case. Hence, we have a contradiction and we cannot take the non-occurrence of the preventor as a necessary condition for issuing the order, together with the claim (TRP). Let us suppose now that the answer instead is negative. That means that according to the answer we are giving, the preventor non-occurrence *is not* a necessary condition for issuing the order. But it follows from the adopted definition of a preventor that if the preventor is the case, then A does not issue the order. But the previous statement is tantamount to the claim that the preventor non-occurrence *is* a necessary condition for issuing the order. Hence, we have another contradiction. That means that the preventor story cannot be told to make sense of cases in the third region without running into contradictions. Hence, it cannot be told at all by an anti-fatalist. On the other hand, the modal fatalist can safely take the non-occurrence of any preventor as a necessary condition for issuing the order. Cases of prevention would safely go in the second region.

Another story that could be told by the anti-fatalist in order to defend the existence of cases in the third region is the one I shall call the reason-story. The story goes like this; the first region contains times where all necessary conditions for issuing the order are met and A does so, simply because he decided to issue the order there. In the third region our admiral does not decide to issue it, and that is why he does not do so there. If the anti-fatalist does not like the concept of decision in order to justify the distinction she needs to make, let us be more general and say that the difference between the first and the third region is the fact that there is a reason for issuing it in the first region whereas that reason is absent in the third one. Here a reason for issuing the order, which we shall call R(issuing the order), might be various things and an anti-fatalist can specify better

the nature of it however she likes: our admiral decided to issue the order, he freely decided to issue it, he intended to issue it, the power to issue the order manifested itself, and so on and so forth. For our purposes here we don't need to enter into these niceties. It should suffice to say that $R(\phi)$ is such that all necessary conditions for ϕ ing are met and the agent ϕ s only if $R(\phi)$ obtains. That being said, (TR) becomes

(TRR) there is at least a time T such that a) all necessary conditions for issuing the order are met at T, b) A does not issue it at T, and c) it is not (or has not been) the case that $R(\text{issuing the order})$.

And now we can ask the following question: is $R(\text{issuing the order})$ a necessary condition for issuing the order? As we did before, we have now to investigate two possible answers. Suppose the answer is affirmative. Then, $R(\text{issuing the order})$ is a necessary condition for issuing the order. (TRR) is a case where all necessary conditions for issuing the order are met. Hence, $R(\text{issuing the order})$ *is* or has been the case. But, (TRR) says that $R(\text{issuing the order})$ *is not* (or has not been) the case. Hence, we have a contradiction and we cannot take the non-occurrence of $R(\text{issuing the order})$ as a necessary condition for issuing the order, together with the claim (TRR). Suppose now the answer were negative. This would mean that $R(\text{issuing the order})$ *is not* a necessary condition for issuing the order. Again, (TRP) is supposed to be a case where all necessary condition for issuing the order are met. We have seen above that all necessary condition for issuing the order are met and our admiral issues the order only if $R(\text{issuing the order})$ is the case. That means that if all necessary conditions are met, then, if he issues the order, then $R(\text{issuing the order})$ is the case. (TRP) is a case where all necessary conditions are met. Thus, at those times, if he issues the order, then $R(\text{issuing the order})$ is the case. But the previous claim is tantamount to the claim that $R(\text{issuing the order})$ *is* a necessary condition for issuing the order. Hence, we have again a contradiction. $R(\text{issuing the order})$ is and is not a necessary condition for issuing the order.

This means that the reason story cannot be told to make sense of cases in the third region without running into contradictions. Hence, it cannot be told at all by an anti-fatalist. On the other hand, the modal fatalist can safely take the occurrence of the reason, whatever it is, as a necessary condition for issuing the order. Cases where this reason is absent would safely go in the second region.

2.3 What are we left with

To sum up, we deemed the principles (NC) and the (MF) right-left side uncontroversial. Those two principles lead to modal fatalism, provided you add the principle (SUF). I then argued that it is hard to justify counterexamples to (SUF). Neither the preventor story nor the reason-story succeeded in justifying the existence of cases in the third region which would falsify (SUF). It is hard to imagine what else could work, and so we are left with the idea that either modal fatalism is a true doctrine or that the distinction of cases in the first and third region is somehow brute.

One way to see that distinction as brute would be to claim that powers are primitive entities. In doing so, one could easily skip (SUF) and directly falsify (MF) by saying that when an agent ϕ s at a certain time, other alternative actions are within the agent's power. That would provide us a genuine counterexample to (MF). However, taking powers as primitives also means that powers cannot be explained at all. In fact, an anti-fatalist willing to make such a move cannot give an answer to a very simple question as why it is within a certain agent's power to do something. There would not be an answer, simply because in such an account powers were primitives. On the contrary, a modal fatalist who embraces (SUF) can provide a perfectly coherent answer to that question. She can simply claim that it is within the agent's power to ϕ because all necessary conditions for ϕ ing are met.

There is actually a way to refute (SUF) through a counterexample, for which no counter-argument along the lines of the previous ones can be given. Suppose two mutually exclusive acts, call them ϕ and ψ have the same necessary conditions. Suppose then the agent as a matter of fact ϕ s at T. Then — because of (MF) right-left side — ϕ is within the agent's power. This in turn means — via (NC) — that all necessary conditions for ϕ ing are met at T. But, by assuming that all necessary conditions for ψ ing are the same of ϕ ing, it follows that also all necessary conditions for ψ ing are met at T. But it cannot be the case that the agent ψ s at T, simply because we are assuming that the agent ϕ s at T, and ϕ and ψ s are incompatible actions. Hence, such a case would be a case in which all necessary conditions for ψ ing are met and it is not the case that the agent ψ s, i.e. it would provide a genuine counterexample to (SUF) and therefore to (MF). However, such a counterexample would be a very small victory for an anti-fatalist.

Say that ϕ and ψ are respectively the act of turning left and the act of turning right. Among the necessary conditions for ϕ ing we would find states of affairs which play a causal explanatory role for the fact that the agent ϕ s, i.e. things that had not been the case, the agent's ϕ ing would

not have been the case as well. Now, if we claim that ψ have the same necessary conditions to be met, it means that the states of affairs which provide a causal explanation of why the agent ϕ s, would also be a good causal explanation of the agent's choice, had the agent ψ ed. It seems to me that this would be a case where whether the agent ϕ s or ψ s is just a matter of chance¹. But this would be a small victory for the anti-fatalist. Pure chanciness cannot ground Free Will, because the outcome of a chancy event is something over which no one can exercise any control. Hence, we are left with two options: either modal fatalism holds in virtue of our reasoning on (SUF), or (SUF) and hence modal fatalism do not hold in virtue of genuine chanciness. In the former case, the power to do otherwise is always ruled out by modal fatalism. In the latter case, the power to do otherwise would be grounded in pure chanciness, a result which cannot be satisfactory to a theorist who wants to argue in favor of Free Will understood as the power to do otherwise.

It should ultimately be noted that my argument for modal fatalism does not suffer the flaw which is shared by the Main Argument and Taylor's. As we have seen in section 1.3.3, both those arguments conflate deterministic and indeterministic scenarios. That is, they apply necessity to outcomes which are clearly contingent, like in cases provided by Quantum Mechanics. My argument does not. It leaves open the possibility for contingent events, which are treated by my argument as cases where two alternative outcomes have the same necessary conditions. This result is due to the fact that my argument employs just necessary conditions which lie in the past with respect to an agent performing a certain act at a time, or in the past with respect to a certain experimental outcome. Taylor's argument instead takes advantage of necessary conditions lying in the future, thereby conflating deterministic and indeterministic cases. Yet, my argument, if correct, scores a point in favor of modal fatalism. In fact, either it completely rules out the power to do otherwise, or it shows that the power to do otherwise is nothing but chanciness, i.e. something which cannot adequately ground our intuitive notion of Free Will.

¹For instance, let us get back to the electron story told in section 1.3.3. Say that ϕ and ψ are respectively the electron exiting the box from the black aperture and the electron exiting the box from the white aperture. These two mutually exclusive states of affairs seem to have exactly the same necessary conditions. For instance, the electron is measured to be white at T, only if a white or hard or soft electron is fed in the box at T_{-1} . And similarly, the electron is measured to be black at T, only if a white or hard or soft electron is fed in the box at T_{-1} .

Chapter 3

Time Travel and Fatalism

In my experience, there's no
such thing as luck

Star Wars - Episode IV

In this chapter we will be dealing with time travel and what I have labeled modal and reasonable fatalism. Time travel raises worries related to time and what people can or cannot do. Unsurprisingly, David Lewis talked about fatalism in his famous paper about time travel. I will argue that reasonable fatalism can correctly account for what is going on with episodes of time travel.

3.1 Time Travel in One-Dimensional Time

In *The Paradoxes of Time Travel* (1976), David Lewis discusses the possibility of time travel. His main point in that article is that time traveling in one-dimensional time is metaphysically possible. He shows this is possible the only way one can show that something is metaphysically possible, namely by telling a story involving that something and arguing that there aren't contradictions in the story. Lewis tells the well-known story of Tim, who goes back in time to a space-time place where he has already been and does what he has already done. Before Tim enters his time machine, there are events which are past according to *external* time and future according to Tim's *personal* time. The distinction between external and a person's personal time is crucial when it comes to time travel. Lewis takes external

time to be time itself¹, whereas what is personal time has to be clarified. First of all, personal time is *not* a further time-dimension. Second of all, personal time applies usually to people, even though it can apply to objects as well. In general, it applies to *things* in a world. Lewis (1976) provides a functional definition of personal time:

it is that which occupies a certain role in the pattern of events that comprise the time traveler's life. (p.146)

Here Lewis is talking about a time traveler, but the notion is applicable to everyone's life. The role Lewis has in mind is the role of that which orders. Suppose you are given in a disordered way all the stages of a person and you are asked to order them with respect to time. You would first start with the birth's stage and end with the death's stage. You would put infantile stages before senile ones. If you are presented with a stage where the person is eating a pizza and another one where the person is digesting it, you would put the former before the latter. If in one stage the person has much more memories than he has in another one, the latter would be put later than the former, and so on. By taking into account more and more things, you will eventually end up having a complete ordering of the stages. This is what personal time *does*. It orders the stages of a person in accordance with the *regularities* we are familiar with in our world. It is a regularity which holds in our world that we first eat pizza and then digest it, or that we are first young and old later on. Such regularities we are familiar with provide a criterion for ordering stages according to personal time. We can even assign coordinates to personal time, as Lewis observes:

But there is one way to assign coordinates to the time traveler's stages, and one way only (apart from the arbitrary choice of a zero point), so that the regularities that hold with respect to this assignment match those that commonly hold with respect to external time. (p.146)

That is, we might assign the number 0 to the birth's stage, the number 1 to the stages where the person is 1 year old, 2 to the stages where the person is 2 years old and so on. Of course, this can be made more precise as much as we like (2 years, 1 day and 3 minutes . . . 2 years 1 day, 3 minutes and 10 seconds, and so on).

Take Tim's events. Tim is born in 1950 (external time), and enters his time machine to travel back to the past at the age of 25 (personal time)

¹"Instead I reply by distinguishing time itself, *external time* as I shall also call it . . . (p.146)"

in 1975 (external time). He then arrives back in the past and steps out of his machine in 1920 (external time). In 1921, say, Tim is 1 year older than when he departed, i.e. he is 26 years old according to his personal time. Before Tim enters his time machine, the events in the external past are the events that *will* happen in Tim's personal future. This episode of time travel gives rise to a *discrepancy* between external time — time itself — and Tim's personal time. The discrepancy has not to do with the fact that, say, Tim is 25 years old in 1920. This is a kind of discrepancy any person displays — one might be 1 year old in 2017, for instance. Rather, the discrepancy has to do with the fact that with episodes of time travel, two same events in a time traveler's life are separated by two different amounts of time. We can illustrate this by means of the graph (3.1).

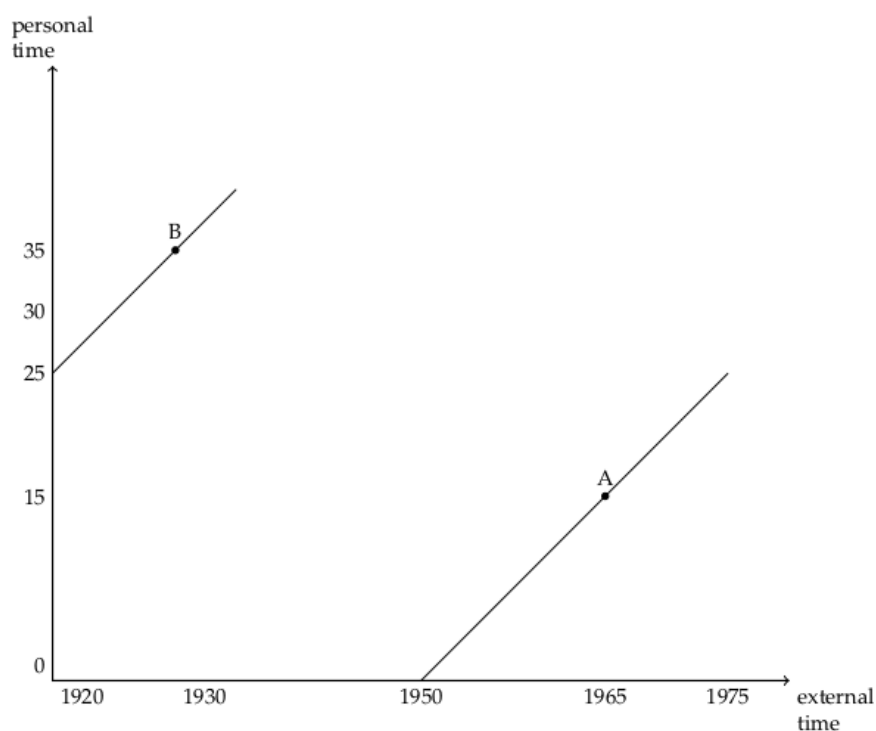


Figure 3.1: Tim's graph.

Take the events A and B which occurred in Tim's life. It is clear from the graph that A and B are separated by 35 years according to external time and by 20 years according to Tim's personal time. This kind of discrepancy, Lewis observes, is *essential* to episodes of time travel.

In Tim's case, there is another discrepancy, which is not essential to

episodes of time travel. The discrepancy has to do with the way events are ordered. For instance, B is earlier than A with respect to external time, whereas A is earlier than B with respect to Tim's personal time². To wrap up Tim's case, his journey to the past is instantaneous and has zero duration. Hence, his line on the graph is a broken line.

We could also imagine a journey to the past which takes time. Suppose that it takes one year of Tim₁'s personal time to go back to the past, from 1975 to 1920. During the travel, Tim₁ could do all the things he usually does in one year. If he usually reads 6 books per year, he can read 6 books during his trip. After one year journey, Tim₁ will find himself at the arrival one year older than when he departed. His line on the graph (3.2) would be a bent one³.

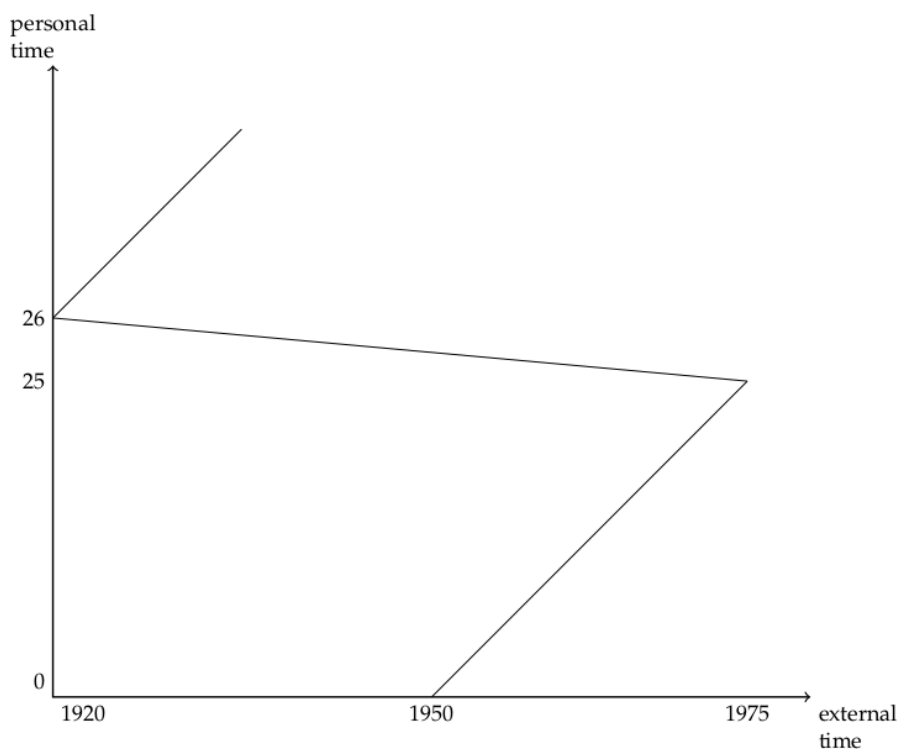


Figure 3.2: Tim₁'s graph.

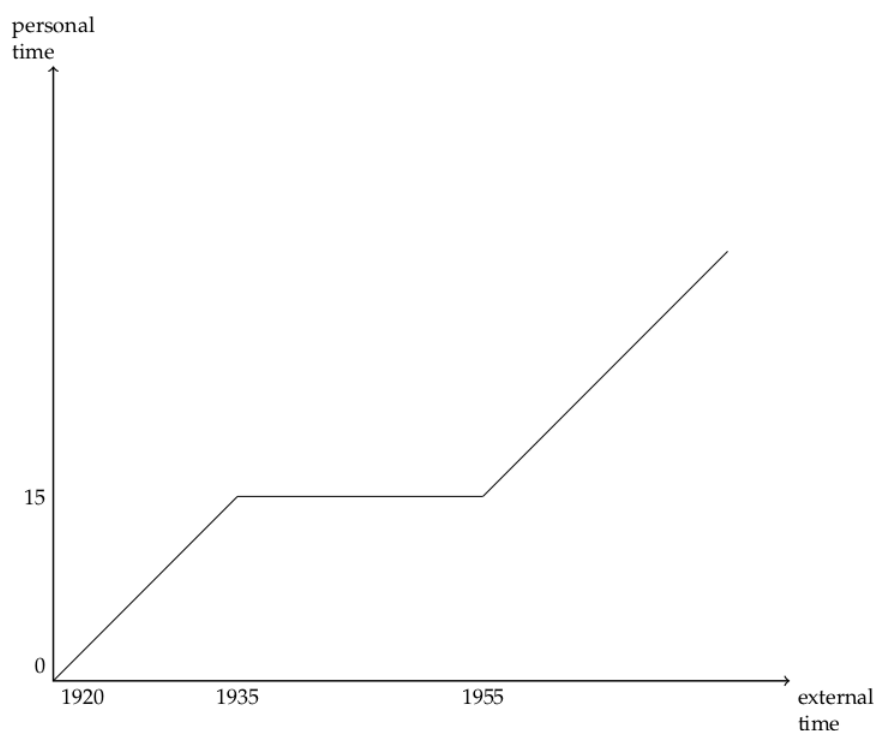
²This disagreement about the way events are ordered is not essential to episodes of time travel. Time travel toward the future provided by Special Theory of Relativity will be a case where all 'times' agree on the way events are ordered. We will treat it later on.

³Tim and Tim₁ can both depart at the same exact time and arrive at the same time even if Tim's journey is instantaneous and Tim₁ trip takes one year. It's just that Tim₁ will arrive at his destination older than Tim.

In both cases thus far considered, two events may be separated by different amounts of time, depending on which time we are considering. In Tim's case, his arrival is immediately after his departure according to his personal time, whereas according to external time, those two events are separated by 55 years. In Tim₁'s case, the departure and the arrival are separated by one year of his personal time and 55 years of external time.

We may imagine other cases of time travel, not treated in Lewis's paper. It is possible to time travel toward the future. Suppose cryogenic sleep will become one day possible, with respect to the development of our technology. Cryogenic sleep would put one person, say Tim₂, into a state of "total sleep", where all body activities and aging processes would be stopped. Time in the outside world would keep flowing, whereas Tim₂'s personal time would be stopped for a while, until he will be eventually safely waked up by a team of future scientists. This would legitimately count as a case of time travel, even if we don't have any time machine as we are used to from science fiction stories. In fact, again, we would have a discrepancy between external and personal time. Tim₂ falls asleep in his cryogenic sleep, and according to his personal time, he wakes up immediately after it, whereas according to external time, several years may pass. We can represent Tim₂'s case with a bent line on the graph (3.3), where the horizontal segment stands for the time where Tim is sent to cryogenic sleep. We assume Tim₂ is sent to cryogenic sleep at the age of 15 and woken up after 20 years of external time has passed.

Another case of time travel comes from Special Theory of Relativity. We can illustrate it by means of the famous Twin Paradox. It is not a real paradox, because it can be fully explained within the Special Theory of Relativity. Yet, it appears to be paradoxical because it has counter-intuitive results with respect to some principle of relativity, like the principle that all frames of references 'are created equal'. That is, if you have a symmetrical situation, and you reverse it, the explanation of what is going on should be the same. Here is one version of the Paradox. Suppose Alice and Bob are two twins; Bob is on planet Earth and Alice sets off with her spaceship to reach a star 3 light years away from planet Earth, and she then comes back to Earth. If she travels with a velocity of $0.6c$, Bob will observe that her round trip takes 10 years. Because of the time dilation principle — moving clocks appear to run slow compared to the clocks of a stationary observer — Bob would also observe that according to Alice's clocks, her trip takes less than 10 years. Bob will observe that Alice's clock ticks 8 years. Alice herself will claim that her trip took 8 years, according to her clocks. Here is where the paradox comes in. Bob on Earth can understand what is going on. He observes Alice moving, and he can tell why she is younger

Figure 3.3: Tim₂'s graph.

than him when she comes back on Earth. 10 years passed according to Bob's clocks, whereas just 8 years passed according to Alice's clock, and this is just the familiar principle of time dilation. What is paradoxical is Alice's perspective. As far as Alice is concerned, she is stationary and she observes Bob moving away from her and then coming back to her at a speed of $0.6c$. Thus, Alice sees Bob's clocks run slowly compared to hers. Given that from the departure to the reunion her clocks ticked 8 years, she will observe Bob's clocks run slow. By doing some simple math, we have the result that she will observe Bob's clock ticking 6.4 years, again because of time dilation. Hence, Bob will expect Alice to be younger than him at the moment of the reunion. But Alice seems to have parallel reasons to expect Bob to be younger than her. Of course this can't happen. Either they have the same age when they reunite, or, one twin is younger than the other one. This paradox can be explained within special theory of relativity⁴.

⁴Sometimes it is argued that the Twin Paradox case cannot be explained within the Special Theory of Relativity, because Special Theory of Relativity only deals with inertial frame of references, where objects move with a constant velocity, whereas necessarily

What happens is that when Alice turns around and change direction to get back on Earth, her lines of simultaneity in the space-time diagram change. When she reaches the star, what happens is that she reads on her clocks that 4 years have passed, and 3.2 years passed on Bob clocks. Bob instead will observe that when she reaches the star, 5 years have passed on his clocks and 4 on Alice's clock. This is due to the relativity of simultaneity and time dilation. On a space-time diagram, Alice and Bob will have different lines of simultaneity, given that they are moving with respect to each other. However, as soon as Alice changes direction, her lines of simultaneity are going to change. Immediately after she turns around, she would observe Bob's clock jump from 3.2 years to 6.8 years. Again, on her inbound trip, Alice would observe that it takes her 4 years to get back. Bob will agree on that, but he will also observe that the inbound trip takes 5 years according to his clocks. So, uncontroversially, Alice is the one who will be two years younger than Bob at the moment of the reunion. What interests us is that this count as a case of time travel toward the future. By traveling across the galaxy and then coming back to Earth, Alice will *age* 8 years, whereas on Earth 10 years will pass. This count as a legitimate case of time travel toward the future, even though Alice does not jump into a time machine, but 'just' in a super-fast spaceship.

Thus far I have been using graphs which used as an axis Lewis's notion of external time, which is time itself, as he says. The problem with Special Theory of Relativity is that there isn't 'one time'. Time is relative to a frame of reference, and Alice and Bob are in two different frames of reference. Moreover, as far as they are concerned, their respective personal time and 'their' external time tick at the same rate. Hence, in their graph (3.4) I take as axes Alice's external-personal time and Bob's external-personal time to highlight that in this case two events are separated by two unequal amounts of time, a feature which is essential to cases of time travel according to Lewis⁵. In the graph, Alice's round trip takes 8 years according to her clocks, whereas it takes 10 years according to Bob's ones. This time the slope of Alice's line will be greater than 1. More precisely, it will have a

Alice has to decelerate and then accelerate when she turns around at the star. But this is not the case. We can idealize the situation and think that the deceleration/acceleration elapsed time is so short that it does not affect the possibility of explaining the paradox within the Special Theory of Relativity.

⁵Here we should note that we don't have the kind of discrepancy which is essential in case of time travel toward the *past*, i.e. the disagreement on how events are ordered. In Tim's case, the event A on his graph is earlier than B according to his personal time and later than B according to external time. On the contrary, Alice and Bob agree on how the events of, say, Alice's trip are ordered. They disagree on how much they are separated in time.

slope of 1.25.

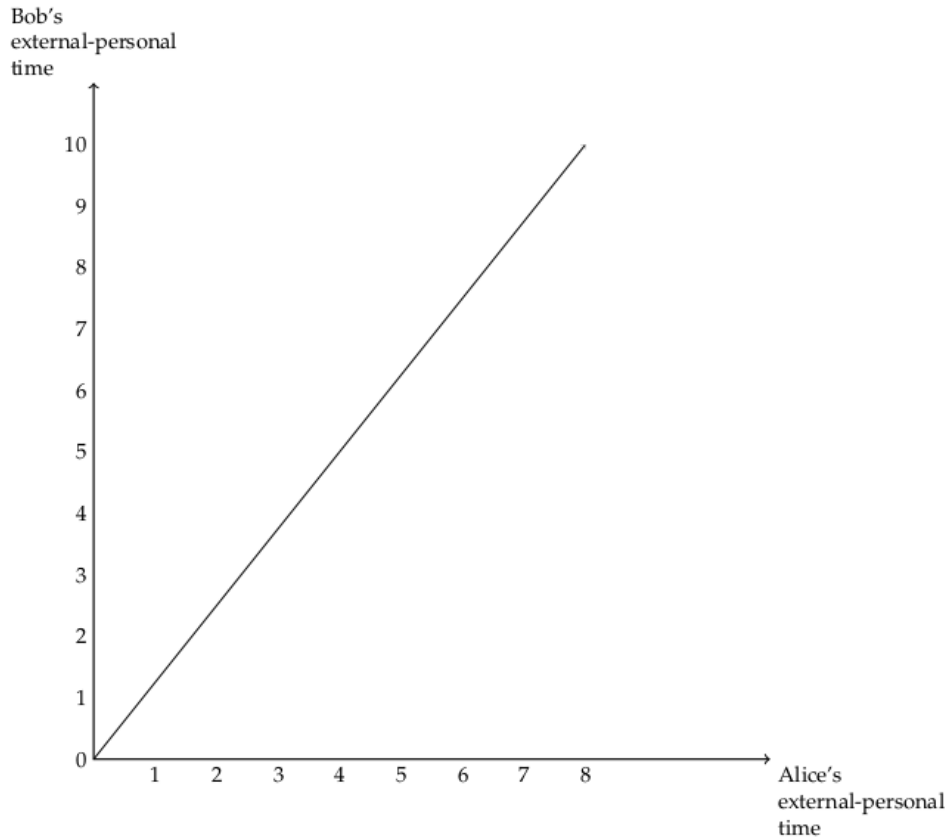


Figure 3.4: Alice and Bob's graph.

Thus far we have seen four cases of time-travel. The first two of them came from Lewis's paper, whereas the third and the fourth involved cryogenic sleep and Special Theory of Relativity. All of them has to do with a discrepancy between times. The first two were cases of backward time travel. The time traveler sets off and she arrives to a point in time which is past with respect to the time of departure. As Lewis observes, time travel toward the past may involve oddities. A world where backward time travel takes place would be a world very different from ours. For instance, time-travel towards the past may generate cases of backward causation, i.e. cases where an effect precedes in time its cause. The time traveler punches himself before jumping in her time machine and when she arrives in the past she has a black eye. Clearly, the punch causes the black eye. A cause is later in time with respect to its effect. Time travelers to the past

might also give rise to causal loops, i.e. cases where every event does have a cause but the whole loop does not. It might be that our time traveler brings along a book and gives it to the author before the author writes it, thereby creating a situation where the author copies his own work which will eventually be brought back in the past by her. The author can write the book because she just copies it, and the time traveler can bring it along because the author already wrote it at the time of her departure. Who wrote the book in the first place? The whole loop cannot be explained. It may also happen that the time traveler visit her younger self and talks to herself on the phone. But, Lewis observes, that is not a contradiction. It is just two distinct temporal parts of the same whole talking to each other, which is not the same thing bilocated in two different places at the same time. The external time of the conversation would be the same, whereas the personal time of the stages composing the whole time traveler would be different.

Suppose you have a time machine in front of you and you can go back in time with it. What can you do with it? Let us take Lewis's example. In 1975 Tim wants to go back to 1920 and kill his grandfather in 1921, who lived until 1957. The year 1921 lies in Tim's personal extended past, 29 years before his birth. It also lies in Tim's personal future before he steps into his time machine, because Tim travels to the past to live his personal future. Grandfather lived until 1957, hence he must be alive in 1921, if we set aside cases of resurrections. As Lewis points out, we may be tempted to speak of two 1921, the one which lies in Tim's personal past and the one which lies in Tim's personal future. Here, though, we are just referring to the same thing, the year 1921, with two different expressions. And, either the year 1921 contains grandfather's survival or it contains his murder. 1921 cannot contain both, or else we would have an inconsistent story. Tim goes back to kill grandfather, who wasn't killed in 1921. Can Tim change the past and make it the case that his grandfather will not be alive at the end of 1921? He cannot. Laws of logic would not permit this. Grandfather survived until 1957, hence he was alive during the year 1921 and nobody killed him. If Tim were to kill him in 1921, the year 1921 would be inconsistent. It would contain both Grandfather survival and his death, which is a contradiction.

To sum up, suppose you are at the time T_{10} and the earlier time T_0 contains the event ϕ . Can you go back to T_0 and make it the case that ϕ is replaced by ψ ? Either T_0 timelessly contains ϕ or it timelessly contains ψ . Either way, there would be no change in the past. Or, if you want to claim that there might be such a change, from ϕ to ψ , *when* this alleged change would take place? At T_{10} when you step in your time machine?

It can't be. You step in your time machine at T_{10} because T_0 contains ϕ . Perhaps T_0 might change from containing ϕ to containing ψ at T_0 , at the moment of your arrival in the past. But again, this can't be. T_0 is earlier than T_{10} , and T_{10} is such that it comes 10 units of time after the event ϕ took place at T_0 . The point is that in a one-dimensional time, it seems to be a non-sense to talk about a change in the past. Does this rule out the metaphysical possibility of time travel in one-dimensional time? We would like to travel in the past in order to change it, maybe to prevent all the evil and unpleasant events which happened in the past. But changing the past in a one-dimensional time is impossible, and just being around in a past place where you haven't been would amount to an alteration of the past. However, the possibility of time travel in a one-dimensional time is not ruled out. One could always go in a past place where she has already been and do what she has already done. This way, the past will not be altered. And in such a case, one could still have the possibility to affect the past. Suppose Tim goes back in the past and buy a packet of cigarettes. He is not altering the past. The event of him buying cigarettes was already part of the past before his departure. Yet, he is surely affecting the past, in the same way I am affecting the present if I go out and I buy cigarettes.

To sum up, changing the past is impossible, whereas affecting it is possible, provided one does what she has already done. I now turn to another story and I will then try to draw some conclusions from time-travel stories.

3.2 Another Odd Story

Here I want to tell another story, which will help us to draw a moral from time travel with respect to issues related to fatalism and the power to do otherwise.

Here is the story. It involves Mit⁶, a guy who goes backwards in one-dimensional time. Mit's birth takes place in 1985 exactly the 31st of December at 11:59 PM and 1 second (external time) and he dies somewhere in time in 1920 (external time). He is an infant between 1980 and 1985, and very old between 1920 and 1930 (external time). As external time goes forward, he becomes younger and younger, from his death to his birth. He has a personal wristwatch, which ticks at the same rate other clocks in the world tick. But his clock also goes backwards. That means that, for instance, when an ordinary clock is at 11:59 PM and 0 seconds in 1985,

⁶His name is Tim spelled backwards, for reasons that will be clear soon.

Mit's wristwatch ticks exactly 1 second, because at that time Mit is 1 second old, according to his personal wristwatch. When an ordinary clock is at 11:59 PM and 1 second in 1985, Mit's wristwatch ticks exactly 0 seconds, the starting point which coincides with Mit's birth. In principle we could even shoot photos of ordinary clocks and Mit's clock and say that they agree about the rate of change, but they have opposite directions. Of course, as far as Mit is concerned, he observes his clock going forward and other people's clocks going backwards. The opposite happens for ordinary people in the same world. They see Mit becoming younger and younger, digesting food before eating it, doing things before planning them. As far as ordinary people are concerned, their clocks are right and go forward, while Mit's clock goes backwards. Of course, it seems that there isn't an absolute way to settle this issue about who is going forward and who is going backwards. This story is odd, yet also possible, provided we accept Lewis's notion of personal time, which is crucial for bringing home his point about the possibility of time travel in one dimensional time.

It is also interesting that many features of Tim's story are mirrored by our Mit. Lewis says that Tim's time travel involves a discrepancy between external time and Tim's personal time. We have a very similar kind of discrepancy also in Mit's case: external time and Mit's personal time go in opposite directions. Tim's departure and arrival back in time are separated by two different amounts of time, 55 years of external time and are next to each other in personal time. In Mit's case, his birth and his death are separated by 67 years, according to both the external time and personal time. Yet, there would be disagreement between different ways of ordering the events. Mit would say that his death is later in time than his birth, while ordinary people would say that his death is earlier in time than his birth.

Tim's case involves backward causation. Tim punches himself before entering the time machine, and Tim has a black eye in the past after his arrival. According to external time, the effect is followed in time by its cause, while according to Tim's personal time the cause precedes the effect. Mit continuously provides episodes of backward causation. According to his personal time, he first lights a cigarette and then he smokes it, while people from the outside would observe that he first smokes a cigarette and then lights it. We have seen that backward time-travel in one-dimensional time might give rise to causal loops. So might Mit. When Mit is young, he might observe a book, copy it and give it to the original author, so that the original author can copy it as well.

Is Mit a whole unique person persisting through time? It seems that if Tim is, Mit also is. Mit has the proper mental connectedness. There is also the proper causality. According to Mit's personal time all his stages

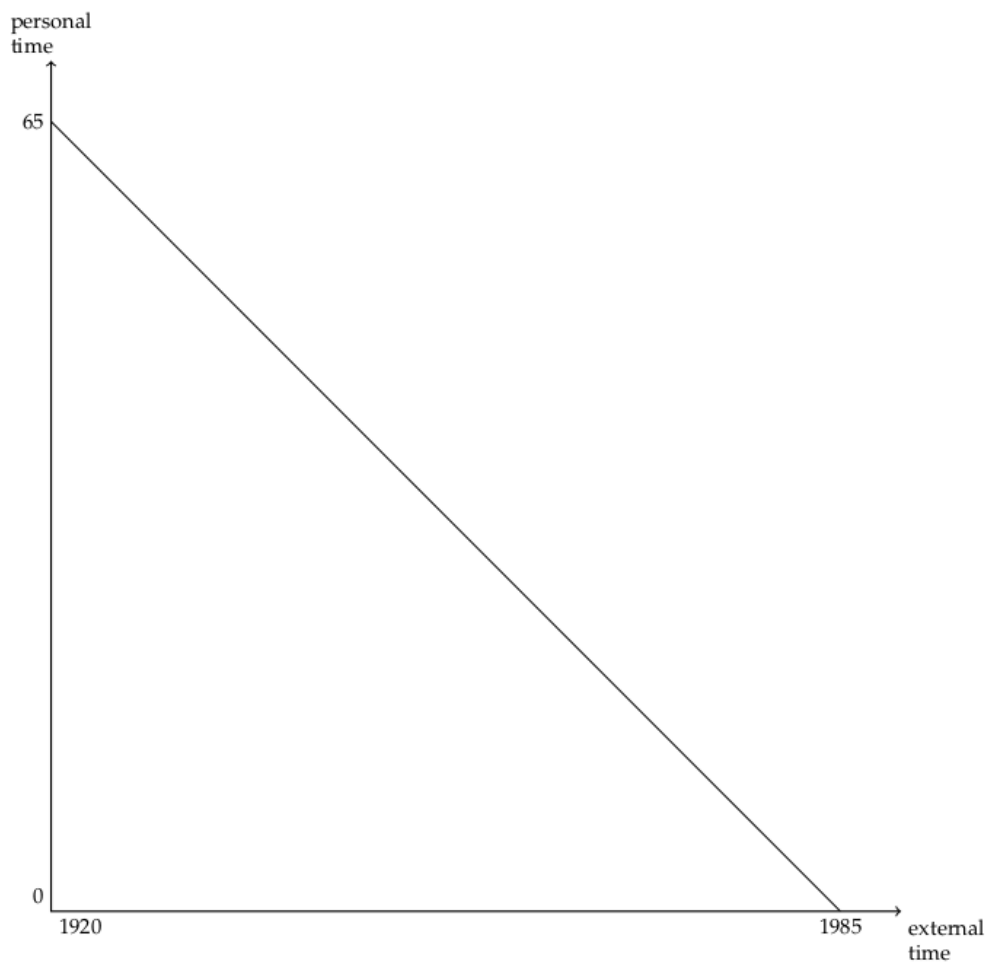


Figure 3.5: Mit's graph.

are continuous and causation works in an ordinary way. In fact, according to Mit's personal time, he eats first, then digests it. He lights a cigarette first, then he smokes it. He plans to do something first, and then he does it, and so on. If we use the graph we used for the previous cases of time travel, Mit's life would be a line with slope equal to negative 1, because as external time goes by, he would become younger and younger.

Is this story metaphysically possible? Probably a lot of issues related to physics might arise. I don't have the time nor the competence to deal with these issues. I just observe that if we think that time-travel in one-dimensional time is possible *and* we further assume that it might take time to go back in time, we would have something similar to Mit's case. Suppose the time-traveler goes back in time 10 years, and the trip takes two hours

of personal time. Let's say that she goes from 2016 to 2006 in two hours. Then, if time is one-dimensional and there are no spatio-temporal tunnels, she must be somewhere in space-time during her trip, at all times from 2016 to 2006. And since she is going back in time, people living normally would observe her in the way I described Mit's case⁷. They would observe her going backward, given that her experience in the time machine would probably be normal and lived forward.

We shall see in the next section why this case could be interesting with respect to issues about fatalism and the power to do otherwise.

3.3 The Specter of Fatalism with respect to Mit and Time Travel

Lewisian time travel to the past and Mit' story raise worries related to fatalism. Consider Tim before his departure. He has already been in the past, or else his time travel would not be possible. We have seen that Tim *can* go back to the past because he has already been there. He also *does* go back to the past, because this is simply what happens in his world. And, there is also a sense according to which Tim *must* go back to the past. Consider him at the age of 25 in 1975, before he steps in his time machine. His personal future stages are back in 1920, and there is nothing he can do to change this fact, given that the past is unchangeable. It seems that he can't refrain from stepping into his time machine, because if he did so, he would change the past, which is impossible. Hence, Tim can go back, does go back and has to go back. The previous sentence seems to be nothing else than the collapse of modalities we have been talking about in section 1.3.2, which is a peculiar feature of modal fatalism.

Now, consider Tim once he is back in the past. The actions he is performing and the events he is witnessing are part of his extended personal past. He might have a record of what happened during those times. History books, personal memories, photographs; all these sources of knowledge could be available to him and he could have brought them along. In such case, all that happens back there in the past, would be just a confirmation of those records. Nothing which differs from what is written in those record could happen, or else the past would change. Tim observes Tom, an ordinary non-time traveler folk, doing such and such at a certain time.

⁷The only difference is that in Mit's case I assumed that Mit' slope in the graph is equal in absolute value to the one of people outside, and just opposite in sign. In this trip from to 2016 the slope would be negative as in Mit's case, but also way greater than one.

This was recorded among his notes. Tom could not have done otherwise, because had he done so, Tom would have changed the past. Not only Tim, in virtue of being a time traveler cannot change the past. Tom as well cannot. Again, this seems to be nothing than the modal fatalistic principle T_c , which states that if something happens, it is necessary that it happens.

Suppose that Tim bought the time machine from a time-machine vendor. They could have had the following conversation. “Can I go back to 1207 with this product that I am about to buy?” and the answer would be “Although this machine is properly functioning, you *cannot* go to 1207, because you weren’t there. But you can go back to 1920.” Tim might then ask “Can I kill my grandfather in 1921?” and the only true answer would seem to be “no, you cannot because you didn’t”. This again sounds as a pure modal fatalistic kind of talk.

Consider now the story of Mit and compare it to Tim’s story. Tim will (personal time) travel towards an unchangeable and fixed (external) past, in order to do what he already did. The same is true for Mit throughout his entire life. We could even imagine that soon after (according to his personal time) he is born, someone gives Mit a book about his entire personal future. That book is completely detailed and it contains just true sentences about his personal future life. It was really easy for Mit’s biographer to write down the biography. The biographer didn’t predict anything at all. He just recorded what he had seen in Mit’s external past. As long as Mit’s life goes by according to his personal time, what is written in the book is continuously confirmed. Very soon, Mit might become a fatalist, and it seems he would have good reasons to do so.

Tim and Mit stories raised worries concerning modal fatalism. That is, what happens to them might incline us to accept modal fatalism with respect to what is going on with their lives. Is this a special feature of their stories? The answer is negative. Mit is going, as far as he’s concerned, toward events which are fixed, immutable and already there, namely his personal future. The same is true for Tim when he is about to take his journey to the past. But their situation is perfectly similar to the situation of ordinary people who live in a block universe, where all past, present and future events are ontologically on a par and fixed. It does not matter that Mit is a backward-moving system and Tim is a time-traveler; in a block universe everyone is going toward events which are fixed, immutable and already there. If what I said in this section about Tim and Mit shows that their lives are governed by fate in the relevant modal sense, then the same should apply to everyone who lives in a block universe.

3.4 The Specter fades away

All the worries we have been talking about in the previous section stem from a common source: a block and fixed universe. In a block universe, all past, present and future events or objects exist and are timelessly out there. In such view, the universe is depicted as a static block which does not itself change, and change is nothing more than qualitative difference between spatio-temporal proper parts of the block. What is present, past or future is just a matter of perspective. That is, there isn't an absolute present that in turn defines an absolute past or an absolute future. Rather, everything is already there and the spatio-temporal place one finds himself defines a relative present.

Tim's trip with his time machine brings him toward a past (from his perspective) part of the block. Mit continuously goes toward a past (from other people's perspective) part of the block. They are both going toward something which is fixed and immutable. But, we observed, ordinary people do the exact same. We are all going toward future (from our perspective) parts of the block. Yet, in a block universe those parts are as fixed as the past ones. To sum up, the worries from the previous section are generated by two features of a block universe. The first one is realism about the relative past and the relative future — the idea that when someone is at a certain spatio-temporal point, her future and her past are as real as her present — and the second one is fixity, i.e. the idea that the relative past and the relative future cannot change.

Are fixity and realism about the future and the past enough to imply modal fatalism? Although considerations as the ones we made in the previous section seem to suggest so, it is not the case that modal fatalism follows from fixity and realism. In fact, fixity and realism just show that the future, like the past, is already written. But modal fatalism is *not* just the idea that the future is already written. It claims something more. According to modal fatalism, the future is already written, *and* the way it is written is the only way it *could* have been written, given how the past has been. Fatalism is a modal doctrine which has to do with what agents *can* or *cannot* do. Fixity and realism just tell us something about what is actual. Something else must be added to jump to modal fatalism. In fact, the Main Argument, Taylor's argument and my argument from the second chapter aim to do so. The Main Argument tries to establish the validity of the schema T_c , Taylor's argument tries to establish that among two alternative courses of action only one is genuinely open and my argument tries to establish the validity of $T_c \diamond$. Fixity and realism by themselves are not enough. In fact, they can be coherently combined with theories which are

anti-fatalist.

One way to combine an anti-fatalist view with fixity and realism comes from the Thin Red Line theory already introduced in 1.2.2. According to the Thin Red Line theory, at any point in time only one future is actual, the one which is lit up by the the Thin Red Line and will actually obtain. Its actuality implies fixity. What happens on the Thin Red Line cannot be changed. Yet, it may be that at some point in time, perhaps because there are free agents equipped with the power to do otherwise, or because the world is indeterministic, other future continuations distinct from the actual one are possible. A standard way to give a Thin Red Line semantics is to say that a sentence $F\phi$ is true if and only if ϕ itself is true at some future point *on the Thin Red Line*. In such view, the operator 'F' takes into account only what happens on the Thin Red Line. We can then introduce the necessity and possibility operator which range on all possible continuations from a given point. This way, $\Box F\phi$ is true if and only if ϕ is true in all possible continuations and $\Diamond F\phi$ is true if ϕ is true in some continuation. It is easy to see that the schema T_c and $T_c\Diamond$, the two schemas which are constitutive of modal fatalism, would no longer be valid. In fact, it may be the case that $F\phi$ is true in virtue of the obtaining of ϕ on the Thin Red Line, whereas $\Box F\phi$ is false because ϕ does not obtain in all continuations. Similarly, $\Diamond F\phi$ might be true because ϕ obtains in some possible continuation, whereas $F\phi$ might be false because ϕ does not obtain on the Thin Red Line. Hence, the Thin Red Line would be a way to avoid modal fatalism and retain realism and fixity about the future.

Another way to combine fixity and realism with an anti-fatalistic stance comes from Lewis and his paper on the Paradoxes of Time Travel (1976). Modal fatalism is a doctrine which is interested in what agents can or cannot do. Lewis offers an account of "can" which has anti-fatalistic implications. Here is his point:

To say that something can happen means that its happening is compossible with certain facts. *Which* facts? That is determined, but sometimes not determined well enough, by context. (p.150)

As he says elsewhere, "can" is equivocal. That is, its meaning varies across contexts. We can illustrate what Lewis has in mind with an example. Alice and Bob are close to a swimming pool at T_1 . Alice took swimming courses and swam a lot of times before T_1 . Bob is afraid of water and has never swam before. As a matter of fact, neither Alice nor Bob swim at T_1 , so there is no water disturbance in the pool at T_2 . Lewis would say that there is a sense according to which Alice can swim at T_1 . She has what it takes to swim. When we say that Alice can swim, we are legitimately ignoring

some facts. In Alice's case, for instance we are ignoring external-future facts such as the absence of water disturbance at T_2 , which is incompatible with Alice's swimming at T_1 . These facts, Lewis would argue, are not relevant in determining what Alice can, in this sense, do at T_1 . The result is that she can swim at T_1 , even if she doesn't. Of course, Bob's situation is different. Bob cannot (in the same sense Alice can) swim at T_1 , because his swimming at T_1 is not compossible with the same type of facts we consider in Alice's case. For instance, the fact that he has always been afraid of water inclines us to say that he cannot swim at T_1 , in the same sense Alice can. As a corollary of this account of "can", we have a counterexample to modal fatalism. Modal fatalism can be described as the doctrine that no agent can perform an action which he or she doesn't actually perform. Take Alice. She does not swim at T_1 , but there is a sense according to which she can swim. Hence, this simple case goes against modal fatalism. And this is compatible with fixity and realism about future events. The event of there being no water disturbance at T_2 is already there and fixed. Yet, Alice can swim at T_1 , even if she does not.

I will have more to say about Lewis's account in the next section. What I wanted to highlight in this one is that realism about the future and fixity do not imply modal fatalism. In fact, we can have accounts like the Thin Red Line and Lewis's which combine anti-fatalistic stances with fixity and realism about the future.

3.5 Powers, Fatalism and Abilities

In this section I want to say something more about Lewis's account of "can" and how it relates to modal fatalism. In the *Paradoxes of Time Travel* (1976) he says:

Fatalists — the best of them — are philosophers who take facts we count as irrelevant in saying what someone can do, disguise them somehow as facts of a different sort that we count as relevant, and thereby argue that we can do less than we think — indeed, that there is nothing at all we don't do but can.
(p.151)

It is surely correct that according to a fatalist there is nothing at all we don't do but can. However, this is not precise enough. According to modal fatalism, if I can ϕ at T , then I ϕ at T , which in turn implies that whenever an agent does not ϕ he can't ϕ . But we must be careful with what sense of "can" we have in mind. It is true, as Lewis observes, that

“can” is equivocal. But, we can pinpoint different senses of “can”. Let us consider the following sentences.

- (50) I could travel faster than light
- (51) I can’t travel faster than light
- (52) I can’t travel faster than light with my brand new spaceship today

The sentences (50) and (51) might not count as contradictions, despite their appearance. In fact, with (50) I might mean that it is logically possible for me to travel faster than light. There is no contradiction in conceiving me as traveling in my spaceship at a speed of $1.5c$. This happens in other possible worlds. Hence, (50) would be true in this sense of “can”. (51) is different. Here I am referring to what is physically possible or impossible in our world in general, given the prevailing laws of nature governing our world. No matter how I try, no matter what I do, I am not going to travel faster than light in this world. (52) sounds odd and even misleading. In fact, it seems to have as an implicature that I could travel faster than light other days, even though I can’t today. This gives us another hint on the distinction between abilities and powers we already encountered when we were talking about Taylor’s argument in section 1.2.3. Suppose I have the ability to swim. One way to express this is to say that I can swim. Saying that I can’t swim *today* seem to express something different. It might mean either that I lost my ability to swim today or that external conditions are such that I can’t exercise my ability to swim today. In general, when we specify a time, we do so to refer to specific circumstances which take place at that time. By doing so, we typically switch the sense of “can” from the one which has to do with abilities to the one which has to do with powers. It is possible for the professional pole-vaulter to pole vault 14 feet. This is an ability he has. We can express this by saying that he can pole vault 14 feet. If we add that he can pole vault 14 feet *now*, we typically mean that conditions for pole vaulting 14 feet are ideal now. For instance, we imply that he is not locked in a room with a ceiling 10 feet high. This way of talking should give us another reason to have a sharp distinction between two senses of “can”. Abilities amount to what an agent can do in general. Powers are what is possible for an agent at a specific time in specific circumstances.

Lewis deals with fatalism in his paper about time travel, because issues related to fatalism seem to imply a contradiction in his story about the time traveler Tim. Here is the problem. As we have seen, Tim goes back in 1920 and tries to kill his Grandfather in 1921, who lived until 1957. As Lewis describes the case, Tim is well prepared to kill. He is at the peak of

his training, he has the best rifle money can buy, and he knows everything there is to know about shooting. As Lewis says here:

There's no doubt that Tim can kill his victim; ... by any ordinary standards of ability, Tim can kill Grandfather. (p.150)

and here:

Tim can kill Grandfather. He has what it takes. (p.149)

Hence, Tim can kill Grandfather. But, Grandfather lived until 1957, hence he was not killed in 1921. If Tim killed him, he would change the past, which is impossible. So, Tim cannot kill Grandfather. Hence the contradiction, we have that Tim can and cannot kill grandfather. If this were a genuine contradiction, the metaphysical possibility of time travel toward the past would be ruled out. Lewis explains away this apparent contradiction by saying that "can" is equivocal. I agree on this strategy. When we say that Tim can and cannot kill Grandfather, we are using two different senses of "can". But, those two senses are well pinpointed by the context provided by Lewis. Surely, Tim has the ability *to kill*. Again, we should note that when we talk about abilities, we tend to talk about what one can do in general. We say things like X has the ability to kill, or X has what it takes to kill. By contrast, with expressions like "to kill Grandfather" we are implicitly referring to a specific time and specific circumstances. That is, we are switching from a context where we are talking about abilities to one where we are talking about powers. It should also be noted that the only relevant sense of "can" which matters in the debate about fatalism and Free Will is the one about powers. In fact, we have already seen that it seems to be a plausible moral principle that if an agent couldn't have done otherwise than what she in fact has done, then she cannot be held morally responsible for what she has done. In general, moral responsibility is related to what agents can or cannot do. For instance, someone might be held responsible if she didn't do something she could have done. Let us now consider the following two situations. In both cases, a terrorist threatens the President of the United States. The terrorist will let a bomb explode, unless the President pole vaults 14 feet within an hour and puts the video on the Internet. In the first case, the President has never pole vaulted before and has no clue about pole-vaulting. She can't pole vault, she has not the ability to pole vault, thus she cannot be held responsible if she does not. In the second case, the President is a professional pole vaulter, yet at the moment when the terrorist threatens her, she is locked in a room with a ceiling 10 feet high there is no way she can escape from.

She has the ability to pole vault, yet she cannot pole vault 14 feet given the prevailing circumstances. That is, it is not within her power to pole vault 14 feet at the moment when the terrorist threatens her. Nobody would hold her responsible for not meeting the terrorist request, even if she has the ability to do so.

To sum up, ability and powers are clearly distinct senses of what is possible, impossible and necessary. Abilities have to do with what is possible in general, whereas powers deals with what is possible in specific circumstances at a given time. Having the ability to ϕ is a necessary condition for having the power to ϕ , but the converse does not hold. I might have the ability to ϕ without having it within my power at a specific time to ϕ . Powers are what matters within the debate about fatalism and Free Will. Modal fatalism can't be so easily dismissed by observing that at some times we have the ability to do something we don't actually do.

Ultimately, in Tim's case, Lewis is on the right track when he says that "can" is equivocal in order to explain away an apparent contradiction. However, a more precise way to explain it away is to say that Tim can kill, and Tim cannot kill *Grandfather*. In other words, Tim has the ability to kill, but it is not within his power to kill Grandfather. If this latter sentence is correct, Lewis's account of "can" does not threaten modal fatalism at all. To this issue I now turn.

3.6 Protecting the Past

Tim goes back in time and tries to kill Grandpa in 1921. But Grandpa survived 1921 and the past cannot be changed, or else we have contradictions. Here is what happens according to Lewis:

You know, of course, how the story of Tim must go on if it is to be consistent: he somehow fails. Since Tim didn't kill Grandfather in the "original" 1921, consistency demands that neither he does kill Grandfather in the "new" 1921. Why not? For some commonplace reason. Perhaps some noise distracts him at the last moment, perhaps he misses despite all his target practice, perhaps his nerve fails, perhaps he even feels a pang of unaccustomed mercy.... Success at some tasks require not only ability but also luck, and lack of luck is not a temporary lack of ability. (p.150)

Someone or something must stop Tim's attempt to kill Grandpa. The past has to be protected!

Tim is perfectly able to kill, Tim tries to kill Grandpa, and eventually Tim fails. Something or someone *must* prevent Grandpa's murder. We can pin down two different senses of the previous sentence. According to one sense, we say that

(53) necessarily, some event prevents Grandpa's death.

According to another sense, we say that:

(54) some event prevents Grandpa's death, and its occurrence is necessary.

According to the former sense, necessarily one preventor or another will make Tim's attempt fail. Say that some noise distracted him. Had the noise not distracted him, his nerve would have failed, or he would have slipped on a banana peel, or he would have felt mercy, and so on. One way or another, something must happen and make it the case that Tim does not kill Grandpa. This seems to me to be a case of Lazy fatalism, as I characterized it in section 1.1.2. According to Lazy fatalism, future events will happen no matter what one does. No matter how Tim is prepared, no matter how Tim is well equipped, no matter how he shoots, Grandpa will survive. What is essential to Lazy fatalism is the denial of the Causal Connectedness thesis, i.e. the idea that what happens at a time may causally depend on what happens at other times. In fact, if we interpret the necessary occurrence of the preventor as the idea that necessarily, something will prevent the murder, we end up lacking the desirable causal connection among events. Suppose that what actually happened is that Tim's nerve failed, and because of that his shot was not precise enough and missed the target by few inches. His nervous breakdown caused the bullet to have a missing trajectory. But, according to (53)'s interpretation, and a standard counterfactual analysis of causation, this is not the case. Take the counterfactual had Tim not had a nervous breakdown, the bullet would not have had a missing trajectory. It turns out to be false in Tim's world. In fact, there isn't a possible world where Tim does not have a nervous breakdown and the bullet kills Grandpa, which is closer to any world where he does not have a nervous breakdown and the bullet has a missing trajectory. Worlds with a missing trajectory are closer than worlds where the bullet hits and kills Grandpa. This is so because Grandpa has to survive 1921, because his survival explains Tim's attempt in the first place. Had Grandpa been killed, Tim would have not been there in the first place. Hence, the relevant counterfactual turns out to be false and the counterfactual analysis (wrongly) does not count the nervous breakdown as an episode of causation. The problem here is probably due to the fact

that according to (53), necessarily, some contingent event must make Tim's attempt fail. As a result, the specter of Lazy fatalism, the worst kind of fatalism according to anyone, arises. If (53) is the correct interpretation, we seem to relinquish causation among events.

What I have said so far should incline us to favor (54). An event, say Tim's nervous breakdown, prevents him from killing Grandpa, and its occurrence is necessary. This is something a modal fatalist, and a reasonable fatalist as well, could happily embrace. In fact, according to a modal fatalist any event is necessary⁸. The nervous breakdown prevents the act of killing. If my argument in Chapter 2 gets through, the presence of this preventor should lead us to say that it is not within Tim's power to kill Grandfather, because the absence of it is a necessary condition for Tim to kill Grandpa. Maybe traveling back in time as Tim did, has as a side-effect a more likelihood of nervous breakdowns. This does not affect Tim's ability to kill. But because Tim was about to shoot a relative of him, he got all emotional and a case of nervous breakdown was bound to happen. A series of causal chains made it the case that it was not within Tim's power to kill Grandpa. That is, it is necessary — in the relevant fatalistic sense — that he fails. This fatalistic modal necessity gives us enough protection of the past. What happened was necessary, it was the only thing that could have happened given Tim's personal past and the external circumstances. It wasn't within Tim's power to kill grandfather, even though Tim has the ability to kill. This should not bother us. Many times we are able to do things which are not in our power to do. For instance, when I am tied to a tree, it is not within my power to swim, even though I am able to. On a side note, the metaphysical possibility of Lewisian time travel in one dimensional time is not ruled by modal fatalism. There is no contradiction to explain away. Tim *can* kill, i.e. has the ability to kill, and Tim *cannot* kill Grandfather, i.e. it is not within his power to do so.

Can (54)'s interpretation make sense of the causation episode with Tim's nervous breakdown? Causation is n-transitive. If event A causes B, and B causes C, C is causally dependent on A, because had A not happened, B would not have happened and neither would C⁹. This can be iterated n times in a causal chain. Suppose John one day turned left instead of right. He turned left on his way to work because he wanted to enjoy the view of the park which was on his left. This might seem a choice with a little impact on his life. A tiny detail in John's life. However, it turns out

⁸or purely chancy. We can set aside pure chanciness here, because it is not a case of nomological indeterminism.

⁹We set aside cases of overdetermination, by saying that without A and B, maybe an event of the same type of C would have happened, but not C itself

that by turning left, John meets a woman and falls in love with her. He later marries her, move to another continent, has kids with this girl and live an happy life. John's life has changed just by making a left on that day. It changed from the unactualized way it would have been, had he not made a left on that day, to the way it actually is. This is not a real and problematic change as the one we were worried about with Tim's attempt to kill Grandpa. It is just another way to express causal dependence among events. It might seem odd, but it is correct to say that had he not made a left that day, John would not have played with his first kid 10 years later in another continent, even though these two events are separated in time by ten years. This again, is due to the fact that causation is n-transitive. Had he not made a left, the whole causal chain I mentioned — the meeting with his future wife, the marriage, the move to another continent and so on — would have been wiped away from actuality.

In Tim's case, we have a causal chain as well. According to Tim's personal time, Tim was born, then he decides to go back to the past, then he steps in his time machine, steps out, then has a nervous breakdown, and eventually the bullet misses Grandpa. All these events are causally connected, exactly in the same way John's events are causally connected. Had he not decided to go back to the past, he would not have stepped out of his time machine in 1920. Likewise, for the other events I mentioned. Had he not decided to go back to the past, he would not have had a nervous breakdown in 1921. However, Tim's case is different. Tim travels back to the past, thereby giving rise to episodes of backward causation. This in turn gives rise to an interesting phenomenon, which seems to amount to a pure case of causal loop. The bullet has a missing trajectory and this event causes Grandpa's survival. By a series of causal connections, then Tim was born. He then goes back, has a nervous breakdown which causes... the missing trajectory. A causal loop. Moreover, because of n-transitivity of causation, it is also the case that the missing causes the earlier Tim's nervous breakdown. This is so because we can move from the missing to the nervous breakdown in n steps within the causal chain. Is it correct to say that had the bullet not missed Grandpa, Tim would not have had a nervous breakdown? It seems odd, especially because the nervous breakdown comes earlier in external time than Tim's nervous breakdown. But we should expect some kind of oddities, given that Tim's world is a world where time travel takes place and hence backward causation takes place too. And, there is a sense according to which this counterfactual turns out to be true. In fact, had the bullet not missed Grandpa, Tim would have never been born and thus we would not have Tim's nervous breakdown.

Back to the relevant counterfactual again. A modal fatalist who accepts

(54) has to claim that the nervous breakdown causes the missing, or else we are back to an unpleasant case of Lazy fatalism. Can he or she do so? What would have happened if Tim's nervous breakdown had not occurred? Probably the bullet would have hit Grandpa, thereby killing him. As a consequence, the whole causal loop — Grandpa's survival until 1957, Tim's birth, Tim's time travel and so on- would be wiped away, the way John's causal chain would have been wiped away, had he made a right. Who would have shot the killing bullet? Probably someone else who is not Tim. Hence, a modal fatalist can claim that the relevant counterfactual is true, and therefore causation is safely rescued and Lazy fatalism avoided. That is, the modal fatalist can safely be *reasonable*.

To sum up, modal fatalism provides a straightforward way to protect the past from changes. It was never the case that it was within Tim's power to kill Grandpa. The event of the bullet missing Grandpa is as necessary as everything else. Moreover, Tim's travel back to the past seems to give rise to a causal loop and odd causal connections. The causal loop, together with n-transitivity of causation, allows the modal fatalist to explain why the missing causes Grandpa's survival and why Grandpa's survival is not a case of lazy fatalism.

Chapter 4

Changing the Past

Spirit walk sounds much safer
than time travel

Heroes - Third season

Thus far, we have been assuming that the past *cannot* change. For instance, the Main Argument we've been talking about in 1.2.2 rests on prior truths that lie in the past, which cannot be changed in virtue of being past. Taylor's argument in 1.2.3 for fatalism toward the past rules out freedom of action with respect to actions which would change the past if performed. In Lewis's account of time travel in one-dimensional time, a time-traveler can affect the past but he or she cannot change it. That the past cannot change is a common and shared assumption. What lies behind us is now unavoidable, in the sense that there is nothing we can do now to prevent what happened or bring about what have not, no matter how we try and what our powers are. Even though we might have a changing past in a multi-dimensional time model, almost everyone agrees that the past cannot be changed if time is one-dimensional. However, one of the most fascinating aspects of philosophy is the fact that almost any assumption, even the ones which are seemingly undoubtable and uncontroversial, can be called into question. In this spirit, Loss (2015) recently argued for the idea that the past *can* change, even if time is one-dimensional. In this chapter I first illustrate the assumption Loss employs to reach his conclusion that the past can change. I then analyze his own example which should show a case where the past changes. I will then argue that two assumptions he makes yield two problematic results. Ultimately, I will try to show why those assumptions should be rejected in favor of an

alternative view, which accounts for Loss's example and does not have as a consequence that the past change.

4.1 An Attempt to Change the Past

In this section I am going to introduce the example Loss employs to show that cases where the past changes are metaphysically possible, even if time is one-dimensional. He argues for this metaphysical possibility by discussing an example of his own, which involves a time-traveler who goes back in the past thereby changing it. I want to stress that Loss's aim is to defend the metaphysical possibility of a past-changing time travel *in one-dimensional time*. Before seeing the example itself, we should first check what assumptions Loss makes.

The first assumption Loss needs is an eternalist ontology, i.e. the idea that worlds are static unchanging blocks made up of a four-dimensional manifold.

The second assumption he makes is an exdurantist account (or stage-theory) of persistence, according to which continuants are instantaneous stages which persist and possibly change in time in virtue of having qualitatively different *counterparts* located at different times. For instance, *I* persist and change in time from T_0 to T_1 , because my instantaneous stage is, say, sitting at T_0 and there is a standing-counterpart of *me* located at T_1 . What makes my stage at T_0 and my stage at T_1 counterparts of each other is a relation of continuity and similarity, which sticks together the stages, even though they are qualitatively different.

The third assumption Loss makes is the fact that the whole world itself can be treated as a continuant. Standardly, the stage-view, or any theory of persistence for that matter, deals with *things* — objects, persons or events — which persist and possibly change in time. Loss extends the stage-view to world-T-slices where a world-T-slice is "is the mereological sum of all the entities x , such that x exists at T , and only at T (p.4)". Under this view, *the world* persists and change in time the same way things in the world persist and change in time. That is, a world persists and change in time in virtue of the existence of T-slice counterparts located at different times.

The last assumption Loss needs in order for his point to get through is the fact that more than one time can be attached to a single T-slice. We shall see what this amounts to by means of an example Loss himself provides. Loss makes us ponder the following possible world W_1 . W_1 is a qualitatively duplicate of our world, except for the fact that in W_1 there

is a 10 years break after the last¹ instant of 1989. All continuants which exist at the last instant of 1989 in W_1 , which is qualitatively identical to our last 1989-instant, are completely annihilated and replaced by a bunch of spheres bouncing around in an otherwise empty space for ten years. After this 10 years break, all the spheres are annihilated and replaced by a slice which is qualitatively identical to our first 1990-slice. W_1 is a block universe as any other possible world, given that we are assuming eternalism, where *all* its T-slices are tied together by an earlier-later relation which tells us how the slices are ordered. However, Loss argues, in some cases, we are entitled to consider just *some* T-slices which make up a world. That is, whereas a world is an unrestricted composition of all its T-slices, sometimes it makes sense to talk of *restricted* compositions of T-slices. Take the initial segment of W_1 from its beginning up to the last 1989-instant *and* its final segment which starts from the first instant after the spheres break. This is a proper part of W_1 , let's name it W_{1a} , which is tied together by a relation of relevant similarity and regularity. That is, within this restricted composition we have the same kind of regularity and continuity which is supposed to hold in our world. As Loss observes, people in W_1 located in the year which is qualitatively identical to our 1990 will say things such as "The Berlin's wall fall occurred one year ago", even though it happened eleven years before, given the 10 years spheres break. This kind of talk gives us a hint for what we were seeking for, i.e. the possibility to attach more than one time to a single T-slice. In fact, it seems we can say that W_{1a} has its own time. Not only all W_{1a} -slices are ordered by an earlier-later relation, but also a time-talk throughout all W_{1a} seems to make sense. That is, it seems that there is room to claim that the first 2000-slice of W_1 is the first 1990-slice, according to W_{1a} *internal time*. We have already encountered the notion of *personal* time when we were dealing with Lewisian time travel in section 3.1. There, the notion of personal time proved to be useful in cases of time travel and it has to be understood as what is measured by things like the time traveler aging processes, the number of heart beats, wristwatches ticking and the like. Loss extends this notion to a whole world-slice. In W_1 , for instance, people in 2000 external time are one year older than they were in 1989 external time, calendars read 1990 and so on. As a result, more than one time can be attached to the year 2000. It is the year 2000 according to external time, and it is the year 1990 according to W_{1a} internal time, where "internal time" refers to the whole world-slice version of personal time.

¹For the sake of simplicity, I shall assume throughout this chapter that time is discrete and non-dense. This way we can safely talk of the first, second and last instant of a given interval. This assumption is neutral with respect to what I want to say.

Having clarified the assumptions Loss makes, we are ready to deal with his example which purports to show the possibility of changing the past by means of time travel. Loss makes us consider the world V_1 . V_1 is qualitatively identical to our world from its beginning up to the last instant of 2009. But then, at the very last instant of 2009, Tim is in his brand new time machine and press the button. The next instant, he finds himself in a T-slice which looks qualitatively identical to the first 1920-slice, except for his presence. As it is seen from the outside, V_1 is a world where the last instant of 2009 is followed by a final segment which starts with a 'second' 1920 and then develops regularly. Here the notion of internal time can be summoned. In fact, V_1 from its beginning up to the last 2009-instant displays the same kind of regularity which holds in our world. Loss names this restricted composition V_2 . And, there is another relevant regular composition, i.e. V_1 from its beginning up to the last 1919-instant together with the segment from 2011's first instant external to V_1 'end. Loss names this restricted composition V_3 ². As we have seen, the fact that there are regular restricted compositions in V_1 , V_2 and V_3 , allows us to attach their own time to them. As a result, the place where Tim finds himself after his time travel is 2011 according to external time *and* the year 1920 according to V_3 's internal time.

Did the past change? Yes, the year 1920 changed from being Tim-free to containing Tim, provided we accept all the assumptions Loss makes. Take the first T-slice of the 'original' 1920. Within Loss's account, something changes as long as there are qualitatively different counterparts located at different times, and this can apply to whole world slices as well. In V_1 , *there is* a counterpart of the 'original' 1920, namely external-2011 which is 1920 according to V_3 's internal time. Hence, 1920 changed. Here is Loss's final point which states that Tim has changed the past by means of time travel:

(Internal) times themselves can thus be seen as persisting and changing in (external) time. Within an exdurantist account of persistence, to say that the year 1920 has changed is to say that the year 1920 is a continuant which has a qualitatively different past counterpart. This kind of change is precisely what is the case in Tim's universe according to the theory I am presenting. The internal 1920 changes from being Tim-free, to being Tim-ful. It persists and changes in (external) time by having qualitatively

²Actually, as Loss observes, V_3 is *quasi*-regular. In fact, in V_3 Tim appears out of thin air. This kind of *local* irregularity should not bother us, though. Tim's case is a time-travel case. Hence, some local irregularity has to be expected.

different temporal counterparts: one occurring in 1920 external time, the other occurring in 2011 external time. (p.10)

To sum up, I think Loss's case shows that, even if time is one-dimensional, it is possible to change the past, *provided we accept his assumptions*. In the next section I will try to show that two assumptions Loss makes yield problematic results.

4.2 Two Problems

In this section I want to highlight two problems which result from Loss's account of change over time and the assumptions we have seen. The first one is that the past seems to change too much. The second one has to do with the fact that it is highly doubtful that Loss's example should uncontroversially count as an episode which shows the possibility of changing the past by means of time-travel.

As for the first problem, take the world W_2 . W_2 displays the kind of regularities which we take to hold in our world. Nothing bizarre and odd like episodes of time travel occur in W_2 . Now, let's take, say, the first and second 1930-slices. Suppose they are qualitatively highly similar — from the first to the second slice, everything stays at rest, except for a single particle which moves 2 cm to the right. What should Loss's account say in such a situation? We have seen that according to his account, continuants change in virtue of having earlier or later qualitatively different counterparts, and this apply to whole-world slices as well. Loss claims that in his example V_1 , the year 1920 *changes* in virtue of having a later counterpart (2011 external time) which is qualitatively different from it. The same applies with our particle case. The first 1930-slice does have a later qualitatively different counterpart, namely the second 1930-slice. The first 1930-slice is also earlier than the second one, i.e. it is past with respect to the second one. Hence, Loss's account yields the result that the first 1930-slice has *changed* from having our particle at a certain position to having it 2 cm to the right. But this clearly should not count as an episode of a changing past. It is simply a case where a continuant, our particle, changes and persist in time in an ordinary way. The moral to be drawn is that if Loss's example should be counted as a case of changing past, then several other possible cases should be accounted in the same way. The problem is that those other cases seems to be ordinary cases of change in time, and not cases where the past itself changes.

Here is the second problem. As we have seen in the previous section,

Loss argues that in his example concerning the world V_1 , we have a case where the past changes by means of time travel. Tim presses the button of his time machine the last instant of 2010 and immediately afterwards he finds himself in what is 1920 according to V_3 's internal time. I want to argue that there is a different way to tell the story about what happens in V_1 . One might say that it is not Tim that time travels to the past. Rather, what happens is that *all* continuants, besides some which would make room for Tim's presence, in the last 1919-slice time travel to the future to the external 2011. Such case would be a case of asynchronic fission — continuants in the last 1919-slice would have counterparts in external 1920 and external-2011. This would be a story according to which Tim does *not* travel to the past, and hence we would not have an episode of time travel which changes the past.

Is there in V_1 a fact of the matter which grounds that what happened is that Tim time travelled to the past rather than all continuants in 1919 travelled to the future? One might argue that Tim should be counted as the time-traveler because he steps into a time machine and activate it by pressing the button, whereas 1919 continuants clearly don't. I don't think this could do. Even though we are familiar with time machines from science fiction stories, their presence seems not to be a necessary condition for having a genuine case of time travel. In the Twin Paradox case we have seen in section 3.1, one is able to time travel to the future by 'simply' using a very fast spaceship, without the help of any time machine. Or, one could imagine an episode of time travel which happens in virtue of some sort of magic power which make a person jump from one space-time to another one. Or, it could be simply a brute fact that in some worlds people time travel without the help of any time machine. Hence, the presence of a time machine could not be taken as a necessary condition, and therefore it cannot ground the distinction between our two possible stories about what happens in V_1 . Another strategy to ground the distinction might come from the fact that if it's Tim that time travels to the past, then we have a *local* case of irregularity, whereas if it's all continuants from 1919 which jump to the future, the irregularity is much bigger. That is, V_3 — the first segment of V_1 up to the end of 1919 external time and 2011 external time onwards — displays just a local irregularity. In V_3 , at some point Tim appears out of thin air and everything else displays the same kind of regularity we have in our world. On the contrary, in the passage from 2010-last instant to 2011-first instant, the irregularity is much bigger. Everything but Tim is annihilated and replaced by something else, whereas Tim persists in an ordinary way. However, this kind of strategy seems to have difficulties. How much local an irregularity must be in order to have

a genuine time travel case? How big enough the irregularity must be in order to *not* have a time travel case? In other words, where should we draw the line? Suppose it's not just Tim who steps in his time machine, but rather it's his entire family. Or suppose it's his whole city who does it. Or, his whole country, continent, planet, galaxy which do the time travel, in some odd way. It is not clear where the line should be drawn. Hence, appealing to local irregularities versus huge ones does not seem promising. Hence, this criterion cannot tell us whether in V_1 Tim time travels to the past or 1919-continuants travel to the future. Moreover, I fail to see what else could ground the distinction. Therefore, it is not clear that Loss's case should count as an episode where Tim time travels to the past thereby changing it.

To sum up, we have encountered two problems for Loss's account. The first one has to do with the fact that the past seems to change too much in his account. The source of this problem is Loss's third assumption, i.e. the idea that T-slices can be treated as continuants which change in time in virtue of the existence of earlier or later counterparts. The second problem stems from Loss's fourth assumption, the idea that we can safely and intelligibly attach more than one time to the same T-slice. In fact, Loss argues, the first 2011 external time-slice of V_1 can be treated as 2011 external time and as 1920 V_3 's internal time as well. As a result, it is not clear whether it's Tim who time travels to the past or it's 1919-continuants which travel to the future. Those two assumptions yield these two problematic results. In what follows I will try to argue why those two assumptions should not be made.

4.3 Times and T-slices

We have just seen that attaching more than one time to the same T-slice can bring about problematic results. In Loss's example, the same T-slice is 2011 according to external time, *and* 1920 according to V_3 's internal time. We have seen what *internal* time in Loss's account amounts to—something which is modeled on the notion of Lewis's personal time and applied to maximal restricted compositions of T-slices, unified by a relation of relevant similarity and continuity. But what is then *external* time in his account? Loss identifies external time with the earlier-later relation, as it is clear from the following two quotes.

Each T-slice can be seen as a world-counterpart, and the earlier-later relation linearly ordering the T-slices can be taken to be the

genidentity relation for the world (that is, the relation making all the T-slices stick together as counterparts of a certain world; (p.4)

and

As we saw above, a T-slice (that is the mereological fusion of all the entities x , such that x exists at time T , and only at T) can be considered both as an external world, and as an internal world. In the first case, the genidentity relation making a T-slice x a counterpart of a certain T-slice y is just the earlier-later relation. (p.9)

The earlier-later relation is commonly taken as a primitive, at least in an eternalist ontology which Loss assumes. It does the job of ordering all the T-slices belonging to a possible world by telling you for any pair x, y of T-slices, whether x is earlier than y , or y is earlier than x ³. With this primitive at hand, you can then account for what it is for something to be future or past, *relative* to something else. For instance, an event A belonging to a T-slice x is future with respect to an event B belonging to a T-slice y if and only if y is earlier than x . Likewise, the earlier-later relation accounts for what it means for something to be past relative to something else. Does the earlier-later relation necessarily tell you that a time-point is, say, *50 years* earlier than another one? The answer is in the negative, because it depends on which nature it turns out to have. Before seeing why things are so, we should make a short digression into the distinction between a substantivalist and relationist account of time.

Substantivalism about time is the view according to which time is a 'substance' that exists independently of events and things located in time. In such view, time is like a 'container' which is — and might not very well be — filled with events and things. On the contrary, a friend of relationism would say that time is construed out of events and things located in time, and as such it is *not* independent of events and things located in time. A lot more could be said about this long-standing controversy. For our purposes here, we could summarize what we need with a slogan. According to substantivalism, first there is time, then events and things. According to relationalism, first there are things and events, then time.

³The earlier-later relation does this job for *any* pair of T-slices, as long as time is linear. In branching-time models we might have pairs of times which are not related by such relation. We set aside this case, given that Loss's example is a case where time is linear. I am setting aside also considerations from special theory of relativity, where what is primitive are spatio-temporal intervals.

With this distinction in mind, we can come back to our previous subject. What does it mean to say, for instance, that World War II's beginning is *6 years-earlier* than its end? I take the previous claim to be significant as long as there are clocks in the world this claim is about. We might agree with Saint Augustine that when we are asked, we don't know what time is. Yet, it seems we have a pretty clear grasp of what is the length of a *time interval*. A time interval is what is measured by a clock. A clock is not necessarily an ordinary clock we are used to in our everyday lives. A clock is any device that keeps going through repetitive cycles. In this sense, the swinging of a pendulum, the Earth rotating around the Sun, a beam of light bouncing up and down in a box with two mirrors, human hearts beating and the like are all clocks. Of course, some of the clocks mentioned are more precise than others. But in principle, they all do the same job. They behave cyclically, and by doing so they determine a time unit, which in turn can be exploited for making measurements. That is, you count the times the cycle completed itself and that tells you how much time has passed. Once you have a master clock, you can set up other clocks which agree with the master one⁴. So, to say that World War II's beginning is 6 years-earlier than its end is to say that the Earth rotated 6 times around the Sun from war's beginning to its end⁵. In terms of T-slices, to say that World War II lasted 6 years — which is equivalent to say that its beginning is 6 years-earlier than its end — is to say that there is a T-slice containing the World War II's beginning and the Earth-Sun system in such and such position, and there is a later T-slice containing World War II's end and the Earth-Sun system in such and such position.

The procedure just described is how *we* make and understand time measurements. Are things fundamentally so? Is the length of a time interval nothing over and above than what is determined by the procedure just described? The answer is in the affirmative, if friends of relationism are correct. In fact, according to relationism, events and things come first. If so, then what determines the length of a time interval can be just the presence of clocks and the measuring procedures just discussed. On the contrary, if substantivalism is the correct view, I think we may have two

⁴Synchronizing clocks separated in space is not as unproblematic as one might think, for at least two reasons. Within Special Theory of Relativity, synchronization of clocks is relative to a frame of reference, as it is simultaneity. Moreover, synchronization relies on the convention that in a round trip the speed of light is the same in any direction

⁵Of course this is not accurate enough. There are clocks more precise than the system Earth-Sun which we should rely on. But again, here we are interested in what is the nature of the length of a time interval, and the system Earth-Sun can serve such a purpose as an example of a clock.

sub-cases. It depends on whether the substantialist 'container' comes with a metric or not. That is, whether the 'container' not only tells you for any pair of times x, y whether x is earlier than y or y is earlier than x , but it also tells you things like x is n -units-of-time-earlier than y . If the 'container' does not come with a metric, then it seems to just be the earlier-later relation as applied to time-points which are then filled by things and events. If so, we would be in a case which is similar to the relationist side when it comes to determining the length of time intervals, and hence the same considerations would apply. If the container does come with a metric, then the length of a time interval depends on the structure of the fundamental metric. To sum up, either the substantialist 'container' comes with a metric or it does not. If it does come with a metric, Loss's example does not get through, because the time intervals are established at the fundamental level. If so, by stepping in his time machine, Tim goes to a point which is fundamentally future and one instant later, and thus it cannot be counted as past as Loss wants it to. No matter how that T-slice is filled — because first comes time and then events and things which fill it — it is future with respect to the one where he he departs. Hence, the only viable option for Loss between the two substantialist sub-cases is that the 'container' does *not* come with a metric, but rather it just determines what is earlier and what is later. To sum up, independently of whether substantialism or relationism is the correct view, Loss can identify external time with an earlier-later relation which does *not* come with a metric, or else his own changing-past example would lose all its force.

The earlier-later relation gives you a time-series, i.e. it gives you all the times, or T-slices composing a possible world, in an ordered fashion. An useful metaphor to introduce people to possible world is the one which says that possible worlds are like fiction books. Our actual is world develops in such and such a way, but, it is said, it might very well have been different. Authors of fiction book are pretty much at liberty when they are writing a fiction. So are we when we imagine a possible world. As long as contradictions are avoided in the story, the result is something which is possible. Now, consider the following misprinted book. On its even pages, Fahrenheit 451 by Ray Bradbury is printed, whereas Crime and Punishment by Fyodor Dostoyevsky appears on the odd ones. Such a book may be thought of as a possible world where a lot of oddities take place. 'At the end of any even page', Guy Montag, his companions and surroundings are annihilated and replaced by Raskolnikov's universe. 'At the end of any odd page', Montag will restart exactly where he has left. In terms of T-slices, such a world would be composed of an alternating sequence of T-slices containing Montag-universe and Raskolnikov-universe, all ordered

in an earlier-later relation. In such world, it would make sense to say that, for instance, Montag meets Clarisse as he walks home, and *1 hour later* he comes home and find his wife Mildred near dead, even though the two events are located 'on different pages'. However, it couldn't be intelligible a claim such as "Raskolnikov murders Alyona *two years later* than Montag meets Clarisse". Each story has its own time, even though they are both tied together in the same possible world by an earlier-later relation. The most we could say would be that a given Fahrenheit 451 event is earlier or later than a Crime and Punishment one, again in virtue of the earlier-later relation.

The misprinted book example and the world associated to it can lead us to a distinction which can be useful, i.e. the distinction between what I shall call a time-series and a time-measurable-series. A time-series is the sequence which result from the mereological sum of *all* T-slices belonging to a world, ordered as the earlier-later relation dictates. A time-measurable-series is a maximal (and not necessarily proper) part — a maximal mereological sum of T-slices— of a world such that: a) it exhibits the same kind of regularity found in the actual world, except for episodes of time travel and b) it is possible to keep the time throughout all T-slices composing it. Clause (b) has to be understood as the requirement of the existence of clocks, defined as persisting objects which behave cyclically. It does not have to be the same persisting clock throughout all the time-measurable-series. Sometimes we may rely on one master clock according to which we synchronize the other clocks, and then switch to another device. The switch is possible as long as there is a time where the two clocks are persisting together and hence we can check at which rate they are respectively ticking.

Here is how the time-measurable-series notion is supposed to work. For instance, our misprinted-book-world has one time-series and two time-measurable-series. Fahrenheit 451-slices compose a time-measurable series with its own measurable time. Likewise for the Crime and Punishment-slices. What must interest us, is the fact that two time-measurable-series cannot overlap, i.e. they cannot have a T-slice in common. In fact, assume *per absurdum* A and B are two distinct time-measurable-series with a T-slice in common. Then, throughout all A it is possible to keep the time, and the same holds for B. Given that A and B have a T-slice in common, it is possible to time-measure events in A and B, and compare the length of time intervals by means of the same unit. If so, A and B are one and the same time-measurable series by definition, and we end up having a contradiction with what we assumed. Hence, two time-measurable-series cannot have a T-slice in common. This gives us the result we were looking

for. It is not intelligible to attach two times to a single T-slice. Given that a T-slice belongs at most to a time-measurable-series — because two time-measurable-series cannot overlap — a single T-slice cannot have attached to it more than one time.

As for the other problem we were talking about in the previous section — the fact that the past seems to change ‘too much’ — we individuated its source in the fact that Loss’s account treats T-slices as continuants which change in time in virtue of the existence of earlier or later counterparts. Such an exdurantist account applied to T-slices has, Loss argues, as an advantage the fact that it allows us to account for the platitude that the world changes, “by considering the world itself as a continuant that changes in time” (p.4). However, this implies the ‘too much’ change problem. An alternative, which is mentioned by Loss himself, is simply to say that the world changes in virtue of things which change in time.

4.4 The Past has not changed

We have seen in the previous section that at most one time can be intelligibly assigned to a T-slice, provided we don’t take the earlier-later relation to have a metric, which in turn would rule out the possibility for the past to change. This should incline us to favor the notion of a time-measurable-series over the notion of internal time used by Loss, because employing time-measurable-series necessarily implies that we can’t assign two times to the same T-slice. Loss’s notion of internal time comes from considering maximal proper parts of a possible world which exhibits the same kind of regularity (or quasi-regularity in cases of episodes of time travel) found in the actual world. On the contrary, a time-measurable-series requires also the presence of ways to keep track of time, i.e. the presence of one clock or more clocks suitably overlapping throughout all the series.

To conclude, I want to show how the notion of a time-measurable-series behaves with respect to Loss’s examples and his notion of internal time. Consider again the world W_1 , the one which has a 10 years spheres break. Internal time can account for broken maximal parts of a world. So can a time measurable-series, since its definition does not require continuity between the T-slices composing a maximal proper part. A time-measurable-series can also account for claims like the ones made by people on what Loss calls internal 1990, which occurs after the spheres break. People located there would say things like “The Berlin’s wall fall occurred one year ago”. According to time-measurable-series, the claim is correct, given that W_1 without the spheres break compose a time-measurable series. Moreover,

time-measurable-series do not entail a claim which I think it's problematic for Loss's account. Loss claim that the spheres break lasts *10 years*. I fail to see how the previous claim can be intelligible, unless the earlier-later relation comes with an intrinsic metric, which as we have seen is not a viable option for Loss's purposes. In fact, saying that something lasts 10 years means that during its occurrence The Earth rotated 10 times around the Sun. But the Earth and the Sun are not there during the spheres break. Of course, one might react by saying that the year-unit can be employed in measurements without using the Sun and the Earth as means. True. But in W_1 , *everything* is annihilated and replaced by spheres bouncing around, and hence we can't compare intervals in the spheres break with intervals outside it. In fact, suppose that during the spheres break, the sphere A moves from the point in space x to the point in space y , where x and y are separated in space by 100 kilometers. This sphere trip takes 10 years, Loss would say. Hence, sphere A moves at an average speed of 10 kilometers per year. However, how can you distinguish this case from a case where sphere A covers the same distance in 5 years at an average speed of 20 kilometers per year, or in one year at an average speed of 100 kilometers per year? Arguably you cannot. Hence, I fail to see what could ground the claim that the spheres break lasts ten years⁶. Time-measurable-series do not yield such a result. According to an account which employs them, the spheres break compose a time-measurable series, and the rest of W_1 compose another one. How long the spheres break lasts? It depends on whether there is some sphere which behaves cyclically, thereby determining a unit — let's name it XYZ — and as such can be taken as a clock. If so, the only intelligible answer would be that the break lasts a certain number of XYZs, and such interval cannot be compared to intervals which belongs to the rest of W_1 , given the *complete* irregularity at the beginning and the end of the spheres break.

What about Tim's case? Tim goes from what Loss calls 'external'²⁰¹⁰⁷ to what he calls an internal-1920. The problem was that there seems to be no way for Loss's account to say that it's Tim that time travels to the past, rather than 1919-continuants which travel to the future. With time-measurable-series we don't have such a problem. In fact, throughout the whole Tim's world there is *always* a way to keep track of time. During the most abrupt change, the alleged episode of Tim's time travel, Tim himself can be taken as a clock. In fact he regularly persists from the last 2010-slice

⁶It should be clear that this holds for whatever the number of spheres is. I talked just about one sphere for the sake of simplicity.

⁷If I persuaded the reader, then the specification 'external' is superfluous. It is 2010 period, given that only one time can be assigned to a T-slice.

to the first 2011-slice. Hence, the *whole* Tim's world is a time-measurable-series and therefore what Loss calls 'external 2011' is just 2011. That being said, the correct and unique answer is that 1919-continuants time travel to the future. Is it an odd result? Definitely yes, but just because the world Loss describes is an odd one. To conclude, Tim didn't travel to the past and the past is safe. Nobody, as we should have expected, changed it.

Chapter 5

Defending A Mutable Future

And now for something
completely different

Monty Python

In this final chapter we move from modal fatalism to a view, originally due to Peter Thomas Geach, according to which the future might literally change. I have already briefly introduced it in section 1.2.2. In chapters 1,2,3 I have been arguing in favor of what I have labeled reasonable fatalism. One of the consequences of reasonable fatalism is that future events are unpreventable, even though I suggested in section 1.1.2 to not cash out reasonable fatalism in terms of unpreventability, because doing so wrongly conveys the idea that there aren't causal connections among events, whereas reasonable fatalism accepts the Causal Connectedness thesis. On the contrary, according to the geachian view, agents *can* prevent future events from happening. As such, it is a view which clearly goes against fatalism. Hence the disjunctive nature of the overall thesis: either reasonable fatalism or there is the mutable future as an alternative.

In this chapter I present the view and provide an example which is supposed to score a point in its favor. My general aim is to highlight the fact that although the idea that the future might literally change received little attention thus far and surely comes with problems, it is a metaphysical view worth discussing.

5.1 Free Will and the Future

One natural way to gloss the notion of Free Will is as the power to do otherwise. Alice made a right at that corner the 10th of February 1967. Could she have turned left? We tend to answer affirmatively, as long as we think that her action was a free action. A billiard ball obeying Newtonian laws of nature seems to behave differently, for it never has that kind of power Alice seemed to have. Given an isolated system of billiard balls bouncing around in a table, the trajectory and the destiny of each particle is just one, namely the one dictated by a single instantaneous state of the system, boundary conditions and the laws of nature. Nobody would want to say that a billiard ball moving with a certain velocity at a certain time could be moving in another way, others things being equal. So, we typically tend to say that we human beings are different from material things like billiard balls; we say that we are free, while billiard balls are not. It is worth thinking what we do care about when free will is involved. It is likely that we as human being do not like the idea that we behave just like billiard balls when it comes to our deliberations. We don't want that our deliberations are fully settled by things like laws of nature and previous states of the physical system we are involved in. But there is at least another thing we care about; we want that our acts have an impact on the future. In short, we want that our free actions shape the future the way we desire.

We can see that by taking under consideration the following very bizarre possible world W_1 . At W_1 there is Bob and Bob is a free agent, in the sense that at any time, Bob has within his power to do otherwise than what he actually does. Yet, the world outside Bob is a very bizarre one. At T_1 , the time measured by Bob's wristwatch, the whole universe but Bob's body is completely empty. At T_2 , Bob is still there but the universe is now completely filled with chocolate peanut butter. At T_3 , the peanut butter has been replaced by a big donut with a small hole suitable for Bob's survival, and so on. In short, W_1 is such that Bob has the kind of regularity which is normally required for someone to be an agent, while the things outside Bob's body are highly irregular, in a way that it could even be hard if not impossible to say that there are laws of nature governing what happens in W_1 . We are also describing Bob as a free agent. Bob raised his right hand at T_1 , but he could have raised his left hand instead. Bob is free, yet Bob's free will is completely useless to him. The reason is that Bob cannot in any way affect his future surroundings, because W_1 is such that Bob's environment abruptly changes every second.

What a world like W_1 shows is the fact that some sorts of predictability

and regularities, in the world outside an agent, are required in order to give the agent the capability of affecting the future. I buy a flight ticket today because I want to bring about the event of me traveling to my holiday destination next summer. I can plan the trip and bring about the future I want as long as I can make reliable predictions based on patterns of regularities with flight companies, airports and so on.

In our actual world, a world full of regularities, we constantly bring about and cause future events. While we cannot affect the past, unless there are episodes of backward time travel, there is no doubt that we constantly affect how the future is, in the sense that we are capable of bringing about events. Can we also *change* the future? Many philosophers think that even though changing the future is an expression frequently used in everyday non-philosophical discourses, it is actually a philosophical non-starter, in the sense that after a quick analysis it turns out to be a meaningless notion. This is so for at least three reasons. One of the reasons is that change requires time. An object may change when it has certain properties at a time and different properties at another later time. For instance, I can change by sitting at T_1 and standing at T_2 . It looks as if times themselves do not change, while things can change in time. Another reason is that when we talk about changing the future, we have in mind situations like the following. On Monday, I decide to have coffee for breakfast on Wednesday, because I am having an important work meeting and I want to be wide awake by then. On Tuesday, I change my mind and I decide to have tea the next morning. Coffee makes me feel too anxious. It's better to have tea. So, one might want to conclude from an everyday perspective that from Monday to Tuesday the future has changed. The future was coffee-for-breakfast-on-Wednesday, and then it became tea-for-breakfast-on-Wednesday. The common philosophical reaction is that such a talk cannot be a correct analysis of what happened. If on Monday coffee-for-breakfast-on-Wednesday was part of the future, so it was my changing my mind on Tuesday. It would seem unjustified to claim that Wednesday was part of my future on Monday, whereas Tuesday wasn't. And then, this means that tea-for-breakfast-on-Wednesday *already* was the future on Monday, even though it didn't seem so. A further reason comes from considerations about the two standard approaches to time, the so-called A-theory and B-theory. According to the standard B-theory, the world is a static block which does not itself change. In such view, there isn't an absolute future and the ordering among events is given by an earlier-later relation among times. What counts as future is always a matter of perspective — what is future depends on where on the block someone finds himself. Typically, in a standard B-theoretic account, if something (objects,

events) is extended in time, then we might say it has temporal parts. Then, change is nothing more than qualitative difference between distinct temporal parts. What constitutes the content of a time, which might be future from someone's perspective, does not change. In a standard A-theoretical framework there is an absolute present and therefore an absolute future. What time is present is fundamentally part of what constitutes reality, and of course what time is present continuously change. What is now present was once future and will become past. Similarly, what is now future will change and become present, and then past. Yet, standardly, this is the only kind of change in the future allowed. What constitutes the content of a future time does not change.

As Todd (2011) interestingly pointed out, contrary to what is commonly believed, there is room for claiming and defending the idea that the future might literally change. Todd discusses the rather unknown position about the changing future, which is originally due to Thomas Peter Geach (1977). In the next section I want to briefly sketch Geach's view on the future.

5.2 Geach and the Changing Future

Geach argued for the idea that the future might sometimes change. He thought that it may happen that at some time the future is going to be in some way, whereas at a later time something else is going to happen at the same future time. As he says:

So what was going to happen at an earlier time may not be going to happen at a later time, because of some action taken in the interim. This is the way we can change the future. (p.50)

In other words, it might be the case that at T_1 the event X is going to happen at T_{10} , whereas at T_2 X is no longer going to happen at T_{10} . So, the future time T_{10} might change from containing X to containing something else. Geach provides the following basic example in order to illustrate his view. Suppose that at T_1 there's an accident on an airplane and the airplane starts to fall down. At T_1 , the airplane is going to crash at T_{10} . But then, at T_2 , the pilot intervenes with a brave maneuver and prevents the airplane from crashing. Because the crash has been prevented by the pilot, the future time T_{10} changed from an airplane crash to a safe landing. The rationale for claiming that in such a case the future has changed is what Geach calls the logic of prevention. When we have a case of prevention, what is prevented is what was going to happen and then did not happen. The pilot's act prevented the airplane from crashing at T_{10} . Before the pilot's

intervention, the airplane was going to crash at T_{10} . After the intervention, the future is such that it does not contain anymore the airplane crash at T_{10} .

In order to understand the boldness of the geachian view, it is worth mentioning his view about future contingent propositions, which behaves in accordance with the view that the future might change. He adopted a tensed view of propositions, namely a view where the tense operators are fundamental and irreducible to tenseless elements. Let us take the proposition expressed by the sentence

(55) The plane will crash at T_{10}

In such view, the proposition expressed comes with a future tense operator as fundamental. Yet, the future tense operator is not filled by a metric content like 'n units of time from now', but rather with a reference to a specific time, in this example T_{10} . (55) uttered at T_1 and, say, T_3 express the same exact proposition. Yet, in this view, the proposition is true at T_1 and false at T_3 . The same proposition changed its truth value, in accordance with the future time T_{10} which itself changed. Propositions of this sort are allowed to change truth value from truth to falsity and from falsity to truth. For instance, the proposition that the plane *will* safely land is first false and then true.

To sum up, according to Geach, our discourses about prevention show that what is prevented is what was going to happen and then did not happen. In order to make sense of prevention talks, Geach argues, a mutable future is required. The natural reaction to the geachian view is that the logic of prevention does not require a mutable future. One might argue that what is prevented when a preventive act takes place is not what was going to happen and yet did not happen, but rather something else. The most natural candidate as the object of prevention is what would have happened, had the preventive act not occurred. This is what might be called the counterfactual analysis of prevention.

Counterfactual Analysis of Prevention When an event X occurs and later in time Y does not occur, X prevents Y if and only if a) X is causally sufficient for not- Y and b) had X not occurred, Y would have occurred.

The counterfactual comes in the (b) clause and we adopt the standard truth conditions we have seen in section 1.1.1. Where A and B are propositions, $A \Box \rightarrow B$ — if A were true, B would also be true — is true at a possible world α if and only if either a) there isn't a possible world where A is true or b) there is a possible world β where A and B are true and such that it

is more similar to a than any other possible world where A is true and B false. In the plane crash case, the pilot's act at T_2 prevents the crash at T_{10} , because the pilot intervention ensures the non-occurrence of the crash at T_{10} *and*, had the pilot not intervened, the plane would have crashed at T_{10} . The advantage of the counterfactual analysis is that we don't need to posit a mutable future in order to account for what Geach calls the logic of prevention. What is prevented is not what was going to happen, but rather what would have happened, had the prevention not taken place. Thus, if the counterfactual analysis succeeds in explaining the logic of prevention, then it is preferable to the geachian view because it does not depart from the widely accepted idea that the future is immutable.

Todd (2011) raises two worries for this counterfactual analysis of prevention. The first one is what he calls the problem of the explosion of preventions. Take the event of the New York citizens going to sleep at a normal time the night of the 1st of March 2016. The following counterfactual is true: had the citizens not gone to sleep at a normal time, the 2nd of March there would have been a great amount of car accidents and fights in New York, because of the irritability and stress due to the lack of sleep. In addition, going to sleep at a normal time the 1st of March ensures that the day after there isn't an unusual amount of car accidents and fights. Thus, according to the counterfactual analysis of prevention, going to sleep at a normal time prevents fights and car accidents. This result is problematic because it can be multiplied ad nauseam with a lot of other ordinary cases. This is not the case with the geachian view. The unusual amount of fights and car accidents on March 2nd wasn't going to happen in the first place, and so in accordance with basic intuitions, nothing has been prevented when people went to sleep at a normal time.

The second problem that Todd points out comes from another example, originally suggested by Kenneth Boyce. Nuclear war is not going to happen because our adversaries have firmly decided that they will not launch their nuclear weapons (in such a way as to causally settle the matter), regardless of whether Obama chooses to sign the peace treaty or not. But Obama does sign the peace treaty, and his doing so is also causally sufficient for nuclear's war not happening. However, in all the nearest worlds in which Obama does not sign the peace treaty, it is because McCain won the election. And in all those worlds, McCain launches our nuclear weapons, thereby causing nuclear war. So we have it that Obama does something that is causally sufficient for nuclear's war failing to occur, and we have it that had Obama not signed the treaty, nuclear war would have occurred. The counterfactual analysis yields the result that Obama's signing is a case of prevention. Yet it seems false that Obama prevented

nuclear war, since our adversaries had already decided not to launch the nuclear weapons regardless of whether he signs the treaty. Once again, the geachian view does not suffer of this problem. The nuclear war was not going to happen in the first place, thus, by signing the peace treaty, Obama did not prevent the nuclear war.

Those two problems might not sound as killing objections against the counterfactual account of preventions, and surely there are ways to resist them. As for the first case, one might say that the going at bed at a normal time really prevented fights and car accidents the day after, it's just that we don't usually consider as relevant and significant those cases of prevention. As for the second case, it appears to be weak because it involves the problematic claim that the nearest possible world where Obama does not sign the peace treaty is a world where Obama is not President. In any case, I don't want to discuss those cases here. Rather, what I will do is considering a new example which sounds problematic for the counterfactual analysis of prevention. To this example I now turn.

5.3 A case in favor of the Changing Future

The example involves two free agents, Alice and Bob, who are playing pool in a bar. We assume that the pool balls behave in perfect accordance with Newton laws of motion. If there is no intervention from the outside, the future behavior of the pool balls is completely settled by the present physical state of the table and the laws of nature. Alice has direct control over the pool ball number 1, whereas Bob has direct control over the pool ball number 2. That means that they can freely use their sticks to impart a velocity to their respective pool balls. Alice and Bob shots are not dictated by Newtonian laws of motion. When they impart a certain velocity to their ball, they do so freely. We consider the game from T_0 to T_{10} and we then make the following assumptions. The game is not an usual pool game, because Alice and Bob are allowed to hit the balls simultaneously or, when it is not the case that all pool balls are at rest. In addition, we assume that at all times of the the interval considered, it is not the case that all pool balls are at rest. Ultimately, Alice and Bob know newtonian laws of nature and they can predict what is going to happen given an instantaneous state of the table and the velocities of the balls. Given their knowledge and their ability to shoot pool balls, Alice and Bob can affect the future the way they want, by imparting a proper velocity to balls number 1 and 2. For reasons of their own, Alice wants the ball number 8 to end up in the upper-left pocket at T_{10} , whereas Bob wants the ball number 8 to end up in the upper-

right pocket at T_{10} , and both Alice and Bob want to intervene as little as possible. Then, the game goes like this. At T_0 , given the present state of the table and newtonian laws of nature, the ball number 8 is going to end up in the upper-right corner at T_{10} . Alice wants to prevent that, and after a quick calculation, she imparts a velocity to ball number 1 at T_1 , so that it will hit the ball number 8 with a velocity such that ball number 8 will end up in the upper-left pocket at T_{10} . After realizing what is going on, Bob is determined to achieve what he wants. Hence, at T_7 , Bob shoots balls number 2 which in turn hits ball number 8 and send it in the upper-right pocket. What actually happens is that ball number 8 eventually ends up in the upper-right pocket at T_{10} .

For the sake of simplicity, we name the events as following: 'A' is Alice shot at T_1 , 'B' is Bob shot at T_7 , 'R' is ball number 8 ending up in the upper-right pocket at T_{10} and 'L' is ball number 8 ending up in the upper-left pocket at T_{10} .

While it is debatable whether A prevents R because R eventually happen, for sure B prevents L from happening. We shall see now how the counterfactual analysis of prevention deals with this case and the events involved. The event A does not ensure the non-occurrence of R because R does eventually happen. This suffices to claim, perhaps correctly, that A does not prevent R. The second case is more interesting, because B does prevent L from happening and the conterfactual analysis makes a wrong prediction. B ensures the non-occurrence of L and hence the first clause of the counterfactual analysis is satisfied. However, the relevant counterfactual is false. It is not true that had Bob not intervened, L would have happened. The counterfactual $\neg B \square \rightarrow L$ is deemed false by standard counterfactuals truth conditions in such scenario. It is false because the worlds similar to Alice and Bob's world where Bob does not intervene are also worlds where Alice does not intervene, given that both want to intervene as little as possible and Bob intervenes because Alice did it as well. But the worlds where they do both not intervene are worlds where R instead of L happen at T_{10} , given the Newtonian dynamics governing the balls' behavior and the state of the system at T_0 . Hence, a world where $\neg B$ and $\neg L$ is closer to the one Alice and Bob inhabit than any world where $\neg B$ and L. The fact that the relevant counterfactual turns out to be false implies that the counterfactual analysis of prevention does not count Bob' shot as a case of prevention and this result is clearly wrong.

The geachian view does not suffer this kind of problem. One who endorses the geachian view would say that at T_6 , before Bob' shot, L was going to happen, but then, Bob prevented that, thereby changing the future time T_{10} from containing L to containing R. Thus, if I am right, this case

scores a point in favor of the geachian view on prevention, the mutable future and against the counterfactual analysis.

5.4 Conclusion

To conclude, I have been arguing in favor of a version of modal fatalism. Fatalism scared generations of philosophers and has often been seen as an untenable doctrine. However, I tried to show that a version of it can be *reasonable*. Reasonable fatalism does not rule out causation, does not have total laziness as a consequence and can account for our moral judgments with respect to punishments and rewards. In chapter 2 I gave an argument in favor of modal fatalism, which together with the Causal Connectedness thesis constitutes reasonable fatalism. In chapter 3 I tried to show that reasonable fatalism can provide a correct account with respect to what agents can or cannot do when it comes to episodes of backwards time travel toward, and how time traveler can affect events in the past.

The unchangeability of the past is an assumption that fatalism shares with several other metaphysical theories. In chapter 4 I argued against the successfulness of a recent attempt to show that the past might change.

According to fatalism, we don't have the power to change the future. However, as we have seen, the fact that we might literally change the future is the main claim of geachianism. In this chapter I provided an argument in favor of this view. The geachian view surely comes with problems — for instance, it seems that it has to relinquish the factivity of knowledge — and it is in need of further developments — for instance, what kind of ontology of time does it need? Yet, it appears to be promising. Geachianism is a view that has as a consequence the falsity of fatalism, cause it claims that agents have the power to change the future. Hence, fatalism and geachianism are incompatible. Moreover, both theories are quite unpopular. The received view is that the future *cannot* change and yet fatalism is false. I argued for the denial of the previous conjunction. Hence the disjunctive nature of my thesis: either we should accept (reasonable) fatalism and deny the *power* to do otherwise, or we should think that the future might change.

References

- Abbruzzese, J. 2001. "On Using the Multiverse to Avoid the Paradoxes of Time Travel". *Analysis* 61(1):3638.
- Abelson, R. 1963. "Taylor's Fatal Fallacy", *The Philosophical Review*. 72: 93-96.
- Albert, D. 1992. *Quantum Mechanics and Experience*. Harvard University Press.
- Albert, D. 2000. *Time and Chance*. Harvard University Press.
- Ayer, A.J. 1963. "Fatalism", in *The Concept of a Person and other Essays*. New York: St Martin's.
- Barnes E. and R. Cameron. 2009. "The Open Future: Bivalence, Determinism and Ontology", *Philosophical Studies* 146: 291-309.
- Barnes E. and R. Cameron. 2011. "Back to the Open Future", *Philosophical Perspectives* 25(1): 1-26.
- Belnap N. and M. Green. 1994. "Indeterminism and the Thin Red Line", *Philosophical Perspectives* 8: 365-88.
- Belnap N, and M.Perloff and M. Xu. 2001. *Facing the Future: Agents and Choices in Our Indeterminist World*. Oxford: Oxford University Press.
- Benovsky J. 2011. "The Relationist and Substantivalist Theories of Time: Foes or Friends?". *European Journal of Philosophy* 19(4):491-06.
- Bernstein M. 2002. "Fatalism". In R. H. Kane (ed.) *The Oxford Handbook of Free Will*. Oxford: Oxford University Press.

- Borghini A. and G. Torrenco. 2013. "The Metaphysics of the Thin Red Line." In F. Correia and A. Iacona (eds.), *Around the Tree. Semantical and Metaphysical Issues Concerning Branching and the Open Future*. Kluwer Academic Publishers. 105-25.
- Bourne, C. 2011. "Fatalism and the Future" In C. Callender (ed.), *The Oxford Handbook of Philosophy of Time*. Oxford University Press. 41-67.
- Brown, C.D. 1965. "Fallacies in Taylor's Fatalism". *Journal of Philosophy* 62: 349-53.
- Cahn, S. 1967. *Fate, Logic and Time*. New Haven: Yale University Press.
- Cahn, S. 1967. "Fatalistic Arguments". *Journal of Philosophy*. 61(10): 295-305.
- Casati R and A. Varzi. 2001. "That Useless Time Machine". *Philosophy* 76(4):581-83.
- Chellas, B. 1980. *Modal Logic, an Introduction*. Cambridge University Press.
- Dennett, D. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. MIT Press.
- Diekemper, J. 2004. "Temporal Necessity and Logical Fatalism". *Proceedings of the Aristotelian Society* 104(3):287-94 .
- Diekemper, J. 2007. "B-Theory, Fixity, and Fatalism". *Nous* 41 (3):429-52. *Proceedings of the Aristotelian Society* 104 (3):287-94.
- Dummett, M. 1964. "Bringing About the Past". *the Philosophical Review*. 73:338-59.
- Dowe, P. 2000. "The Case for Time Travel". *Philosophy* 75 (3):441-51.
- Dowe, P. 2003. "The Coincidences of Time Travel". *Philosophy of Science* 70 (3):574-58.
- Dowe, P. 2007. "Constraints on Data in Worlds with Closed Timelike Curves". *Philosophy of Science*. 74 (5):724-35.
- Everett A. and Roman T. 2012. *Time Travel and Warp Drives*. Chicago: Chicago University Press.

- Fine, K. 1994. "Essence and Modality". *Philosophical Perspectives* 8:1-16.
- Fischer J.M. and Ravizza M. 1998. *Responsibility and Control: a Theory of Moral Responsibility*. Cambridge University Press.
- Fischer J.M and Todd P. 2011. "The Truth about Freedom: a reply to Merricks". *The Philosophical Review* 120(1): 97-115
- Fischer J.M and Todd P. (eds.). 2015. *Freedom, Fatalism and Foreknowledge*. Oxford: Oxford University Press.
- Frankfurt, H. 1969. "Alternate Possibilities and Moral Responsibility". *Journal of Philosophy* 66(23):829-39.
- Geach, P. 1977. *Providence and Evil*. Cambridge University Press.
- Haack, S. 1974. "On a Theological Argument for Fatalism". *Philosophical Quarterly* 24(95):156-59.
- Harrison, J. 1971. "Dr. Who and the Philosophers or Time-Travel for Beginners". *Aristotelian Society Supplementary Volume* 45 (1):1-24.
- Hawtorne, J., T. Sider and W. Zimmerman (eds.) 2008. *Contemporary Debates in Metaphysics*. Blackwell.
- Horwich, P. 1987. *Asymmetries in Time: Problems in the Philosophy of Science*. Cambridge: MIT Press.
- Iacona, A. 2007. "Future Contingents and Aristotle's Fantasy". *Critica* 39 (117):45-60.
- Iacona, A. 2014. "Ockhamism Without Thin Red Lines". *Synthese* 191 (12):2633-52.
- Kripke, S. 1980. *Naming And Necessity*. MA: Harvard University Press.
- Lewis. D.K. 1968. "Counterpart Theory and Quantified Modal Logic." *Journal of Philosophy* 65(5):113-26.
- Lewis, D.K. 1973. "Causation". *Journal of Philosophy* 70(17):556-67.
- Lewis, D.K. 1976. "The Paradoxes of Time Travel". *American Philosophical Quarterly* 13(2):145-52.
- Lewis, D.K. 1981. "Are We Free to Break the Laws?" *Theoria* 47(3):113-21.

- Lewis, D.K. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
- Lockwood, M. 2005. *The Labyrinth of Time: Introducing the Universe*. Oxford University Press.
- Lukasiewicz, J. 1967. "On Three-valued Logic". in S. McCall, ed. *Polish Logic*. Oxford: Oxford University Press.
- Lukasiewicz, J. 1968. "On Determinism". *The Polish Review*. 13(3):47-61
- Loss, R. 2015. "How to Change the Past in One-Dimensional Time". *Pacific Philosophical Quarterly* 96(1):1-11.
- MacFarlane, J. 2003. "Future Contingents and Relative Truth". *Philosophical Quarterly*. 53(212):321-36.
- Markosian, N. 1995. "The Open Past". *Philosophical Studies* 79(1):95-105.
- Markosian, N. 2013. "The Truth About the Past and the Future". In F. Correia, A. Iacona, eds. *Around the Tree*. Berlin: Synthese Library, Springer Verlag.
- Mackie, P. 2003. "Fatalism, Incompatibilism, and the Power to Do Otherwise." *Nous* 37(4):672-89.
- Merricks, T. 2009. "Truth and Freedom". *Philosophical Review* 118 (1):29-57.
- Meyer, U. (forthcoming). "Fatalism as a Metaphysical Thesis". *Revista Internacional de Filosofia*.
- Øhrstrøm, P. 2009. "In Defense of the Thin Red Line". *Humana.mente* 8 Models of Time: 17-32.
- Pereboom, D. 2001. *Living Without Free Will*. Cambridge University Press.
- Pooley, O. 2013. "Relativity, the Open Future, and the Passage of Time." *Proceedings of the Aristotelian Society* 113:321-63.
- Prior, A. 1953. "Three-Valued Logic and Future Contingents". *The Philosophical Quarterly* 3(13): 317-26.
- Prior, A. 1967. *Past, Present, and Future*. Oxford: Oxford University Press.
- Prior, A. 1976. "It Was to Be". In P. Geach and A.J.P. Kenny (eds.) *Papers in Logic and Ethics*. London: Duckworth.

- Putnam, H. 1962. "It Ain't Necessarily So". *Journal of Philosophy* 59 (22):658-71.
- Rea, M. 2006. "Presentism and Fatalism". *Australasian Journal of Philosophy* 84(4):511-524.
- Rea, M. 2015. "Time Travelers Are Not Free". *Journal of Philosophy* 112(5):266-79.
- Rietdijk, C. 1966. "A Rigorous Proof of Determinism Derived From the Special Theory of Relativity". *Philosophy of Science* 33(4):341-44.
- Ryerson, J. 2011. "A Head that Throbbled Heartlike: the Philosophical Mind of David Foster Wallace" in Wallace D.F. *Fate, Time and Language: an Essay on Free Will*. Columbia University Press. 1-34.
- Rosenkranz, S. 2012. "In Defence of Ockhamism". *Philosophia* 40(3):617-31.
- Ryle, G. 1954. "It Was to Be" in *Dilemmas*. Cambridge University Press.
- Saunders, J. T. 1962. "Professor Taylor on Fatalism". *Analysis* 23(1):1-2.
- Saunders, J.T. 1965. "Fatalism and Ordinary Language". *Journal of Philosophy* 62(8):211-22.
- Sharvy, R. 1964. "Tautology and Fatalism". *Journal of Philosophy* 61 (10):293-295
- Shoemaker, S. 1969 "Time Without Change". *Journal of Philosophy*. 66: 363-81.
- Sider, T. 2002. "Time Travel, Coincidences and Counterfactuals". *Philosophical Studies* 110: 115-38.
- Slater, M. 2005. "The Necessity of Time Travel (on Pain of Indeterminacy)". *The Monist* 88(3):362-69.
- Smith, N.J.J. 1997. "Bananas Enough for Time Travel?". *British Journal for the Philosophy of Science* 48(3):363-89.
- Smith, N.J.J. 2005. "Why Would Time Travelers Try to Kill Their Younger Selves?". *The Monist* 88(3):388-95.
- Taylor, R. 1962. "Fatalism". *Philosophical Review* 71(1):56-66.

- Taylor, R. and Makepeace P. 1962. "Fatalism and Ability". *Analysis* 23(2):25-29.
- Taylor, R. 1963a. *Metaphysics*. Englewood Cliffs, N.J.: Prentice Hall.
- Taylor, R. 1963b. "A Note on Fatalism". *Philosophical Review* 72(4):497-99.
- Todd, P. 2011. "Geachianism". *Oxford Studies in Philosophy of Religion*. 3:222-51.
- Todd, P. 2016. "On Behalf of a Mutable Future". *Synthese* 193(7):2077-95.
- Torre, S. 2011. "The Open Future". *Philosophy Compass* 6:360-73.
- Van Inwagen, P. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.
- Van Inwagen, P. 2009. "Changing the Past", in D. Zimmerman (ed.) *Oxford Studies in Metaphysics*, vol. 5. Oxford: Oxford University Press, 3-40.
- Van Rensselaer Wilson, H. 1955. "Causal Discontinuity in Fatalism and Indeterminism". *The Journal of Philosophy* 52(3):70-72.
- Vihvelin, K. 1996. "What Time Travelers cannot do". *Philosophical Studies* 81, 315-30.
- Vranas, P. 2005. "Do Cry Over Spilt Milk: Possibly You Can Change the Past". *The Monist*. 88(3):370-87.
- Wallace, D.F. 2011. *Fate, Time and Language: an Essay on Free Will*. New York: Columbia University Press.
- Wasserman, R. 2017. "Time Travel, Ability, and Arguments by Analogy". *Thought*. 5(4)

no proprietary software has been used to write this thesis.