Logica Universalis

CrossMark

Karma Theory, Determinism, Fatalism and Freedom of Will

Ricardo Sousa Silvestre

Abstract. The so-called *theory of karma* is one of the distinguishing aspects of Hinduism and other non-Hindu south-Asian traditions. At the same time that the theory can be seen as closely connected with the freedom of will and action that we humans supposedly have, it has many times been said to be determinist and fatalist. The purpose of this paper is to analyze in some deepness the relations that are between the theory of karma on one side and determinism, fatalism and free-will on the other side. In order to do that, I shall use what has been described as the best formal approach we have to indeterminism: branching time theory. More specifically, I shall introduce a branching time semantic framework in which, among other things, statements such as "state of affairs e is a karmic effect of agent a", "a wills it to be the case that e" and "e is inevitable" could be properly represented.

Mathematics Subject Classification. Primary 03A05; Secondary 03B44, 03B45, 03B80, 03C55, 03C98.

Keywords. Branching time theory, temporal logic, logic of action, karma theory, determinism, free-will, fatalism.

1. Introduction

The so-called *theory of karma*, or law of karma, is one of the distinguishing aspects of Hinduism and other non-Hindu south-Asian traditions. There have been many descriptions of the theory of karma, from the usually obscure passages of the *Upanishads* and accounts made by several figures of *Mahabharata* and *puranas*, to the sometimes overconfident descriptions of modern Indian gurus. From the western scholar point of view, we have witnessed an increase in interest in the theory of karma during the last decades (see, e.g.,

Birkhäuser

[8, 10, 17, 25, 26, 28] and [30]); the recent debate on the theory of karma as a solution to the problem of evil (see [7, 19, 20] and [34]) is a good instance of this. However, and despite of this, there is an important issue which remains understudied: the connection between the theory of karma on one side and determinism, fatalism and free-will on the other side.

The theory of karma[1] has many times been said to be determinist and fatalist: "if we are justified in our acceptance of the causal dogma [of the theory of karma], there does not seem to be any legitimate way to avoid fatalism. If the present is determined by the past so as to admit of an accurate prediction of the past [...], how can we avoid the conclusion that the future is similarly determined by the past and the present?" ([35], p. 82). In fact, according to the theory of karma, our present actions somehow determine our future states of being, including our disposition to act and desire in specific ways. So, if the theory of karma is true, it seems we are faced with something very close to determinism and fatalism (at least in a localized way) as well as with an absence of free-will.

On the other hand, there is a sense in which the law of karma is closely connected with the freedom of will and action that we humans supposedly have. Radhakrishnan, for example, says as follows: "Man is not a mere product of nature. He is mightier than his karma. [...] The law of karma, which rules the lower nature of man, has nothing to do with the spiritual in him [...] The essence of spirit is freedom. By its exercise man can check and control his natural impulses." ([29], p. 246). In fact, the whole purpose of the theory of karma is to function as a moral retributive law regarding the actions we humans do. But if there is no freedom of will, and the future, including our future actions, is completely determined by our past actions, in which sense can these actions be said to be *ours*?

My purpose in this paper is to analyze in some deepness the relations that are between the theory of karma on one hand and determinism, fatalism and free-will on the other hand. In order to do that, I shall use what has been described as the best formal approach we have to indeterminism: *branching time theory* ([1, 3–6]). Besides being an interesting framework for saying what determinism, indeterminism and fatalism are, branching time theory has been quite successfully used to formalize several kinds of agentive sentences, notably sentences describing action-like events (see, e.g., [2, 9, 16, 33] and [36]). This is important because a minimally detailed description of the theory of karma should take into account the actions agents do and the karmic effects coming from them. Therefore, one of my primary tasks will be to develop a semantic framework in which statements such as "state of affairs e is a karmic effect of agent a", as well as "a wills it to be the case that e" and "e is inevitable", might be represented; following the standard in philosophical logic, this will be done

---

[1] As it shall be elaborated in the next Section, the theory of karma can be described as follows: besides its purely physical effects, our actions have also moral and psychological effects, which might occur either in this life or in future ones, and which affect our environment, genetic make-up, physical characteristics, social status at birth, length of life, etc., as well as our psychological dispositions and tendencies to act, desire, etc.

with the help of modal operators defined inside this semantic framework. With the operators and semantics at hand I will be able to define in a minimally precise way the notions of determinism, fatalism and (lack of) free-will so that their relations with the theory of karma might become clearer.

The structure of the paper is as follows. In the next Section I lay down the basic principles of karma theory, which together form the subject matter of my analysis. In Sect. 3 I informally introduce the modal operators I shall use in the rest of the paper and present a preliminary but important analysis of some of those principles. In Sects. 4 and 5 I introduce the semantic framework which will be used in the next Section to conclude the analysis I have started in Sect. 3. Finally in Sect. 7 I lay down some concluding remarks.

2. General Principles of the Theory of Karma

Before starting my analysis, it is useful to make a couple of considerations about the theory or law of karma. First of all, there is a problem, someone might claim, with the use of the words "theory" and "law" here: it is not clear at all in which sense the theory of karma is a theory or law. Although no defendable synonymous for these words appear in the Sanskrit texts which inspired the doctrine, these terms have become standard among modern Indian and western writers. A more cautious and realistic position would be to take these words in a broader sense, meaning the same as "doctrine", "concept", "idea", etc. It is cautious because in this way we do not have to explain the relation of the doctrine with modern and contemporary uses of the words "theory" and "law"; it is realistic because it takes into account the variety of different understandings that exist about the theory of karma in Indian tradition. This takes us to the next consideration.

From a strict point of view, we cannot talk about *the* theory of karma. Indian philosophical and religious tradition has produced many different views of what the doctrine is. This point can be illustrated in several ways. Rocher, for instance, has found five systems of karma and rebirth in one single chapter of one Indian scripture: the twelfth book of the *Manusmrti* [32]. Aspects such as the possibility of karma transfer, divine intervention in the karmic process, nullification of karma, the way karma is fructified, the role of *samskaras* in the karmic process, just to mention the most prominent ones, are interpreted quite differently in different traditions. We have Buddhist views on karma, which are considerably different from Jain ones, which differ from Advaita Vedanta ones, which are different from the view of theist versions of Vedanta, and so on and so forth (see [10, 25, 26] and [30]).

This might lead one to argue that it is simply misleading to speak of something like *the* theory of karma. One cannot speak of the theory of karma outside a specific Indian religious tradition. You have to pick up one tradition or a specific set of texts and see what it says about the doctrine; this will give you the only thing you can have: *a* theory of karma. This is one of the criticisms Chadha and Trakakis [7] make about Kaufman's analysis [19] of the role of the theory of karma in the problem of evil. And it is a quite reasonable

criticism. However, that one cannot speak about *the* theory of karma does not imply that there are not principles and presuppositions common to all different theories of karma. The doctrine of rebirth is surely part of most, if not all, post-Vedic theories of karma; the same can be said about the causal aspect of the law of karma (see, e.g., [11,18,21,23] and [27]). That these principles do not form a theory of karma does not mean that their study cannot shed light on the existing versions of the theory of karma. On the contrary: by being, we suppose, the common core of all theories of karma, their analysis might tell a lot about these theories. That is the purpose of the rest of this section: to try to disclose some key presuppositions or principles common to all theories of karma.[2]

To start with, I will say that the basic idea of the theory of karma is this: besides its purely physical effects, our actions have also moral and psychological effects, which might occur either in this life or in future ones, and which affect our environment, genetic make-up, physical characteristics, social status at birth, length of life, etc., as well as our psychological dispositions and tendencies to act, desire, etc.

Here we have the popular idea related to karma theory that while morally good actions cause proportionally good states of affairs, morally bad actions cause proportionally bad states of affairs ([26], p. xi). These effects or consequences are also connected with the agent's physique: in order for an action to be classified as good or bad, for instance, the agent's intentions have to be taken into account.[3] This connection however is two-sided. Besides including our environment, genetic make-up, physical characteristics, social status at birth, length of life, etc. (that is, virtually everything that matters in one's life), the effects of one's actions also include one's psychological dispositions and tendencies to act, desire, etc., corresponding to the famous notion of *samskara* (see [30], pp. 19–21).

From this brief description it is also clear that the theory of karma is a causal theory (see [30], pp. 24–43); after all, it prescribes, even if in a vague and general way, how certain effects follow from certain agentive causes. Adopting the view that if a full description of a causal relation is possible at all it should be done in terms of sets of sufficient conditions (see [22]; Mill [24] seems to also endorse this view), I can state as follows: If, according to the theory of karma, (1) if such and such (kind of) action c is performed and certain antecedent conditions are met, then such and such (kind of) effect e follows and (2) action c (or an instance of the action-kind c) is performed and the

---

[2] I shall from now on use the expressions "theory of karma" and "law of karma" to refer either to this set of basic presuppositions or to some idealistic or popular view of the doctrine which might help us in the task of finding out about these basic principles.

[3] Some traditions go as far as saying that the condition of one's next life is fully determined by his or her psychological condition at the moment of death. E.g., verse 6 of chapter 8 of the *Bhagavad-gita*: "Moreover, whatever state of being he remembers when he gives up the body at the end, he goes respectively to that state of being, Arjuna, transformed into that state of being."

corresponding conditions are met, then (3) e (or an instance of the effect-kind e) occurs.

Based on this, I define the notion of *karmic effect* as follows: a karmic effect of an agent a is an effect of one or more actions already performed by a whose antecedent conditions were met. With this notion at hand, I lay down what I see as the first basic principle of karma theory. Let s be a state of affairs and a an agent.

Cause–Effect Principle If s is a karmic effect of a, then s is (or will be) the case.

This principle guarantees that the theory of karma really works, that is to say, that a karmic effect of an agent indeed occurs (or will occur).

Besides this causal side, the theory of karma has also, as we know, a moral side: a karmic effect s of an agent a should be morally proportional to the actions a did which caused s. This is the idea behind the popular saying that one reaps what one sows and it is why the law of karma is many times saw as a retributive law, or still, as the implementation of some principle of cosmic justice ([11], p. 2). As such, it seems obvious that a's karmic effects must somehow affect a: after all, they have to do with a's environment, a's genetic make-up, a's physical characteristics, a's social status at birth, a's length of life, as well as with a's psychological dispositions and tendencies to act. Therefore the following principle:

Person-Affecting Principle If s is a karmic effect of a, then s affects a.

Along with the Cause–Effect Principle, which is a predictive principle, the Person-Affecting Principle allows one to speak about the way future will be with respect to an agent. The most appealing side of the theory of karma, however, has been its supposedly explanatory power ([26], p. xi). Defenders of the theory have promoted it by saying that only by making reference to theory of karma can we explain otherwise unexplainable events regarding one's condition of life.[4] From a general viewpoint, the idea is that all that matters in one's life must be explained as the karmic result of his or her previous actions. We thus have the following principle:

Explanation Principle If an actual state of affairs s affects a, then s is a karmic effect of a.

Here is another principle:

Change Principle The pool of karmic effects of an agent can change.

This principle is a very important tenet of karma theory. Since the actualization of a karmic effect might be temporally very distant from the actions

---

[4] Why, for instance, one child is born in a very comfortable situation, in a wealthy family, good looking, healthy, etc., and other is born in a very different environment: unhealthy and in a poor and problematic family? The answer is that while the first one had a virtuous previous life, the second one must have had a sinful previous life, or at least committed a couple of immoral acts; they are just reaping what they sowed. From a theist perspective, this would free God from the threat of being unjust.

which caused it (in fact, they might take place in an agent's future life), some traditions have come up with the notion of a *karmic pool*, a pool of created but not 'enjoyed' karmic effects ([30], pp. 19–22). The idea of this principle is that this pool of karmic effects can (and usually do) change. This has to do with the obvious fact that since a is performing new actions at every moment, he or she is creating new karmic effects, which will be added to the old ones. Second, when a karmic effect becomes the case, it is not part any more of the karmic pool of the agent. Finally, some versions of karma theory admit the possibility of a's nullifying or transferring a karmic effect which has already been created but has not yet become the case ([27], pp. 10–13).

But for this principle to work, it is necessary that there might be possible ways the world could be which are incompatible with a's karmic pool at a given moment of time. Let t be a specific moment of time. If all ontological possibilities the world could be are compatible with a's karmic pool at t, then there would be no room for a's karmic pool to change at a moment posterior to t. By a's karmic pool at t being incompatible with a way the world could be I simply mean that at least one of the karmic effects of a's karmic pool at t is not the case in this way things could be. Therefore we have a further principle:

Incompatibility Principle There are ways the world could be which are incompatible with the karmic effects of a at t.

3. Freedom of Will, Determinism and Fatalism: A Preliminary Approach

Now, how do these principles work together? How are they related to the issues of free-will and determinism? And how might they be represented inside a formal framework? I will in this section make some preliminary but important steps towards answering these questions. I will work on a very elementary logical level, leaving the full formal work to the next sections. I will also postpone the analysis of the two last conditions; as it shall be clear in the course of the text, the analysis of the Change Principle and the Incompatibility Principle will have to wait until Sect. 6.

Let us consider the Cause–Effect Principle first: If s is a karmic effect of a, then s is (or will be) the case. First of all, we need a way to represent that s is a karmic effect of a. I shall do that with the help of a modal operator, KE. Let a be an agent and α a sentence, formula or proposition.[5] Formula [a]KE:α means that the state of affairs represented by α is a karmic effect of a.[6] In this way, the Cause–Effect Principle would be represented as follows:

---

[5] I shall here overlook the technical differences that are between the terms "sentence", "proposition" and "formula".

[6] Most of the time I shall drop reference to the fact that what is a karmic effect of a is not α itself, but the state of affairs represented by α, and simply speak of α's being the karmic effect of a. A similar movement will be done with the other formal sentences to be introduced later.

(1) [a]KE:α → α[7]

The situation where α is the case in a future moment could be represented with the help of the tense operator F: if α is a formula, Fα means that α will be the case. Thus:

(2) [a]KE:Fα → F α

But there is a problem here. Due to the questions I am addressing, there should be room in this simple language for an indeterminist view of the future. According to such a view, writing Fα does not mean that α will be definitively true, but simply that in one way the world could be in the future, α is true. See however that this understanding of Fα conflicts with what we want (2) really to say. The idea of this tense version of the Cause–Effect Principle is not that if Fα is a karmic effect of a then α is true in a possible way the future could be: we want Fα to be true independently of which of the possible future courses of the history becomes the case. In other words, we want the truth of Fα to be inevitable. Letting I be a modal operator such that I:α means that α is inevitable, (2) should be rewritten then as follows:

(3) [a]KE:Fα → I:(F α)

A similar reasoning can be made to generalize this idea and show that the correct way to represent the Cause–Effect Principle is as follows:

(CEP) [a]KE:α → I: α

As far as the Person-Affecting Principle—if s is a karmic effect of a, then s affects a—is concerned, I will use a very simple approach to represent the notion of affectability: a state of affairs e affects agent a if and only if a wills it to be the case that e or a wills it to be the case that non-e. This is of course a psychological approach: insofar as a has some attachment or aversion towards e, e has some impact on a's mental states and attitudes; therefore e affects a.

But if I am to represent the notion of affectability in terms of a's will, I need a way to represent a's willing a certain state of affairs to be the case. Let a be an agent and α a formula. [a]Will:α means that a wills it to be the case that α. The notion of affectability will then be introduced as follows:

(DEF$_w$) [a]Affect:α =$_{def}$ ([a]Will:α)∨([a]Will:¬α)

[a]Affect:α means that a is affected by the state of affairs represented by α. Having said this, here is the Person-Affecting Principle:

(PAP) [a]KE:α →[a]Affect:α

The representation of the Explanation Principle is straightforward:

(EP) α ∧[a]Affect:α →[a]KE:α

Addressing now the second question I have posed in the beginning of this section, the threat to free-will generally associated with the theory of karma might be represented as follows:

---

[7] The symbols →, ¬, ∧ and ∨ are been used here with their usual meaning.

(FW) [a]Will:α →I:([a]Will:α)

That is to say, if the fact that a wills it to be the case that α implies that this willing is inevitable, that is to say, that it had to happen, then there is no freedom of will. I am assuming here an incompatibilist view of free-will: for an agent a to freely will it to be the case that α requires that there were a plurality of futures open to a so that he or she could have willed otherwise (see [14]).

From (EP) and (CEP), through the transitivity of material implication, we obtain what I name the *Karmic Inevitability Principle*:

(KIP) α ∧[a]Affect:α →I:α,

of which the following formula is a trivial instance:

(4) [a]Will:α ∧[a]Affect:([a]Will:α) →I:([a]Will:α).

(4) basically says that if a wills it to be the case that α and is affected by his or her willing, then his or her willing is inevitable. Or, in a non-abbreviated form:

(4′) [a]Will:α ∧([a]Will:([a]Will:α)∨[a]Will:¬([a]Will:α)) →I:([a]Will:α)

We could easily get (FW) from (4) if the following principle were true:

(5) [a]Will:α →[a]Will:([a]Will:α)∨[a]Will:¬([a]Will:α)

But clearly it is not. There are many instances of willing in which I will that α but neither will that I will that α nor will that I do not will that α: I might will this paper to be finished soon without willing that I will this paper to be finished soon, nor willing that I do not will this paper to be finished soon.

However, there are many cases of willing in which either [a]Will:([a]Will:α) or [a]Will:¬([a]Will:α) is true. For instance, I might be addicted to alcohol and will to drink as much as possible. However, being aware of this addiction and willing to stop it, both propositions might quite reasonably be true:

(6) [I]Will:(I_drink_as_much_as_possibe)

(7) [I]Will:¬([I]Will:(I_drink_as_much_as_possibe))

The same can be said about my willing to help others. I know that it is a good thing and I want to keep doing it. Therefore, the two propositions below might be true as well:

(8) [I]Will:(I_help_other_people)

(9) [I]Will:([I]Will:(I_help_other_people))

These examples seem to be a sort of second-order willing where besides willing α the agent also wills to will that α or wills not to will that α. I could perhaps say that conscious moral willing have this feature: if I am aware of the moral aspect of my willing, I shall have one of those two attitudes towards it, depending on whether I see its content as good or bad. And assuming that a moral agent must be aware of the moral quality of his acts and wills, that is to say, that moral willing is always conscious moral willing, I could further say that if α has a moral content and a is moral agent, then (5) is true. As a

consequence of that, (FW) would also be true; in other words, for this kind of second order willing in general and moral willing in particular there would be no freedom of will.

This is certainly odd, for while there is no threat to freedom of first-order will, second-order will would be inevitable. Perhaps we could try to make sense of this through the notion of *samskara*. Usually, *samskaras* are understood as tendencies or dispositions to act, think, desire, etc. caused by our previous actions ([30], pp. 19–21). The performance of good and bad deeds produce a tendency to act and will accordingly. If it is correct that moral willing is conscious moral willing, that is, that (5) is true in the case α has a moral content and *a* is moral agent, then the fact that (FW) is also true for α and *a* can be seen as a representation of the impact of karma upon the will of moral agents.

Unfortunately this attempt to soft down the problem does not work. Suppose that α means "this paper is finished soon" and *a* means myself. As I have stated above, it is true that [a]Will:α but false that [a]Will:([a]Will:α)∨[a]Will:¬([a]Will:α). However, suppose that at a posterior moment of time I get so stressed with my working that I will that I had not the desire to finish the paper soon. Therefore, [a]Will:¬([a]Will:α) is now true, as it is the antecedent of (4). As such, the consequent, I:([a]Will:α), is also true. So, although earlier my willing this paper to be finished soon was free, my mere willing not to will it transformed it into an inevitable event.

About the charge that the theory of karma leads to determinism and fatalism, I am interested here in what is called causal determinism (see [12] and [15]). Among the several kinds of causal determinism found in the philosophical literature, we can distinguish between what might be called *global determinism* and *local determinism*:

Global Determinism The world is determinist (or is governed by determinism) if and only if, given a specified way things are at a moment of time *t*, the way things go in moments posterior to *t* is fixed.

Local Determinism A state of affairs *e* happening at moment of time *t* is determinist if and only if *e*'s being the case was fixed by the way things were at a moment anterior to *t*.

Given the conceptual tools we have so far, the formula bellow is the closest we can get to these two definitions of determinism:

(Det) α →P(I:Fα)

where P is the past tense modal operator: if α is a formula, Pα means that α was the case. (Det) says that if α is the case, then that α will be the case was inevitable.

By saying that (Det) is the closest we can get to these definitions of determinism I mean two things. First, that (1) if state of affairs *e* is determinist and α is its representation in our language, then (Det) is true and that (2) if the world is determinist, then (Det) is true for every proposition α. Second, although the other way round is clearly not true, that is to say, (Det) might be true with no causal determinism, from the point of view of the formal representational machinery we have so far, it is the farthest we can get. In

the next section, when I introduce the full formal apparatus of branching time theory, I will be able to give a much more refined account of determinism. From now on, I shall take (Det) as implying both sorts of determinism.

As far as fatalism is concerned, it has been understood in at least two different ways: as an attitude of resignation in the face of some future events which are thought to be inevitable, or as the view that whatever will happen in the future is inevitable or unavoidable, in the strict sense of the words (not avoidable or evitable, by no human means at least) [31]. I am concerned with this second sense of fatalism. There are alike at least two sorts of fatalism in this sense:

Global Fatalism A fatalist view is one that claims that whatever will happen in the future is inevitable.

Local Fatalism A fatalist view of future event e claims that e is inevitable.

Both views can be represented as follows:

(Fat) $F\alpha \rightarrow I{:}F\alpha$

Now, (KIP) says that if α is true and agent a is affected by α, then α is inevitable. It is quite reasonable to suppose that most propositions will affect some agent. It is hard to do any sort of quantification here, but calling these propositions *agent-affecting propositions* (α is an agent-affecting proposition if and only if [a]Affect:α is true for some agent a), we have that the following is true:

(10) $\alpha \rightarrow I{:}\alpha$, where α is an agent-affecting proposition

If α is inevitable, it is reasonable to suppose that in the past the proposition that α will be true was also inevitable; in other words, I:α seems to imply P(I:Fα). Therefore, from (10) we get

(Det$_{aa}$) $\alpha \rightarrow P(I{:}F\alpha)$, where α is an agent-affecting proposition

Also from (10), through a very simple substitution, we get

(Fat$_{aa}$) $F\alpha \rightarrow I{:}F\alpha$, where $F\alpha$ is an agent-affecting proposition

(KIP) therefore implies qualified versions of both local determinism and local fatalism, as well as of (FW)—(4) is basically an agent-affecting proposition version of (FW).

Two points have to be made before I end this section. First, the conclusions I got about free-will, determinism and fatalism were possible due to (KIP). In its turn, (KIP) was obtained through (EP), (CEP) and the transitivity of material implication. Therefore, at least inside this simple representational framework, the connection between the theory of karma on one side and free-will, determinism and fatalism on the other rest on these two principles: (EP) and (CEP); rejecting either one or the other would dissolve the connection. Otherwise said, if a karma theorist does not want to commit himself or herself to qualified, but quite strong versions of a determinist, fatalist and free-will-less view of the world, he or she will have to reject either (EP) or (CEP).

Second, there is another way we can see the conjunction of (EP) and (CEP) as problematic. A quite reasonable principle of our inevitability operator is this:

(Inev)   I:α → α

From it along with (CEP) we get [a]KE:α → α. Along with (PAP) we get [a]KE:α → α ∧[a]Affect:α. But this, along with (EP) allows us to obtain

(11)   α ∧[a]Affect:α ↔[a]KE:α,

that is to say, that α's being a karmic effect of *a* is equivalent to the α's being true and affecting *a*. This is problematic because it completely empties the concept of karmic effect, which now can be defined in terms of the concepts of truth and affectability.

4. Branching Time Theory

The purpose of the last section was to show how some relevant consequences can be drawn from a simple representation of the principles and concepts involved in our discussion. It also had the purpose of providing a motivation for a deeper investigation of the theme, preferentially inside a formal framework in which the meanings of the elements of the language, in special the modal operators, are satisfactorily defined. This is what I shall be doing from now on.

To start with, here is the formal definition of the language (I shall limit myself to the propositional case):

Definition 4.1. Let A be a countable set of names of agents and P a countable set of propositional symbols. The *language of karma* (in symbols: L_K) is defined as follows:

(i)   If α ∈P, then α ∈ L_K;
(ii)   If α, β ∈ L_K, then ¬α,α ∧ β,α ∨ β,α → β ∈ L_K;
(iii)   If α ∈ L_K, then Pα,Fα ∈ L_K;
(iv)   If α ∈ L_K, then HI:α,LI:α,GI:α,KI:α ∈ L_K;
(v)   If α ∈ L_K and a ∈ A, then [a]KE:α,[a]Will:α ∈ L_K;

As we saw, the modal operator Affect is defined in terms of Will, ∨ and ¬:

(DEF$_w$) [a]Affect:α $=_{def}$ ([a]Will:α)∨([a]Will:¬ α)

The only items in this definition I have not explained yet are the ones contained in (iv). They are inevitability operators: HI, LI, GI and KI stand, respectively, for *historical inevitability*, *local inevitability*, *global inevitability* and *karmic inevitability*. HI:α means that α is historically inevitable, LI:α that α is locally inevitable, GI:α that α is globally inevitable and KI:α that α is karmically inevitable. They are meant to replace the inevitability operator I; why this is required will be clear after I lay down the semantic components needed to evaluate these operators.

Definition 4.2. A *model* M is a sextuple $< T, \prec, V, G, || \; ||, W, K >$ where

(i)   T is a non-empty set of moments of time;

(ii) $\prec$ is a non-reflexive, transitive and asymmetric relation between moments of time such that, for any $t_1, t_2, t_3 \in$ T, if t_3 is such that $t_1 \prec t_3$ and $t_2 \prec t_3$, then either $t_1 = t_2$ or $t_1 \prec t_2$ or $t_2 \prec t_1$ (*no backward ramification condition*) and for any $t_1, t_2 \in$ T, there is $t_3 \in$ T such that $t_3 \prec t_1$ and $t_3 \prec t_2$ (*historical connection condition*);

A moment of time is a spatially complete and instantaneous event or, if you wish, an instantaneous world-slice. It is a nonrelativistic idea, but one which helps us to settle our intuitions about time and causation ([1], p. 370). $\prec$ is a causal temporal ordering relation between moments of time. If $t_1, t_2 \in$ T are such that $t_1 \prec t_2$, then we say that t_2 is posterior to t_1 (or that t_1 is anterior to t_2) as well as that t_2 is partially, at least, caused by t_1. $\prec$ sets our view of time as branching instead of linear: there might be $t_1, t_2 \in$ T such that neither $t_1 \prec t_2$ nor $t_2 \prec t_1$. However, as non-connected as they might be, there is always a moment $t_3 \in$ T anterior to t_1 and t_2 (that is: $t_3 \prec t_1$ and $t_3 \prec t_2$) which 'links' them to the same 'tree'; this is what the historical connection condition says. We say in this case that t_1 and t_2 are incompatible future possibilities of t_3. But there cannot be a moment t_4 anterior to $t_1 (t_4 \prec t_1)$ and different from t_3 such that neither $t_3 \prec t_4$ nor $t_4 \prec t_3$. In other words, while the future is branching, the past is linear; this is what the no backward ramification condition says. Finally, the condition of asymmetry indicates that moments of time are non-repeatable events.

From these two first components of our model a couple of important concepts can be defined:

Definition 4.3. A *history* is a maximal chain defined on $< T, \prec >$. We call H the set of all histories.

Definition 4.4. Let h,h'$\in$ H and t $\in$ T. h and h' *share the same past at t* (in symbols: h $\cong_t$ h') iff, for all $t' \prec t$, $t' \in h$ and $t' \in h'$.

Definition 4.5. Let h,h'$\in$H. h and h' *split at t* (in symbols: h $\leftrightarrows_t$ h') iff h $\cong_t$ h' and there is one and only one $t' \in h$ such that $t \prec t'$ and h $\cong_{t'}$ h'.

Definition 4.6. The set of *alternatives* $\mathcal{T}$ is a partition of T such that t,t'$\in$ T belong to the same partition (which might be referred to either as $\mathcal{T}_t$ or $\mathcal{T}_{t'}$) iff, for all h,h'$\in$ H, h $\cong_t$ h' sss h $\cong_{t'}$ h'.

Definition 4.7. Let t,t'$\in$ T and h,h'$\in \mathcal{T}_t$ be such that t $\in$ h but t$\notin$h'. t' is the *splitting moment* of $\mathcal{T}_t$ (in symbols: $t' \leftrightarrows \mathcal{T}_t$) iff h $\leftrightarrows_{t'}$ h'.

Definition 4.8. The set of *instants* I is a partition of T such that t,t'$\in$ T belong to the same partition (which might be referred either as I_t or $I_{t'}$) iff $\mathcal{T}_t = \mathcal{T}_{t'}$ or t'' and t''' belong to the same instant ($I_{t''} = I_{t'''}$), where t'' is the splitting moment of $\mathcal{T}_t$ and t''' is the splitting moment of $\mathcal{T}_{t'}$.

Our concept of history corresponds to the intuitive notion of history, except that the intuitive notion encompasses only the past, whereas

ours encompasses the future as well. In the figure below, while moments $t_{11}, t_{21}, t_{31}, t_{41}$ and t_{51} belong to history h_1, moments $t_{11}, t_{21}, t_{31}, t_{42}$ and t_{54} belong to history h_2.
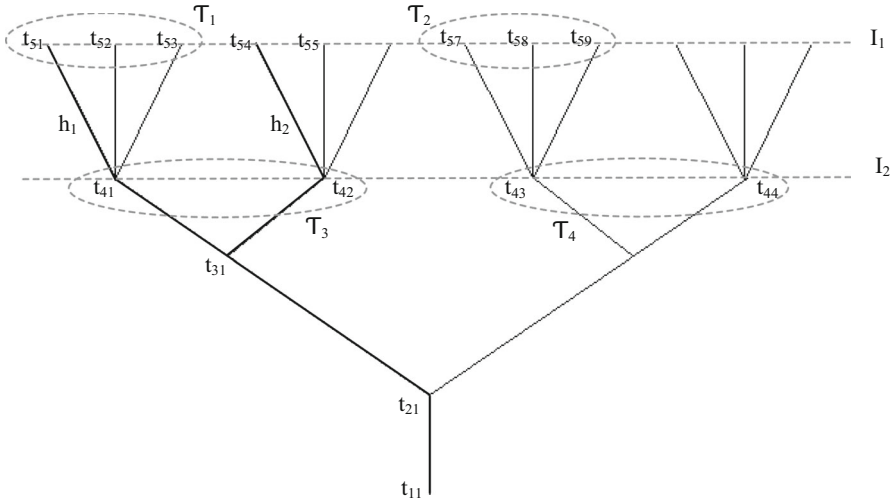
Despite being different, h_1 and h_2 share the same past at t_{41} as well as at t_{42}, t_{31}, t_{21} and t_{11}. That is because they split at t_{31}. t_{41} and t_{42} belong to the same alternative $\mathcal{T}_3$ (which can also be referred to as $\mathcal{T}_{t_{41}}$ or $\mathcal{T}_{t_{42}}$); t_{43} and t_{44} belong to alternative $\mathcal{T}_4$; t_{51}, t_{52} and t_{53} belong to $\mathcal{T}_1$; and t_{57}, t_{58} and t_{59} belong to $\mathcal{T}_2$. Finally, while t_{51}, t_{52}, $t_{53}, t_{54}, t_{57}, t_{58}, t_{59}$ and the other unmentioned moments of the same level belong to instant I_1 (which might also be referred to as $I_{t_{51}}, I_{t_{52}}, I_{t_{53}}$ and so on), t_{41}, t_{42}, t_{43} and t_{44} belong to instant I_2.

Definition 4.2 (*continuation*). A *model* M is a sextuple $< T, \prec, V, G, ||\ ||, W, K >$ where

(iii) V: PxT $\rightarrow$ {true, false} is a function which, given a moment of time, maps propositional symbols to truth-values;

V is the truth function which evaluates the truth value of atomic formulas for each one of the moments of T.

Even though I have not finished introducing the elements of the model, it might be useful to stop for a while and comment on the way tense and inevitable formulas are evaluated. In order to do that, I will have to start defining the evaluation relation $||\!\!-$. In the logic of branching time, the evaluation relation $||\!\!-$ (which, given a model M, says whether a formula α is true at M) has as parameter, besides a moment of time, also a history ([36], p. 323). This is because of the ambiguity concerning future tense formulas. Intuitively, Fα is true at moment t if and only if α is true at a moment posterior to t. But if we say that Fα is true, say, at t_{31}, do we mean by this that for *each* history h passing through t_{31} α is true at some moment t' posterior to t_{31} such that t $\in$ h? Or is it just for *at least* one history passing through t_{31}? Or, still, is it for a

specific history h passing through t_{31} so that besides a moment of time t the evaluation of formulas has also a history h as parameter? Usually branching time theorists pick the third option. M $\Vdash_{t,h} \alpha$ then means that formula α is true or satisfied in model M, moment of time t and history h; naturally, t and h are such that $t \in h$.

Definition 4.9. Let M $=< T, \prec, V, G, \| \, \|, K, W >$ be a model, $t \in T$ a moment of time, $h \in H$ a history such that $t \in h$, $\alpha \in L_K$ a formula and $p \in P$ a propositional symbol.

(i)   M $\Vdash_{t,h}$ p iff V(p,t) = true;

(ii)   M $\Vdash_{t,h} \neg \alpha$ iff M $\nVdash_{t,h} \alpha$[8];

(iii)   M $\Vdash_{t,h} \alpha \wedge \beta$ iff M $\Vdash_{t,h} \alpha$ and M $\Vdash_{t,h} \beta$;

(iv)   M $\Vdash_{t,h} \alpha \vee \beta$ iff M $\Vdash_{t,h} \alpha$ or M $\Vdash_{t,h} \beta$;

(v)   M $\Vdash_{t,h} \alpha \rightarrow \beta$ iff M $\nVdash_{t,h} \alpha$ or M $\Vdash_{t,h} \beta$;

(vi)   M $\Vdash_{t,h}$ Fα iff, for some $t' \in h$ such that $t \prec t'$, M $\Vdash_{t',h} \alpha$;

(vii)   M $\Vdash_{t,h}$ Pα iff, for some $t' \prec t$, M $\Vdash_{t',h} \alpha$;

(viii) M $\Vdash_{t,h}$ HI:α iff, for all $h' \in H$ such that $t \in h'$, M $\Vdash_{t,h'} \alpha$;

(ix)   M $\Vdash_{t,h}$ LI:α iff for all $t' \in \mathcal{T}_t$ and all $h' \in H$ such that $t' \in h'$, M $\Vdash_{t',h'} \alpha$;

(x)   M $\Vdash_{t,h}$ GI:α iff for all $t' \in I_t$ and all $h' \in H$ such that $t' \in h'$, M $\Vdash_{t',h'} \alpha$;

As it can be seen, the semantic definition of logical connectives and propositional symbols follow the standard in modal logic. The same holds for tense operators. As far as the inevitability operators are concerned, HI, LI and GI form a gradation in our view of inevitability: (1) GI:α entails LI:α and (2) LI:α entails HI:α; also (3) LI:α entails P(HI:Fα). In order to illustrate the use of these operators, suppose a model M such that $V(p, t_{51}) = V(p, t_{52}) = V(p, t_{53}) = V(p, t_{41}) = V(p, t_{42}) = V(p, t_{43}) = V(p, t_{44}) = V(p, t_{55}) =$ true and $V(p, t_{54}) =$ false, where p is a propositional symbol of M. In this way, we have that M $\Vdash_{t_{41},h_1}$ HI:Fp, M $\Vdash_{t_{51},h_1}$ LI:p and M $\Vdash_{t_{41},h_1}$ GI:p; but since $V(p, t_{54}) =$ false, M $\nVdash_{t_{54},h_1}$ LI:p and M $\nVdash_{t_{51},h_1}$ GI:p. The definition of KI will have to wait until we introduce the other elements of M.

At this point it should be clearer why I have proposed such a variety of inevitability operators in the place of Sect. 3 single modal operator I. Given a moment of time t, if a future proposition Fα is true in all histories passing through t, α is going to happen no matter which history becomes the case. Therefore we have an operator of historical inevitability. On the other hand, if α is true at t, but also at all alternative moments to t, then α is not only the case: it had to be the case, for even if the course of the world were different, α would still be true. Therefore the operator of local inevitability. If besides that α is true at all co-instantaneous moments to t, even looking at different courses the world could have followed at distant moments in the past, α would still be true. Therefore the notion of global inevitability.

This variety of inevitability operators allows us to state in a more precise way the issues of determinism, fatalism and free-will. While (Det) can be stated in terms of both HI and LI:

---

[8] M $\nVdash_{t,h} \alpha$ stands for it is not the case that M $\Vdash_{t,h} \alpha$.

(Det) $\alpha \rightarrow P(HI{:}F\,\alpha)$
(Det$'$) $\alpha \rightarrow LI{:}\,\alpha$

(Fat) is stated in terms of HI:

(Fat) $F\,\alpha \rightarrow HI{:}F\,\alpha$

It can also be stated, in both weaker and stronger ways, in terms of LI:

(Fat$'$) $F\,\alpha \rightarrow F(LI{:}\,\alpha)$
(Fat$''$) $F\,\alpha \rightarrow LI{:}F\,\alpha$

While (Fat) implies (Fat$'$), the other way round is not true; and while (Fat$''$) implies (Fat), the other way round is not the case. The threat to the existence of free-will can also be represented in terms of LI:

(FW) [a]Will:$\alpha \rightarrow$LI:([a]Will:α)

Now, branching time theory is often seen as a theory of indeterminism (see [1,4,5]). But while branching time theory in general and (Det) in particular captures the notion of local determinism, it is, at first sight, at odds with the notion of a global determinism: in one sense, if our world is globally determinist, there is no branching. There are two ways to answer this. First, by pointing out that our semantic framework allows branching, but does not force it. There can be linear models:

Definition 4.10. Let M $=< T, \prec, V, G, ||\,||, K, W >$ be a model. M is a *linear model* if and only if, for every t_1, t_2, $t_3 \in$ T, if t_3 is such that $t_3 \prec t_1$ and $t_3 \prec t_2$, then either $t_1 = t_2$ or $t_1 \prec t_2$ or $t_2 \prec t_1$. A non-linear model is called a *branching model*.

It is easy to see that if M is a linear model then (Det) holds for every formula α in M. The same can be said about (Fat). Further, in a linear model our three notions of inevitability are extensionally identical.

Second, global determinism and fatalism can be defined in branching models by postulating a *thin red line* (see [1,4] and [5]).[9] The idea is that among the several histories of a (branching) model passing through a distinguished moment, say t, there is a special history, a thing red one,[10] which is how things are really going to happen. For example, suppose h_2 is our thin red line; even though in t_{42} both p and $\neg p$ are historically possible (p is false at t_{54} but true at t_{55}), it is $\neg p$ which turns out to be true in the instant immediately posterior to t_{42}. Therefore, even though the future is in some sense branching, there in fact only one way the world could actually be. I will come back to this idea in the next section.

---

[9] I shall overlook here the issue about the tenability of the doctrine of a thin red line. For an overview of the debate and a defense of it see [6].

[10] The expression comes from the saying that the British Empire was maintained by a thin red line of soldiers in service to the Queen ([1], p. 366).

5. The Semantics of Karma

Continuing the definition of the components of our model, we have as follows:

Definition 4.2 (*continuation*). A *model* M is a sextuple $< \text{T}, \prec, \text{V}, \text{G}, || \; ||, \text{W}, \text{K} >$ where

(iv) G is a non-empty set of agents;
(v) $|| \; ||$: A $\rightarrow$ G is a function which maps (names of) agents to elements of G;
(vi) W: GxT $\rightarrow$ $\aleph$(H)[11] is a function which, for each moment of time t and agent g, gives the will-history of g at t. W is such that, for all $t \in \text{T}$ and $g \in \text{G}$, W(g,t) $\neq \varnothing$ (*non-contradiction condition*);

A *will-history* is a set of histories containing, for a given agent g and moment of time t, all compatible histories with the will of g at t. In symbols: W(g,t). I shall call the members of W(g,t) *will-histories*. With the help of W, I can define the truth-conditions of the modal operator Will:

Definition 4.9 (*continuation*). Let M $=< \text{T}, \prec, \text{V}, \text{G}, || \; ||, \text{K}, \text{W} >$ be a model, t $\in$ T a moment of time, h $\in$ H a history such that t $\in$ h, $\alpha \in L_K$ a formula and p $\in$ P a propositional symbol.

(xi) M $||{\vdash}_{t,h}$ [a]Will:α iff (1) for all h$'\in$ W($||a||$, t) and t$'\in I_t$ such that t$'\in$h$'$, M $||{\vdash}_{t',h'}$ α and (2) there is at least one t$'\in I_t$ such that M $||{\not\vdash}_{t',h'}$ α, for some h$'\in$H such that t$'\in$h$'$;

[a]Will:α is true at t in the case α is true at all histories compatible with the will of a (at t) at the instant to which t belongs. Condition (2) is there to exclude tautologies and inevitable propositions; agents are therefore assumed to satisfy a criterion of minimal rationality according to which there cannot be willing towards tautological or inevitable propositions. Neither can an agent will contradictory propositions; this is guaranteed by the non-contradiction condition in the definition of W.

Here is the final component of our model:

Definition 4.2 (*continuation*). A *model* M is a sextuple $< \text{T}, \prec, \text{V}, \text{G}, || \; ||, \text{W}, \text{K} >$ where

(vii) K: GxT $\rightarrow$ $\aleph$(H) is a function which, for each agent g and moment t, gives the karmic pool of g at t. K is such that K(g,t) $\neq \varnothing$ (*non-contradiction condition*); if h$\in$K(g,t), then for all h$'\in$H such that t$\in$h$'$, h $\cong_t$ h$'$ (*historical relevance condition*); if $\mathcal{T}_t = \mathcal{T}_{t'}$, then K(g,t) $=$ K(g,t$'$) (*alternative condition*); there is at least one h $\in$ H such that h $\in$ K(g,t) for all g $\in$ G (*collective karma condition*); and there is at least one h $\in$ H such that t $\in$ h and h$\notin$K(g,t) (*incompatibility principle condition*).

A *karmic pool* is a set of histories containing, for a given agent g and moment of time t, all karmically compatible histories with the intentional actions done

---

[11] $\aleph(\Gamma)$ is the powerset of Γ.

by g up to t as well as with the psychological state of g at t. The idea is very simple. Up to moment t, agent g has done several actions with specific states of mind, either in the life he or she is living at moment t or in past ones. According to karma theory, these actions have produced several effects, to be actualized and 'enjoyed' by g at t or at moments posterior to t. In some histories these effects are the case; in others they are not. The histories in which these effects are the case are karmically compatible with the actions done by g up to t; the set containing all of them is called the karmic pool of g at t (in symbols: K(g,t)).

With the help of K, I can finally define the karmic effect modal operator:

Definition 4.9 (*continuation*). Let $M = < T, \prec, V, G, ||\,||, K, W >$ be a model, t $\in$ T a moment of time, h $\in$ H a history such that t $\in$ h, $\alpha \in L_K$ a formula and p $\in$ P a propositional symbol.

(xii) M $||\!\vdash_{t,h}$ [a]KE:α iff (1) for all h$' \in$ K($||a||$, t) and t$' \in \mathscr{T}_t$ such that t$' \in$ h$'$, M $||\!\vdash_{t',h'} \alpha$, (2) there is t$' \in I_t$ such that M $||\!\nvdash_{t',h'} \alpha$, for some h$' \in$ H such that t$' \in$ h$'$ and (3) M $||\!\vdash_{t',h'} \alpha$ or M $||\!\vdash_{t',h'} \neg \alpha$ for all h$' \in$ W($||a||$, t) and t$' \in I_t$ such that t$' \in$ h$'$.

[a]KE:α is true at t in the case α is true at all karmically compatible histories with the past actions of a and alternative moments to t. Similarly to Will, (2) is there to exclude tautologies and inevitable propositions from the set of karmic effects of a. Condition (3) guarantees that if α is a karmic effect of a, then α affects a; as one would expect, the Person-Affecting Principle is a tautology in our system:

(PAP)   [a]KE:$\alpha \to$ [a]Affect:α

As far as the conditions in the definition of K are concerned, the non-contradiction condition is there to prevent formulas of the kind [a]KE:($\alpha \wedge \neg \alpha$) to be valid. The historical relevance condition guarantees that all histories belonging to the karmic pool of a at t pass through moments alternative to t. The rationale behind it is that the karmic pool of a at t was produced by actions and psychological states of a at moments anterior to t; at the limit, it was produced at the moment t' immediately anterior to t. As such, since the karmic pool contains histories which might be lived by a given t, the karmic pool only makes sense if it contains just those histories which are possible considering t'. Another consequence of the karmic pool being produced at t' is that the karmic pool of alternative moments is the same; that is the alternative condition. That is why we have, in the definition of KE, considered alternative moments and not instantaneous moments, as we have done in the definition of Will.

Finally,[12] the collective karma condition guarantees that there is at least one history common the karmic pools of all agents. It could be called the

---

[12] The explanation of the incompatibility principle condition will have to wait until Sect. 6, where I shall explain how the Incompatibility Principle is taken into account in my semantics.

some-thing-happens condition as well, and the reason for this is crucial for my analysis. To understand why, consider the following definitions:

Definition 5.1. Let M $=< $ T$, \prec, V, G, || ||, K, $W $>$ be a model and t $\in$ T a moment of time. The *karmic world* at t (in symbols: $\mathcal{K}_t$) is the intersection of all karmic pools of all agents; in symbols: h $\in \mathcal{K}_t$ iff h $\in$ K(g,t) for all g $\in$ G.

Definition 5.2. Let M $=< $ T$, \prec, V, G, || ||, K, $W >$ be a model and t $\in$ T a moment of time. *t* is *karmically normal* iff there is at least one h $\in \mathcal{K}_t$ such that t $\in$ h. $T_K \subseteq$T is the set of all karmically normal moments of T.

Definition 5.3. Let M $=< $ T$, \prec, V, G, || ||, K, $W >$ be a model and t $\in$ T a moment of time. If t $\in T_K$, then the set of histories passing through it which belong to the karmic world at *t* (in symbols: $\mathcal{H}_t$) is called the *karmically normal histories of t*; in symbols: $\mathcal{H}_t = \{$h $\in \mathcal{K}_t|$t $\in$ h$\}$.

   With these definitions at hand, I can define the truth conditions of our last modal operator: the karmic inevitability operator.

Definition 4.9 (*continuation*). Let M $=< $ T$, \prec, V, G, || ||, K, $W >$ be a model, t $\in$ T a moment of time, h $\in$ H a history such that t $\in$ h, $\alpha \in L_K$ a formula and p $\in$ P a propositional symbol.

(xiii) M $||\vdash_{t,h}$ KI:α iff for all t$'\in T_K$ such that t$'\in \mathcal{T}_t$ and all h$'\in \mathcal{H}_t$, M $||\vdash_{t',h'} \alpha$.

   In other words, α is karmically inevitable if and only if α is true at all karmically normal moments alternatives to *t* and histories passing through these moments. Since only moments alternative to *t* are being considered, LI:α implies KI:α. We then have the complete list of the relations holding between our inevitability operators:

Theorem 5.1. Let M $=< $ T$, \prec, V, G, || ||, K, $W >$ be a model, t $\in$ T a moment of time and h $\in$ H a history and $\alpha \in L_K$ a formula.

   (i)   If M $||\vdash_{t,h}$ GI:α then M $||\vdash_{t,h}$ LI:α;
   (ii)   If M $||\vdash_{t,h}$ LI:α then M $||\vdash_{t,h}$ HI:α;
   (iii)   If M $||\vdash_{t,h}$ LI:α then M $||\vdash_{t,h}$ KI:α.

   Now, going back to the some-thing-happens condition, if the theory of karma is really true, the histories belonging to $\mathcal{K}_t$ cannot be possible courses of the way things could be in the same way that histories not contained in $\mathcal{K}_t$ are. To suppose they can is to trivialize the theory of karma; if $\mathcal{K}_t$ has no distinguished status, then there is hardly a sense in saying that an individual is going to suffer or enjoy the karmic effects of his or her actions. Therefore, if anything is to happen at all, $\mathcal{K}_t$ cannot be empty.

   But there is more than this. What seems to be at stake here is a doctrine very similar to the thin red line, with the difference that instead of a line we have a *thin red tree*, or a *karmic thin red tree*. The other difference is that since at every moment possibly new actions will be done and new karmic effects become the case, at every moment there will be a different karmic thin red tree, which is to be distinguished from the rest of the tree formed from <T, $\prec$>. But distinguished in which sense? Usually branching time theorists adopt

a realist view of time: all moments of time have the same ontological status; they are all equally existing entities ([6], pp. 106–107). Following this stream, one first answer could be to say that the histories of $\mathcal{K}_t$ are distinguished in an ontological sense: from the point of view of t, histories belonging to $\mathcal{K}_t$ are all that is ontologically possible; they are the real possibilities; the world cannot be in a way which is not contained in $\mathcal{K}_t$.

This position poses a couple of problems. First, what is the status of histories not contained in $\mathcal{K}_t$? Are they some sort of abstract or linguistic representation? But representation of what? It cannot be representation of ways the world could be, for we have assumed the world cannot be different from $\mathcal{K}_t$ (see [5], pp. 156–159). Second, even if we could somehow solve this issue, at every moment there will possibly be different karmic histories. This means that some histories would suddenly become ontologically possible, while others will cease to be so. It is certainly not trivial how one could support such an odd metaphysical doctrine.

One way to avoid these problems is to adopt a more standard position and take all histories as having the same status, that is to say, as being ontological possibilities of the way the world could be; altogether, they would form what Belnap calls *our world* ([1], pp. 370–371). As far as $\mathcal{K}_t$ is concerned, perhaps all we could say is that it is a brute fact that only histories of $\mathcal{K}_t$ can become the case. Although in one sense solving the problem, this path is clearly unsatisfactory, for we do want to know why the histories of $\mathcal{K}_t$ have such a distinguished status: to say that they have it because they have it is frustrating, to say the least.[13]

I am going to follow a path which, if does not solve these problems, at least avoids them. In fact, this path allows us to remain silent about the ontological status of moments of time and histories. In order to explain it, I need to go back a little and speak about the standard way this issue appears in possible world semantics.

A model M in modal logic is composed by a non-empty set of possible worlds W, an accessibility relation between worlds R and a truth function which maps propositional symbols and possible worlds to truth-values. The evaluation relation $\Vdash$ is defined in terms of models and possible worlds: M $\Vdash_w \alpha$ means that α is true at world w of model M; if α is true at all worlds of M we write M $\Vdash \alpha$. As far as the consequence relation is concerned (something we have not yet defined for our logic), the most complete way to define it is to make reference to two sets of premises: a set of global premises and a set of local premises [13]. Formally we have as follows: α is a logical consequence of A and B (in symbols: $A \dotdiv B \vDash \varphi$) if and only if, for every model M such that M $\Vdash \alpha$ for all $\alpha \in A$, and for every w of M such that M $\Vdash_w \beta$ for all $\beta \in B$, M $\Vdash_w \varphi$; while A is the set of global premises, B is the set of local premises. From a proof-theoretical point of view, the distinction appears in the application of the necessitation rule (from α conclude $\Box \alpha$); it can be applied only to global premises.

---

[13] See [5] and [6] for a discussion on these questions on the context of branching time theory and the thin red line.

Now, despite the debate over the nature of possible worlds and whether or not there are distinguished worlds (such as the actual world), from an inferential point of view, at the same time that there is no need of engaging in such sort of debate here, there are preferred or distinguished models and worlds. When we write a derivation from $A \dot{\dashv} B$ to ϕ, all steps of the derivation can be seen as valid only inside a very specific set of models and worlds. If formula ϕ is derived only from global premises, we say that ϕ is true in all models M (and in all worlds of M) such that M $\Vdash \alpha$ for all $\alpha \in A$. It is from this perspective that we can write $\Box \phi$: this does not mean that ϕ is true in all worlds of all models, but only in the worlds of those models selected by A. And if there is a local premise involved in the derivation of ϕ, then by writing ϕ we simply mean that it is true in these models and worlds w such that M $\Vdash_w \beta$ for all $\beta \in B$. But by saying that we are not giving any special status to those models and worlds.

Before defining the notion of logical consequence, let me introduce some quite straightforward definitions:

Definition 5.4. Let M $= < T, \prec, V, G, ||\,||, K, W >$ be a model and $\alpha \in L_K$ a formula. M $\Vdash \alpha$ iff, for all $t \in T$, M $\Vdash_{t,h} \alpha$ for all $h \in H$ such that $t \in h$; M $\Vdash_K \alpha$ iff, for all $t \in T_K$, M $\Vdash_{t,h} \alpha$ for all $h \in \mathcal{H}_t$.

Definition 5.5. Let A$\subseteq L_K$ be a set of formulas. The sets of models $\mathcal{M}_G$ and $\mathcal{M}_K$ are defined as follows: $\mathcal{M}_G(A) = \{M | M \Vdash \alpha$ for all $\alpha \in A\}; \mathcal{M}_K(A) = \{M | M \Vdash_K \alpha$ for all $\alpha \in A\}$.

Now the definition of logical consequence:

Definition 5.6. Let A,B,C $\subseteq L_K$ be three sets of formulas and $\varphi \in L_K$ a formula. φ is a *karmic logical consequence* of A, B and C (in symbols: $A \dot{\dashv} B \dot{\dashv} C \vDash \varphi$) if and only if, for all models M $\in \mathcal{M}_G(A) \cap \mathcal{M}_K(B)$, $t \in T_K$ (of M) and h$\in \mathcal{H}_t$ such that M $\Vdash_{t,h} \phi$ for all $\phi \in C$, M $\Vdash_{t,h} \varphi$. $\vDash \varphi$ is an abbreviation for $\varnothing \dot{\dashv} \varnothing \dot{\dashv} \varnothing \vDash \varphi$, in case of which I call φ a *tautology*.

While A is the set of global premises, B and C are the set of karmic premises and local premises, respectively. From a proof-theoretical point of view, which shall not be developed here, while all inevitability rules (α/GI:α, α/LI:α, α/HI:α and α/KI:α) would be applied only to formulas derived exclusively with the help of global premises, for formulas derived with the help of some karmic premise only the karmic inevitability rule (α/KI:α) could be used; if the formula uses local premises no inevitability rule could be used.

Following the idea of global and local premises of modal logic explained above, I am doing three restrictions in the selection of the models, moments of time and histories to be used in the evaluation of $A \dot{\dashv} B \dot{\dashv} C \vDash \varphi$. The first one is that I pick up only those models M such that M $\Vdash \alpha$ for all global premises α. Recall that in order for a formula α to be satisfied by M (M $\Vdash \alpha$) it must be satisfied by all moments t of M and all histories h passing through t (M $\Vdash_{t,h} \alpha$ for all $h \in H$ such that t $\in h$). Out of these models, I pick up the models M in which all karmic premises $\beta \in B$ are true in all its karmically normal moments t and karmically normal histories of t. This is the second restriction. Finally, for each of these moments, I pick up only those karmically

normal moments t and karmically normal histories of t that satisfy all local premises. It is in those models, moments and histories that our conclusion φ has to be true for the consequence relation to hold.

The relevant point here is that I do not postulate any kind of ontological distinction of a specific set of moments and histories. But as far as the consequence relation is concerned, I pick up only karmic normal times and, consequently, histories belonging to the karmic world, solving thus, in one sense, the issue about the some-thing-happens condition and the distinguished status that histories belonging to $\mathcal{K}_t$ should have. The relation $A \mathbin{\vdots} B \mathbin{\vdots} C \vDash$ φ can perhaps be read as follows: if the theory of karma is true and if those premises are true according to their respective roles, then φ is also true.

6. Karma, Freedom of Will, Determinism and Fatalism

With the semantic framework fully defined, I can analyze more rigorously the questions posed at the beginning of Sect. 3. To begin with, let me recall the way I have represented the notions of determinism, fatalism and the threat to free-will in Sect. 4: (Det) α →P(HI:Fα); (Det′) α →LI:α; (Fat) Fα →HI:Fα; (Fat′) Fα →F(LI:α); (Fat″) Fα →LI:Fα and (FW) [a]Will:α →LI:([a]Will:α). It should be noted that none of these formulas are tautology in our system:

Theorem 6.1. The following relations are valid[14]:

(i) $\not\vDash$ α →P(HI:Fα);
(ii) $\not\vDash$ α →LI:α;
(iii) $\not\vDash$ Fα →HI:Fα;
(iv) $\not\vDash$ Fα →F(LI:α);
(v) $\not\vDash$ Fα →LI:Fα;
(vi) $\not\vDash$ [a]Will:α →LI:([a]Will:α).

Therefore, strictly speaking our logic of karma does not embrace determinism or fatalism, neither is it against free-will.

However, there is another inevitability operator—KI—, which is very alike to the local inevitability operator LI: both take into consideration only alternative moments and histories passing through those moments, the difference being that KI restricts the moments to karmically normal moments t and histories to karmically normal histories of t. Now, if all these principles can be represented solely with the help of LI, even if in stronger versions, they can also be represented with the help of KI. We have therefore the following karmic versions of (Det), (Fat) and (FW):

(Det$_K$)   α →KI:α
(Fat$_K$)   Fα →KI:Fα
(FW$_K$)   [a]Will:α →KI:([a]Will:α)

---

[14] $\not\vDash$ α stands for it is not the case that $\vDash$ α.

We have seen in Sect. 3 how these principles—in qualified versions and expressed in terms of our generic inevitability operator I—could be derived from our karma principles. A similar movement can be made in terms of KI. From (EP) and (CEP) (now expressed in terms of KI)

(EP)   α ∧[a]Affect:α →[a]KE:α
(CEP) [a]KE:α →KI:α

we obtain (KIP):

(KIP) α ∧[a]Affect:α →KI:α

From (KIP) we obtain a qualified version of (FW)

(FW$_{K\text{-}aa}$) [a]Will:α →KI:([a]Will:α), where [a]Will:α is a second-order will-ing[15]

as well as of (Det$_K$):

(Det$_{K\text{-}aa}$) α →KI:α, where α is an agent-affecting proposition,

from which we obtain the following qualified version of (Fat$_K$):

(Fat$_{K\text{-}aa}$) Fα →KI:Fα, where α is an agent-affecting proposition.

Finally, from (Inev), now represented with the help of KI—

(Inev) KI:α → α

—, (PAP), (EP) and (CEP) we get what I have called the trivialization of KE:

(8) α ∧[a]Affect:α ↔[a]KE:α

The important question now is: are all these principles valid in our logic of karma? It is a trivial task to show that (CEP), (PAP) and (Inev) are:

Theorem 6.2. The following relations are valid:

(i) $\models$ [a]KE:α →KI:α
(ii) $\models$ [a]KE:α →[a]Affect:α
(iii) $\models$ KI:α → α

The same, however, cannot be said about (EP). To see this, suppose that $\models$ α ∧[a]Affect:α →[a]KE:α. If this is so, then for all models M, M $\Vdash_K$ α ∧[a]Affect:α →[a]KE:α, that is to say, for every t ∈ T$_K$, M $\Vdash_{t,h}$ α ∧[a]Affect:α →[a]KE:α for all h∈ $\mathcal{H}_t$. But there is no contradiction at all in a model M, moment t ∈ T$_K$ (of M) and history h ∈ $\mathcal{H}_t$ such that M $\Vdash_{t,h}$ α, M $\Vdash_{t,h}$ [a]Will:α but M $\not\Vdash_{t,h}$ [a]KE:α: it suffices that there is h′∈ K(||a||, t) and t′∈ $\mathcal{T}_t$ such that t′∈h′ such that M $\not\Vdash_{t',h'}$ α. Therefore, $\not\models$ α ∧[a]Affect:α →[a]KE:α.

---

[15] As I have defined in Sect. 3, [a]Will:a is a second-order willing if and only if it is true that [a]Will:([a]Will:α) or that [a]Will:([a]Will:¬α).

In order to see why this is so, let us ponder on what it would take for (EP) to be a tautology in our logic. Since one part of the antecedent of (EP), namely [a]Affect:α, guarantees conditions (2) and (3) of KE's semantic definition, all it takes for α ∧[a]Affect:α →[a]KE:α to be a tautology is to establish a connection between the other part of the antecedent (α) and condition (1) (which requires that for all h′∈ K(||a||, t) and t′∈ 𝒯ₜ such that t′∈h′, M ||⊢ₜ′,ₕ′ α). From the semantic structure we have, the truth of α would guarantee the truth of [a]KE:α in one condition only: if the model at hand were a linear one:

Theorem 6.3. Let M be a model. If M is a linear model, then M ||⊢ₖ α ∧[a]Affect:α →[a]KE:α.

However, even if this works from a formal point of view, it does not do it from a conceptual standpoint; and this has to do with the two remaining principles we have introduced in Sect. 2. The Change Principle says that the karmic pool of an agent can change. In our semantics this is represented by the fact that K is a moment of time function and that the histories belonging to K(g,t) are just a small subset of H. That is to say, there are probably many incompatible ways the world could be which are not part of K(g,t) and which might be part of K(g,t′), where t′ is a moment immediately posterior to t. Also, moments which are part of K(g,t) might cease to be so. This is made possible by our tree-like time structure, where time branches in the future.

It has also to do with the Incompatibility Principle, which says that there are ways the world could be which are incompatible with the karmic effects of a at t. This principle is guaranteed by the restrictions imposed upon K, but also by the tree-like structure of our semantics. If there were not different and incompatible ways the world could be from the perspective of moment of time t, there would not be ways the world could be which are incompatible with the karmic effects of a at t. From this is clear that the branching of future is not an optional feature of our model: in order for the Incompatibility Principle to hold, there must be a branching of future histories. As a consequence of this, there cannot be linear models in our logic. We therefore have to reformulate our definition of karmic logical consequence.

We start by defining what we call a karmic model—

Definition 6.1. Let M be a model. M is a *karmic model* if and only if M a branching model.

—to then redefining the notion of karmic logical consequence so that is applied only to karmic models:

Definition 5.6 (*reformulation*). Let A,B,C ⊆ Lₖ be three sets of formulas and φ ∈ Lₖ a formula. φ is a *karmic logical consequence* of A, B and C (in symbols: A ∸ B ∸ C ⊨ φ if and only if, for all karmic models M∈ 𝓜_G(A)∩𝓜_K(B), t ∈ Tₖ (of M) and h ∈ 𝓗ₜ such that M ||⊢ₜ,ₕ ϕ for all ϕ ∈ C, M ||⊢ₜ,ₕ φ.

In this new framework, there is no place for (KE) and for the threat that it poses to the freedom of will and to a fully non-deterministic world.

One might however object that our semantics still does not fully guarantee that the Incompatibility Principle is satisfied. It might be that K(g,t) contains all histories which pass through t, which makes that, from the point of view of t, there are no ways the world could be which are incompatible with the karmic effects of a at t. That this does not happen is guaranteed by the incompatibility principle condition present in the definition of K.

7. Conclusion

I have presented in this paper what I could call 'a logic of karma'. It has been done from a semantic point of view only, and its formal properties were barely scratched here. Although to formally develop this framework is certainly an important task—and I plan to do this in future works—, the main reason for introducing it was to provide a minimally rigorous framework in which the relations that exist between karma theory on one hand and determinism, fatalism and free-will on the other could be analyzed.

My main conclusion was twofold. First, that the basic principles of karma theory, together with a couple of very simple logical and conceptual principles, do imply a sort of determinism, fatalism and lack of free-will. There is however some more to say about these principles. Despite their apparent individual intuitiveness, taking them together seem to trivialize the very notion of karmic effect. Second, after coming up with a full semantic framework in which the notions of karmic effects, inevitability and will could be represented, it turned out that one of these principles, the explanation principle—if an actual state of affairs s affects a, then s is a karmic effect of a—is not generally valid. As a consequence of that, the threat to free-will, as well as the link between karma theory and determinism and fatalism, were vanquished. Although perhaps a good news for the karma theorist, there is a high price to pay: the loss of karma theory's explanatory power, which is allegedly one of its greatest virtues.

This conclusion of course depends on the features of the semantic framework inside of which the principles were represented. A critic could bring into question virtually any of the several theoretical decisions I made in the construction of this framework. Unfortunately I would not have space to anticipate and respond to these criticisms; all I can say is that I have followed the standard way of representing agentive sentences in branching time theory. Also, my approach to karma theory was a minimal one: besides taking into consideration only the core of karma theory, the theoretical decisions I have made were also minimal (as well as reasonable, I believe). Naturally the principles I laid down in Sect. 3 are not exclusive to theories of karma: although I have argued that any theory of karma should satisfy these principles, some theory might satisfy them without having nothing to do with karma or reincarnation.

References

[1] Belnap, N., Green, M.: Indeterminism and the thin red line. Philos. Perspect. **8**, 365–388 (1993)

[2] Belnap, N., Perloff, M.: Seeing to it that: a canonical form for agentives. Theoria **54**, 175–199 (1988)

[3] Belnap, N., Perloff, M.: The way of the agent. Studia Logica **51**, 463–484 (1992)

[4] Belnap, N., Perloff, M., XU, M.: Facing the Future: Agents and Choices in Our Indeterministic World. Oxford University Press, Oxford (2001)

[5] Benovsky, J.: Branching and (in)determinism. Philos. Papers **42**, 151–173 (2013)

[6] Borghini, A., Torrengo, G.: The metaphysics of the thin red line. In: Correia, F., Iacona, A. (eds.) Around the Tree: Semantic and Metaphysical Issues Concerning Branching and the Open Future, pp. 105–125. Kluwer, Dordrecht (2013)

[7] Chadha, M., Trakakis, N.: Karma and the problem of evil: a response to Kaufman. Philos. East West **57**, 533–556 (2007)

[8] Chapple, C.: Karma and Creativity. State University of New York Press, Albany (1986)

[9] Chellas, B.: Time and modality in the logic of agency. Studia Logica **51**, 485–518 (1992)

[10] Clooney, F.: Evil, Divine omnipotence, and human freedom: Vedanta's theology of karma. J. Religion **69**, 530–548 (1989)

[11] Creel, A.: Contemporary philosophical treatments of karma and rebirth. In: Neufeldt, R. (ed.) Karma and Rebirth. Post Classical Developments, pp. 1–14. State University of New York Press, Albany (1986)

[12] Earman, J.: A Primer on Determinism. Reidel, Dordrecht (1986)

[13] Fitting, M.: Basic Modal Logic. In: Gabbay, D., Hogger, D., Robinson, J. (eds.) Handbook of Logic in Artificial Intelligence and Logic Programming, vol. 1, pp. 365–448. Oxford University Press, Oxford (1993)

[14] Ginet, C.: On Action. Cambridge University Press, Cambridge (1990)

[15] Hoefer, C.: Causal Determinism. In: Zalta, E. (ed.) Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/determinism-causal/ (2016)

[16] Horty, J., Belnap, N.: The deliberative stit: a study of action, omission, ability and obligation. J. Philos. Logic **24**, 583–644 (1995)

[17] Humphreys, C. Karma and Rebirth. Curzon Press, London (1983)

[18] Jaini, P.: Karma and the problem of rebirth in Jaininsm. In: O'Flaherty, W. (ed.) Karma and Rebirth in Classical Indian Traditions, pp. 217–238. Motilal Banarsidass, Delhi (1983)

[19] Kaufman, W.: Karma, Rebirth, and the problem of evil. Philos. East West **55**, 15–32 (2005)

[20] Kaufman, W.: Karma, Rebirth, and the Problem of Evil: A Reply to Critics. Philosophy East and West **57**, 559–560 (2007)

[21] Long, B.: The concepts of human action and rebirth in the mahabharata. In: O'Flaherty, W. (ed.) Karma and Rebirth in Classical Indian Traditions, pp. 38–60. Motilal Banarsidass, Delhi (1983)

[22] Mackie, J.: The Cement of the Universe: A Study on Causation. Clarendon Press, Oxford (1974)

[23] McDermott, J.: Karma and rebirth in early budhism. In: O'Flaherty, W. (ed.) Karma and Rebirth in Classical Indian Traditions, pp. 165–192. Motilal Banarsidass, Delhi (1983)

[24] Mill, J. S.: A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation. Cambridge University Press, Cambridge (2011)

[25] Neufeldt, R. (ed.): Karma and Rebirth. Post Classical Developments. State University of New York Press, Albany (1986)

[26] O'Flaherty, W. (ed.): Karma and Rebirth in Classical Indian Traditions. Motilal Banarsidass, Delhi (1983)

[27] O'Flaherty, W.: Karma and Rebirth in the vedas and puranas. In: O'Flaherty, W. (ed.): Karma and Rebirth in Classical Indian Traditions, pp. 3–37. Motilal Banarsidass, Delhi (1983)

[28] Perrett, R.: Karma and the Problem of Suffering. Sophia **24**, 4–10 (1985)

[29] Radhakrishnan, S. Indian Philosophy, vol. I. Macmillan, New York (1923)

[30] Reichenbach, B.: Law of Karma: A Philosophical Study. University of Hawaii Press, Hawaii (1991)

[31] Rice, H.: Fatalism. In: Zalta, E. (ed.) Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/determinism-causal/ (2014)

[32] Rocher, L.: Karma and rebirth in the dharmasastras. In: O'Flaherty, W. (ed.) Karma and Rebirth in Classical Indian Traditions, pp. 61–89. Motilal Banarsidass, Delhi (1983)

[33] Segerberg, K.: Getting started: beginnings in the logic of action. Studia Logica **51**, 347–358 (1992)

[34] Sharma, A.: Karma, rebirth, and the problem of evil: an interjection in the debate between Whitley Kaufman and Monima Chadha and Nick Trakakis. Philos. East West **58**, 572–575 (2008)

[35] Suryanarayanan, S.: Karma and fatalism. Philos. Quart. (India) **16**, 70–95 (1940)

[36] Vanderveken, D.: Attempt, success and action generation: a logical study of intentional action. In: Vanderveken, D. (ed.) Logic, Thought and Action, pp. 315–342. Springer, Dordrecht (2005)

Ricardo Sousa Silvestre
Philosophy Department
Federal University of Campina Grande
Rua Aprígio Veloso, 882
Campina Grande, PB
CEP 58429-900
Brazil
e-mail: `ricardoss@ufcg.edu.br`