

## THE PRINCIPLE OF DEONTIC REFLEXIVITY AND THE KANTIAN AXIOM

SERGIO GALVAN

### *Introduction*

Some principles of propositional deontic logic present aspects of particular philosophical importance. Two of them are the subject of the present study. The first is the principle of deontic reflexivity  $\bigcirc A \rightarrow A$  (in short the deontic axiom T, i.e. OT) according to which if something ought to be then it is. The second one is the Kantian principle  $\bigcirc A \rightarrow \Diamond A$  (in short OP) according to which ‘ought-to’ entails ‘can’. Despite their similarities, the two principles are profoundly different both from the formal point of view and as regards their truth value within plausible normative interpretations of deontic logic systems.

In the first place, the principle of deontic reflexivity OT is a principle that can be formulated in a language comprising deontic operators alone, while the Kantian axiom OP can only be formulated in a mixed language comprising both deontic and alethic modalities — or alternatively in an alethic language of deontic logic in which the deontic modalities are defined by the alethic modalities and the constant of idealization.

However, the greatest difference between the two principles resides in the fact that the OT principle does not appear as an axiomatic principle or theorem in any plausible system of normative logic — to the extent that its non-presence is viewed by some scholars as a specific feature of deontic systems which delimits their domain with respect to alethic systems or others of different type — whereas the Kantian principle is contained in a large group of mixed or alethic systems of deontic logic.

Firstly, the Kantian axiom is valid in any mixed system (normal or regular) in which the principle of correlation between means and end  $\bigcirc A \wedge \Box(A \rightarrow B) \rightarrow \bigcirc B$  holds. This feature, given the almost unquestionable plausibility of the principle, vouches for the Kantian principle. Secondly, it is valid in all the alethic systems of deontic logic — those endowed with a normal alethic base — in which the axiom Q holds, i.e. the axiom declaring that the state of perfection designated by the idealization constant is possible ( $\Diamond Q$ ). The

plausibility of  $Q$  is therefore a further argument in favour of the validity of the Kantian principle.

Conversely, the impossibility of the OT principle's appearance either as an axiom or as a theorem in a plausible system of normative logic is confirmed by the fact that it is incompatible with Hume's law. This can be demonstrated in the usual formulation — where normative (deontic) propositions cannot be derived from a set of descriptive (non modal alethic) propositions — with regard to a large proportion of the most accredited systems of deontic logic.<sup>1</sup>

It is perhaps surprising that OT is incompatible with Hume's law, which is valid for a number of deontic systems, whereas the Kantian OP axiom (within the context of alethic systems, of course) is not. However, the reason is immediately apparent if one considers the formulation of Hume's law set out above. In it a reference is being made to a set of non-modal descriptive propositions. What does this mean? It means that these are non-deontic formulae; and, moreover, ones without modal operators. The non-modality condition is essential because Hume's law, in the formulation given, no longer holds for a set constituted by modalized formulae. A violation in this sense would be the Kantian principle itself derivable in each of the systems mentioned.

<sup>1</sup> It is not possible here to give precise and exhaustive treatment of the results of non-derivability (of deontic formulae from alethic formulae) concerning the formulation and proof (within particular systems of normative logic) of specific but important exemplifications of Hume's law. However, as support to the arguments set out in the main text some of the most significant of these results should be mentioned, given that they are closely connected with the underderivability of the OT principle under discussion. These results concern the important set of pure deontic systems 0-KD, 0-KD4, 0-KD5, 0-KD45 and the alethic systems (of deontic logic) KQ, K4Q, K5Q, K45Q, KT5Q. However close the equivalence relationship between the systems in the former group and those in the latter — and between K45Q and KT5Q — may be as regards their deontic parts, it is advisable to consider them separately. First, by virtue of their structure, the alethic systems contain a series of bridge-principles (i.e. principles which establish the relationship between modally different propositions) which are of great importance as regards Hume's law. Second, the group of alethic calculuses contain the KT5Q system, which still today is regarded as the most powerful system of normal deontic logic and for this reason deserving of particular attention. And yet the following formulation of Hume's law applies to all these systems (K in short). Let a consistent set M of non-modal alethic formulae be given, and let the so-called condition of the logical non-validity in K of the deontic proposition A hold (i.e.  $\vdash_K A$ ), then A is not derivable in K from M either (i.e.  $M \not\vdash_K A$ ). The relevance of these results to my argument is obvious. From them follows the underderivability of OT in each of the systems mentioned. Suppose, in fact, for the sake of argument, that the OT-scheme is derivable in some K. Then also  $O\neg p \rightarrow \neg p$  would be derivable in K — since it is a particular instance of the principle — and therefore also  $p \rightarrow Pp$  by contraposition. But this is impossible because of the underderivability result for K mentioned above: however much  $\vdash_K Pp$  may hold, one nevertheless has  $p \not\vdash_K Pp$ . More details on this topic can be found in Galvan (1988). See also Schurz (1997).

Does this mean that the Kantian OP principle finds itself in the same situation *vis-à-vis* Hume's law as the OT principle? No it does not, because if the set of descriptive propositions — from which one must establish whether or not the normative propositions can be derived — also contains modal propositions, then Hume's law assumes a different formulation suited to the new situation generated by the presence of bridge-principles like the Kantian OP axiom. There are several of these formulations, but common to them all is the idea that possible candidates for breach of Hume's law are not normative propositions but relevant normative propositions. For example, one can consider only elementary obligations to be relevant normative propositions, and the non-derivability of these is sufficient to state that Hume's law, in its essential nucleus, is still valid.<sup>2</sup>

In conclusion, despite their apparent similarity, the OT and OP principles have very different formal statuses. Their formal diversity is, moreover, the consequence of their differing truth-values in the context of plausible models of normative language. I shall begin with analysis of the OT principle in two different contexts: the context of the teleological explanation of so-called anthropic coincidences, and that of the intentional explanation of human action. I shall then pass to analysis of the Kantian OP principle, this too viewed from a twofold standpoint, theoretical and practical.

### 1. OT and teleological explanation

The problem of teleological explanation has recently attracted renewed scholarly attention because of the impact of the anthropic principle — proposed by a number of cosmologists in order to explain so-called anthropic coincidences — not only on the narrow circle of cosmologists but also on wider contemporary philosophical and cultural debate. From the point of view of my present discussion, it is interesting to ask why the debate aroused by the anthropic principle has revived the age-old problem of teleological explanation. To answer the question it is necessary to reconstruct, albeit in outline form, the terms of the debate.

<sup>2</sup> More precisely, let us assume as valid the formulation of the law for the powerful mixed system  $KT5-OKD-O\Box O-\Box O$  ( $K-\Box O$  in short) equivalent to the alethic system (of deontic logic)  $KT5Q$ . Let a set  $M$  of alethic formulae and the elementary obligation  $OA$  (i.e. where  $A$  is a formula not containing deontic operators) be given. Let the conditions of logical non-validity (i.e.  $\vdash_{K-\Box O} OA$ ) and of non-necessitation (i.e.  $M \vdash_{KT5} \Box A$ ) of the obligation  $OA$  also be valid. Then the law states that the obligation in question is not derivable in  $K-\Box O$  even from formulae belonging to  $M$  (i.e.  $M \vdash_{K-\Box O} OA$ ). Again see Galvan (1988) and Schurz (1997).

According to the most accredited cosmological theories, the existence of the universe in its current form — including the appearance of intelligent life on earth — depends on the occurrence of a series of highly specific and improbable coincidences made possible by imperceptibly fine measures relative to the unfurling of the initial conditions of the universe and to the formation of its fundamental physical constants. Even minimal alterations in these magnitudes would have had such incalculable effects on the evolution of the universe that the birth of intelligent life would have been prevented. And yet the *anthropic principle* (at least in its strong version) has been advanced to explain the occurrence of those infinitesimally improbable events that are anthropic coincidences. The principle thus states that the universe must be such to allow, at a certain moment, the creation of observers within it. To paraphrase Descartes «cogito ergo mundus talis est».<sup>3</sup> Now, given this formulation of the anthropic principle, it is entirely natural that discussion of the principle should also involve the problem of teleological explanation. In my view the principle is amenable to only two interpretations.

The first is that anthropic coincidences are only necessary conditions for the birth of intelligent life, which amounts to saying that the evolution of the cosmos in order for intelligent life to appear could not have occurred without the initial conditions, the fundamental physical constants and the interaction times that go by the name of anthropic coincidences; but this is as far as it goes. However, if the interpretation of the anthropic principle is restricted to these terms, one fails to understand what its explanatory capacity can be, since the principle simply states that without the initial anthropic coincidences life would not exist. But it does not provide any information as to how these initial coincidences came about. The correct interpretation of the principle must therefore be another, and precisely a genuine teleological explanation.

My mention of a *genuine* notion of teleology when dealing with the anthropic principle is no coincidence. Philosophical reflection, in fact, has not always been clear on the matter, for sometimes presented as teleological explanations are particular forms of causal explanation. These are certainly endowed with greater complexity but they are no different in their essential structure from the usual causal explanations or those akin to them (usually referred to as nomological explanations). It is therefore advisable to dwell for a moment on the characteristics that, in my opinion, are distinctive of teleological explanations in the strict sense. Moreover, the issue of the explanatory significance of the anthropic principle has emblematic significance in this respect.

<sup>3</sup> Carter (1974), p. 291.

Generally speaking, two main trends with regard to teleology can be identified within the contemporary philosophical scene. On one side there are those who maintain that the teleological analysis of a specific event (i.e. a fact like the heartbeat) has the principal purpose of *describing* the event itself, but *not of explaining* it. The event in question, according to this interpretation, can be described by saying that it serves some purpose, that it has the function of doing something (in the example, making the blood circulate), whereas its presence — its happening — has a causal explanation like any other event. The other interpretation instead maintains that the teleological analysis of a specific event does not have a simple descriptive purpose, for it is also intended to furnish a proper *etiological* explanation of the event itself. According to this interpretation, therefore, teleological explanations are real, not causal explanations of the happening or of the presence of the event or fact considered, even if several positions can be taken regarding the interpretation of this explanatory rather than causal role.

Belonging to the first group are authors such as Nagel or Hempel,<sup>4</sup> who maintain that teleological explanations can be reformulated in terms of standard nomological (in particular causal) ones. In synthesis: the fact that the heart beats in order to make the blood circulate means (according to these authors) that the beating of the heart in normal conditions is the necessary condition for the blood to circulate. If, in addition to this, we take account of the fact that in the given conditions the heartbeat is also the sufficient condition for the blood to circulate, we may argue with Mackie<sup>5</sup> that the heartbeat is actually the cause of blood circulation. Hence, the teleological explanation according to which "the purpose of the heartbeat is to circulate the blood" is replaced by the causal (i.e. formulated in terms of antecedent conditions and the mechanisms that produce the effect) explanation which says that "the blood circulation is caused by the beat of the heart".

Of course, the point of view related to such approaches does not gainsay either the possibility of ascribing functions to objects or even the use of teleological explanations. Nevertheless, this is allowed because the authors who share this perspective assume that the attributions of functions can be completely reformulated in terms of sufficient and necessary conditions and that the use of teleological explanations is in the last analysis a question of style and not of substance.

This point of view is rejected out of hand by the authors belonging to the second group. In this respect, L. Wright assumes what can be taken to

<sup>4</sup>Nagel (1961), (1968) and Hempel (1965). See also: Beckner (1968) and Braithwaite (1953).

<sup>5</sup>As we know, according to Mackie's theory, the cause is an INUS condition. See Mackie (1974).

be a representative position. He provides the basis for an analysis of the concept of function that acknowledges its non-dispensability in the use of teleological explanations.<sup>6</sup> According to Wright, saying that the function of X (heartbeat) is Z (to make the blood circulate) means: (i) X (the heartbeat) is there because it does Z (circulates the blood); (ii) Z (blood circulation) is a consequence (or result) of X's being there (heartbeat). One notes that the peculiarity of Wright's analysis resides in the first clause, which states that the presence of X is *etiologically* explained by the function performed by X, namely Z. However, problems arise when we try to interpret the nature of this etiological explanatory relationship between Z and X.

First of all, it is rather implausible to interpret the teleological order expressed by the relationship between X and Z as a simple tendency. A tendency, in fact, consists of a propensity to behave, given certain conditions, in a pre-established manner; in other words, it is a particular type of statistical regularity. On this interpretation, therefore, the teleological explanation coincides with some form of common nomological explanation.

Secondly, the teleological explanatory relationship should not be confused with any form of retrocausation. The profound revolution in physics brought about by quantum mechanics, and more recently by the physics of complexity, has obliged numerous physicists to address the question of the subjective or non-subjective nature of temporal direction. If the temporal relation is defined in subjective terms as something that depends on the observer but is not intrinsic to the nature of things, the normal form of causal connection characterized by the relation of temporal priority is flanked by that of retrocausation. This creates room for a new and more comprehensive notion of causal explanation as performed not only *a parte ante*, but also *a parte post*. Furthermore, if one starts from a broad conception of cause, retrocausation is justifiable even regardless of the subjectivity or otherwise of the temporal relation. If a cause is understood as any element whatever of the explanatory antecedent of any nomological explanation, and if the explanatory law is not formulated in terms of temporal priority — which is legitimate if the explanatory scheme being used is not subject to the limits of Hempel's model — it is entirely reasonable to view a cause as something different from the initial conditions, and taken to the extreme, as a state of affairs subsequent to the event-effect. Indeed, some scholars are convinced that the possibility of explanations *a parte post* legitimates teleological explanations. The argument usually adopted is as follows. Explaining a phenomenon in teleological terms, it is said, means explaining it on the basis of the end towards which that phenomenon is directed. But the end is subsequent to the phenomenon

<sup>6</sup>Wright (1976), (1998).

itself, so that explaining it on the basis of its end entails that temporal subsequent events are causally able to influence antecedent ones. Consequently, the teleological explanation is identical with, or at least presupposes, retrocausation.

However, this analysis of the question is not entirely satisfactory. I do not believe that the teleological explanation is identical with, or presupposes, retrocausation. I do so for various reasons which cannot be thoroughly expounded here but can be summarised by saying that there are analyses of the teleological explanation which neither presuppose nor identify themselves with the retrocausation relationship.

Among these analyses, mention must necessarily be made of the evolutionary interpretation. This latter, because it is a looking-back theory, does not presuppose any retrocausation because it justifies the presence of X in order to Z via the evolutionary history which has selected feature X precisely because of the good consequences ensuing from function Z.<sup>7</sup>

Of greatest interest, nevertheless, is the interpretation of the teleological explanatory relationship in terms of *ought to be*.<sup>8</sup> To my mind, this is the most historically rooted modality with which to understand the teleological explanation, and for this reason the most authentic and, as explained above, *genuine* one.

An authentically teleological explanation consists, in fact, in the explanation of certain events in terms of ought-to-be. These events occur because they ought to occur either as an intermediate stage with respect to a final state or immediately as a final state. Whereas, according to the causality thesis, events occur because they are produced by other events — and it is then the task of individual theories of cause to interpret the causal relationship in some or other sense — according to the teleological thesis, events occur because they ought to occur. Now, the ought-to-be of something — i.e. the exigency of this something — is inherent to the situation before the event-end occurs. Hence the teleological explanation — assuming that it works — does not in the least presuppose that the previous situation exerts a causal

<sup>7</sup> There is a large body of literature on the evolutionary interpretation of the notion of function and analysis of the teleological explanation. See, for example, besides the works by L. Wright, Millikan (1984), (1998), and Neander (1998).

<sup>8</sup> M. Bedau puts forward a position similar to this one. In Bedau (1998), in fact, he emphasises the importance of the valuational point of view in teleology, upholding the role performed by values in both analysis of the concept of function and explication of the teleological explanatory relationship. I would like to express my gratitude to the anonymous referee for pointing out the works of this author, previously unknown to me.

influence on the event-end. Of course, one must establish whether the exigency of something is able to explain the occurrence of that something and how this comes about, but the retrocausation relationship is not involved.

Before assessing the plausibility of teleological explanations in the genuine sense just outlined, it is worth stressing that there is a profound difference between the two opposing notions — causal and teleological — of explanation. This difference is most evident in the teleological interpretation of the anthropic principle; the only interpretation, in my view, able to give explanatory capacity to the principle. This point is crucial. It is not at all my intention to assert that the anthropic principle (in its strong version) is epistemologically valid. I merely wish to point out that if one claims explanatory capacity for it, then one must confer the status of teleological principle upon it. When shortly below I show the implausibility of genuine teleological explanations, rejection of the anthropic principle in its strong version will also be deducible from my treatment. However, as we shall see, this does not do away with the task of explaining the givenness of the anthropic conditions; it only means that the explanation should be looked for in a direction other than that indicated by the teleological explanation.

The genuine teleological nature of the anthropic principle has recently been forcefully argued by J. Leslie,<sup>9</sup> who explicitly draws on the neo-Platonic tradition in order to assert that the best explanation for anthropic coincidences lies in the assumption — in keeping with neo-Platonic theory — of the so-called *ethical creationist requirement* (ECR). It is true that neo-Platonic philosophy is favourable to the teleological explanation in its rigorous formulation. In fact, neo-Platonic theory — especially in the most mature and intelligent version of it in Plotinus — has a conception of the Good whereby the Ought-to-be has original explanatory force, in the sense that it brings into Being that which is worthy of Being in that it is Good. The neo-Platonic law of the efficacy of the Ought-to-be does not apply, though, only in the sphere of the One, which imposes itself *qua* pure Ought-to-be, but unfolds in the various forms of teleology that more and more weakly — as one moves away from the One — characterize reality as a whole.<sup>10</sup>

That the core of the neo-Platonic conception has been taken up by Leslie is manifested by the meaning of the ECR which he advances in explanation. What does this principle consist of? I believe that the best analysis is provided by a key passage in Quentin Smith’s detailed review of Leslie’s two works,<sup>11</sup> which states that Leslie’s idea of the ECR as the cause of the

<sup>9</sup> Leslie (1989) and (1990).

<sup>10</sup> *The Enneads*, Treatise 8, §18.

<sup>11</sup> Smith (1994).



universe is best formulated in the following logical, not causal terms: proposition (1) "A universe endowed with intelligent life ought to exist" entails the proposition (2) "A universe endowed with intelligent life exists". This is evidently an explicit argument of genuine teleological type, which, when articulated into its intermediate steps and related to the actual world, becomes: (1) the actual universe is a positive world because it contains forms of intelligent life; (2) that which is positive ought to exist; (3) that which ought to exist, exists; (4) therefore the actual universe exists. Moreover, (5) the existence of the actual world implies the occurrence of anthropic coincidences during the initial instants of the universe; (6) hence it follows that the anthropic coincidences occurred in the initial instants of the universe.

In conclusion, the explanation of the anthropic coincidences offered by the ECR principle is clearly teleological in nature. The coincidences are entailed by the fact that the world with forms of intelligent life ought to exist. However, it is precisely because of the clarity with which the explanation of anthropic coincidences, in terms of ought-to-be, highlights the profound nature of teleological explanations that it is easy to marshal criticisms against it.

Why is a teleological explanation implausible? It is implausible because a teleological explanation of event A explains A in terms of the positivity of A (that is, it states that A is because OA, i.e. that the reason for the existence of A is constituted by OA). And this is the same as saying that  $OA \rightarrow A$  is a logically valid principle (i.e. that A logically follows from OA without the mediation of other principles). Now, as we saw in the first section, the deontic reflexivity principle OT does not appear as either an axiom or a theorem in any of the most accredited systems of deontic logic. Indeed, in these systems the deontic modality of the ought-to-be is entirely independent of the alethic modality of factual actuality.<sup>12</sup>

<sup>12</sup>There is a substantial consonance between this analysis and that suggested by Bedau (p. 275) for the teleological subject. Bedau cites as an example of the most demanding grade of teleology (grade three teleology) a form of teleological scheme similar to the following (where there are no temporal parameters, given that they are irrelevant to the logical argument): C is good; B means (necessary) for C; therefore, B. Now, this argument is correct only if the OT principle is valid. Therefore, the correctness of the argument suggested by Bedau is attributable to the correctness of the OT principle in the same way as it is stated in the main text that genuine teleological explanations are attributable to the use of the OT principle. Nevertheless, I do not share Bedau's conclusions. In fact, he (p. 283) argues: "On the other hand, so grade three teleology in biology does not exist. There are no true full-blooded teleological explanations in biology. Except for the teleology traceable to the mind, the conditions required for grade three explanations are never present in the natural biological world". Yet, a little later (p. 286) when discussing a non mentalistic interpretation of the Aristotelian teleology, he states: "Aristotle's picture of non-mental biological teleology happens not to fit the biological facts, but the picture is nevertheless *possible*. Full grade three teleological explanations *could* be true in non-mental worlds. Thus, the grade three theory does not collapse into mentalism, even if the two are accidentally co-extensive". On the

Of course, this is not to rule out that the reflexivity principle  $OA \rightarrow A$  holds, locally though not universally, for particular propositions  $A$ . It is important to realise, however, that in these cases, even assuming the relevance of  $OA$ , the reason for  $A$  resides not in the simple fact that  $OA$ , but in a more complex fact in determination of which the specific meaning of  $A$  is also of relevance; that is, in a state of affairs in which  $OA$  plays a role only in the context of specific hypotheses about  $A$ . In other words, it is true that we are aware of the occurrence — often regular — of events to which we attribute positivity — if not optimality — in the light of some shared axiological system. However, the existence of these positive or optimal facts cannot be explained on logical grounds — that is, on the basis of the mere fact that they are optimal facts. Instead, assuming that a reason for these facts exists, it should be looked for among reasons of non-logical kind. In short, there are no logical reasons for the existence of certain optimal states of affairs; their occurrence must be, in this case, explained in terms of other reasons.<sup>13</sup>

My arguments regarding the implausibility of the teleological hypotheses, proposed in order to explain valuatively positive facts and the consequent need to develop explanatory hypotheses of another kind, are borne out by the tendency in contemporary debate on the anthropic principle to rely on explanations for anthropic coincidences based on non logical-deontic arguments, like those expressed by the ECR principle. Thus, analysis of the cosmological debate on the anthropic principle is also useful as regards the search for non-teleological explanations for the occurrence of anthropic coincidences.

There are essentially two non-teleological explanatory hypotheses in circulation. The first, by far the most widely current in the scientific community, is rigorously naturalistic, and it is significant from the epistemological point of view because the positivity of the world endowed with intelligent life plays no role in it. I refer to the so-called hypothesis of the *plurality of worlds*. This hypothesis is supported (many specialists in the field report) by

contrary, I contend (and will defend my thesis) that this is *impossible*. Genuine teleological explanations, which is to say those explanations that are not mentalistic of grade three, are logically incorrect and consequently non-valid. As we shall see later, only intentional explanations are valid (they correspond to Bedau's teleological mentalistic explanations). But these have a logical structure which differs from that of genuine teleological explanations.

<sup>13</sup> This is the reason why Bedau's argument mentioned in the previous note, if analysed in evolutionary terms, becomes conclusive. The argumentation was:  $C$  is good;  $B$  is a means (necessary) for  $C$ ; therefore  $B$ . Now, we obtain the conclusion from the two premises plus the principle  $C$  is good implies  $C$ . But, if " $C$  is good" is interpreted as " $C$  is advantageous for the fitness", then it is tautological to obtain " $C$  is good implies  $C$ " and therefore the conclusiveness of the argumentation. Yet, in this case, it is *not because*  $C$  is a value that  $C$  is given, but  $C$  is already given as a fact and it is only by accident that  $C$  happens also to be good. This is enough to prevent the argument from being a genuine teleological explanation.

the theoretical principles of quantum physics, and it states that the universe in which we live is not the only one that exists. Besides our own universe there are many others; indeed, alongside our own universe and physically independent of it, as well as from each other, are all physically possible worlds: that is, those distinguished by initial conditions compatible with the laws of physics at the moment of the Big Bang but nevertheless globally different. The actual world in which we find ourselves, therefore, is only one of many worlds — perhaps infinite in number — that exist but are physically inaccessible. It is thus not surprising that our world came into being and that, for this purpose, anthropic coincidences occurred in the first instants of the universe. Since every possible world came into being, it is entirely natural that this is also the case of the positive world in which we live. However improbable the occurrence of our world as a system isolated from the set of other parallel worlds may seem, it is no longer improbable if we consider it as one of the many possible actualised worlds.<sup>14</sup>

It is not my concern here to examine this hypothesis from the point of view of its truth value. What interests me instead is the peculiarity of its epistemological structure.

First of all, this is an explanatory hypothesis of the existence of our world which hinges not on its positivity but on its necessity. Herein lies the most important difference with respect to the teleological hypothesis based on the neo-Platonic ECR principle. To be sure, it has a feature in common with the latter, for it starts from the acknowledgement that the simple fact is not self-explanatory, and that for this reason it is necessary to move from the typical modality of the conditioned to some typical modality of the non-conditioned. Those who invoke a causal explanation, like the one based on the plurality of parallel worlds, nonetheless believe that this role cannot be performed by the deontic modalities. Although the latter express a specific instance of unconditionality — the ought-to-be can be constitutive and therefore necessary — they are deemed inefficient with regards to the actual world and independent of it. Efficacy is wholly transferred to the alethic modalities, so

<sup>14</sup> It might be argued that the hypothesis of the parallel existence of all worlds is physically implausible. In reality, as suggested by the anonymous referee, the hypothesis of the plurality of the worlds could be interpreted as follows: "There was a (hidden) cosmic evolution of the universes (inflations of vacuum fluctuations) — most of them dying out after a very short time; several of them existing for a long time. In this way, not all but many many of the physically possible universes have already come into existence. This will not make the existence of our universe necessary, but much more probable (or less improbable) than otherwise". Nevertheless, I feel I should say that even under the interpretation of the existence of our universe as more probable rather than necessary, the analysis of the hypothesis of a plurality of worlds carried out here does not lose any of its value. Probability is nothing more than depowered necessity or, which is the same thing, the necessity is nothing more than maximal probability.

that everything that is possible — including the world with its biosphere and noosphere — necessarily urges towards being and, in the end, is.

The second difference is that explanation is wholly transferred to the alethic modalities, in the sense that the positivity of the world does not play an explanatory role, not even a mediated one. In other words, the explanatory hypothesis of the plurality of worlds has nothing to do with a local exemplification of the OT principle made possible by the fact that the antecedent of the principle is part of a network of alethic relationships which mediate its connection with the consequent. This is because the event to be explained — the existence of the actual world — is a consequence of its pure being-possible, in the definition of which its ought-to-be *qua* positive world plays no part.

The second non-teleological explanation of the occurrence of the anthropic coincidences consists, by contrast, in a reformulation of the so-called *argument from design*. Just as until the mid-nineteenth century (until Darwin to be precise) it was claimed that the phenomenon of life could not be explained without the intervention of a Designer — or at least that without such intervention life would be highly unlikely — so today it is claimed, by those who embrace the argument, that only the hypothesis of a Designer can lend likelihood to anthropic coincidences.

Evidently, this explanatory argument is profoundly different from the previous one. The essential difference lies in the fact that the key role in the argument from design is played, not by the positivity of the actual world, nor by its pure being-possible, but by the intention of the Designer. Once again, of course, I am not so much interested in establishing the truth or otherwise of this argument as in comparing its structure with those of the other explanations discussed above. Nonetheless, this purpose obliges me to shed some light on the nature of intentional explanations, given that the argument from design hinges essentially on explanations of this type.

## 2. OT and intentional explanation

Put briefly, intentional explanations rest on the principle that the intention of A produces *ceteris paribus* A (formally:  $IA \rightarrow A$ , where I is an intentional operator and the clause *ceteris paribus* is implied), so that explaining A intentionally means viewing A as the product of an intention regarding A. This principle is obviously akin to the OT principle in that it shares its reflexive form, and it is for this reason that I call it the principle of intentional reflexivity and denote it with IT. However, since this is a practical principle, its validity depends on satisfaction of an important condition of general relevance which must be verified case by case. Without such satisfaction the principle is left suspended, so to speak, in its truth value. Simplifying

somewhat and referring the interested reader to Galvan (1992) for technical details, the meaning of this condition is broadly the following: in a practical setting, the intention to do something does not give rise to concrete action until this intention becomes the intention which, 'all things considered', takes priority over rival intentions and is thus able to engender action or, on a larger scale, fulfilment of the intentioned project. In other words, the practical domain is characterized by a holistic context which differentiates it from the theoretical one in which both the types of explanation discussed above are located. Apart from this contextual difference, the IT principle is isomorphic with OT. This is the reason for the widespread conviction that explanation based on the argument from design is teleological in nature.

However, this conviction is mistaken, for it is an error to believe that the intentional explanation is teleological in the rigorous sense with which the term is used here. It is true — indeed obvious — that a crucial role in the intentional explanation is played by the notion of end, given that the action of the agent is explained in terms of the end that the agent sets out to accomplish. It is also true, that the notion of end is an evaluative one, in the sense that an end is pursued according to its value. However, the relation between the end and the action — that is, between the ought-to-be of the end induced by its value and the action that seeks to accomplish it — is not an immediate one. Instead, it is a relation mediated by the epistemic context of the agent's beliefs in which the ought-to-be of the end is inscribed. In other words, the reason for fulfilment of the end does not lie analytically in its ought-to-be — almost as if the end followed logically from the latter — but only in the ought-to-be as intentioned by the agent. As a consequence, the cause of the action is the intention of the subject and not the ought-to-be of the end, although the latter constitutes the content of the intention. In short, the end is pursued not because it is good but because it is believed to be good by the acting subject.

Significant corroboration of the non-immediacy of the relationship between the ought-to-be of the end and its accomplishment is also provided by analysis of the intentional operators. As we know, formally intentional operators can be defined by means of specific epistemic operators applied to deontic contexts, so that intentional contexts can be viewed as specific deontic contexts in an epistemic framework.<sup>15</sup> For example, once a suitable preferential order (of a single subject) has been established and can be used to define the deontic concept of optimality (with respect to the subject's expectations), the proposition expressing the (subject's) will to achieve A (formally  $WA$ ) can be defined as the (subject's) belief that A in this particular moment is for him the best (most sensible) thing to do (formally  $B(O_tA)$ ).

<sup>15</sup> Galvan (1992).

Also this formal analysis of intentional logic, therefore, shows that intentional explanation does not coincide with teleological explanation.

In conclusion: intentional explanations are not teleological explanations. This means that the former do not rest on the OT principle, which for that matter is deontically invalid. However, intentional explanations are consistent with the IT principle, which displays only structural identity with the deontically invalid OT principle. This isomorphism derives from the fact that intentional operators are definable as deontic operators embedded in the scope of epistemic operators. It does not derive from the subsumption, by virtue of the rule of epistemic necessitation, of the OT principle in a suitable epistemic context. This confirms the practical relevance of IT but not of OT.

### 3. Analysis of the Kantian OP principle

The Kantian principle is plausible from two points of view. It is plausible from the theoretical point of view of an ontology of the ought-to-be able to give an acceptable meaning to the existential notion of sense, and it is plausible from the practical point of view of a rational theory of action.

Why does the Kantian axiom have to do with analysis of the existential concept of sense? First of all, in order to answer this question, it is necessary to clarify the *existential* notion of sense, and to do so it is necessary to make some preliminary points on the *general* notion of sense.

It is frequently asserted that the sense of an action (or of a project) lies in the accordance of the action (or project) with values. There is certainly some truth to this assertion. The action of an agent is *sensate* if it is performed with a view to a fulfilling end. Likewise, a project is *sensate* if it corresponds to the real requirements of the situation in which the project has been undertaken. This *general* notion of sense can also be interpreted in subjectivist terms. It is sufficient for *sensateness*, in this meaning, not to be conceived as a relation between the action (or project) and the end (or the real requirements) inherent in the reality of things, but as a relation between the action (or project) and the end (or the expectations) *qua* intentioned by the subject and therefore deemed by the latter to be worth pursuing, even though it may not be so.

However, there is a second notion of sense presupposed (antecedent), so to speak, by (to) the notion just illustrated. Why should sense — in its first acceptance — coincide with conformity to values? The reason, I believe, is that it is good for values to be realised, and therefore the actions or projects which represent even modally different forms of fulfilment of these values are endowed with sense. It is *sensate*, therefore, that the positive should come into being, although it is possible that this may not happen, as we saw in the previous section. Yet if it is *sensate* that the positive should exist, it is

*a fortiori* sense that it should be possible for the positive to exist, i.e. that the conditions for its realization should exist.

It is this second notion of sense that I wish to discuss, the *existential* one. Here sense does not lie in the simple convenience that the positive should exist but in that it should be possible. This is presupposed by the former in the same way as the possible is presupposed by the actual. Moreover, this notion is existentially more important than the former one because, in this acceptance, it is not whatever conforms with the value — i.e. the state of affairs that satisfies the value requirement — that is sense but the existence of the value itself. In other words, this second notion of sense allows us to deem nonsensical a situation characterized by the inaccessibility of a certain value. It does so for the simple reason that the value preserves its requirement of fulfilment intact even in this situation, and yet the value cannot be fulfilled. Nonsensical, for example, is the situation created by a legislator who legitimately issues laws that cannot be obeyed, because the promulgation of these laws entails the obligation of obeying them and yet this is impossible.

This second notion of sense was first defined by Kant in his practical philosophy, specifically when he posited freedom as the necessary condition for being responsible with regard to moral duty: you ought to because you can choose. However, it is a notion with a meaning and a range that go beyond Kant's conception. In particular, it acquires the entirety of its existential meaning within the context of an ontology of value which views it as rooted in an exigency of being, and therefore as the reason for the fact that a state of affairs is directed towards an end. In the light of this interpretation of value, sensateness (or otherwise) concerns the existence itself of end-directedness. If a situation is directed towards a certain end but is marked by the impossibility of fulfilling this end, then it is a locus of non-sense in which an exigency is set that cannot be satisfied. Conversely, a situation in which the conditions exist whereby the limit present within it can be overcome is a situation endowed with sense. Of course, this manner of conceiving the relationship between being and ought-to-be presupposes an axiological option based on a realistic ontology, thorough discussion of which is not possible here. Suffice it to say, that the relationship can be analysed in the modern terms of a supervenience relation where the axiological (and deontic) predicates are implicated by corresponding ontic predicates (i.e. concerning the dimension of being) so that they can be defined as 'supervenient' on the latter. Now, given this analysis, the meaning of the assertion that a situation (or possible world) is capable of inducing an ought-to-be is obvious. A situation induces an ought-to-be because the ontic predicates defining it declare its limit with respect to the state of fulfilment that they require. In other words, just as when certain characteristics are given a situation can be judged positive as if the axiological quality of the positivity supervenes on those characteristics, so one can say that the situation has an axiological

quality of opposite sign when these characteristics are lacking, i.e. when the situation is such to require them but at the same time has been determined in such a way that it cannot possess them. Likewise it is evident why the Kantian axiom is the law governing the institution of end-directed situations satisfying the relative sense condition. These situations satisfy this condition only if the ought-to-be that they induce can be realized — that is, only if the end-directedness that they express is not endlessly thwarted.

I have said that the root of non-sense lies in an unachievable ought-to-be. I have also said that this assertion does not concern solely the moral meaning of the ought-to-be. All forms of ought-to-be are subject to the law of possibility as the condition for their sense. Yet a particular form of ought-to-be is whatever the agent deems to be the best course of action at a given moment; in other words, whatever concretely and globally is first in the agent’s order of preferences, so that it induces him to will and in consequence to act. The Kantian axiom thus acquires specific relevance to a theory of will and action that starts from the idea that the agent always wills whatever according to his overall order of preferences appears at that moment to be optimal.

And this, in fact, is the most significant acceptance of the Kantian principle from the practical point of view. Given the interpretation of the ought-to-be expressed by the principle in terms of a preferential optimality, the principle can be defined — by appropriately inserting the deontic context expressed by the order of preferences into the epistemic context of the agent’s beliefs — as the basic principle of practical rationality **SP**; the principle that intention (will) entails the belief that the content intentioned (by the will) is achievable — that is to say, possible. In fact, using  $O_t$  to denote the deontic operator of the obligation defined as optimality, **OP** in the above sense becomes:

$$O_t A \rightarrow \Diamond A$$

from which, by the rule of epistemic necessitation one has:

$$B(O_t A) \rightarrow B(\Diamond A)$$

and therefore, by definition of the intentional operator of will **W** (something is willed when it is deemed optimal with respect to the order of preferences at a particular moment), the basic principle of practical rationality or the sense postulate **SP** is:

$$WA \rightarrow B(\Diamond A)$$

Just as the **IT** principle results from the transposition of **OT** from the deontic level (where it is invalid) to the intentional level (where instead it is valid), so the **SP** principle results from the corresponding transposition of



the Kantian OP principle. There is, however, a substantial difference between the two principles. In fact, although the former is clearly analogous to OT, it is novel with respect to it because it does not derive from the insertion of OT in the epistemic context of the subject's beliefs. This is owing to the simple fact that OT is invalid. Conversely, the SP principle is some sort of logical consequence of the Kantian principle, for it derives from it through the straightforward application of the rule of epistemic necessitation, and also acquires its truth value from it.

The interest of OP from the practical point of view is that it constitutes a reason at once explanatory and justificatory for the fact that the SP postulate of the sensateness of will is a cardinal principle of rational action. It is difficult to deny, in fact, that human will — like any other form of intentional activity — behaves in conformity with the sense postulate SP. Will is by necessity always accompanied by the conviction that the object of will is at least possible in the long term. If, then, one of the subject's beliefs is that OP, then the derivability of SP from OP explains why the set of the subject's intentions is grounded on the SP sense postulate. With respect to the subject, finally, OP also figures as a clear justificatory reason for the same postulate.

The cardinal principle of practical rationality mentioned in the previous section is, of course, neutral with respect to the theory of good embraced by the individual subject. What appears optimal to the subject may also be in his exclusive interest: his happiness. However, it may be that the subject's order of preferences also comprises an objective axiological order — or at least commitment to impartial principles like the principle of universalizability — such to constrain him to a behaviour not always dictated by the personal exigency of happiness. It is also possible, however, that the subject may embrace a pluralistic theory of good which deems it inadmissible to fulfil certain values and not others — in particular, fulfil moral duty to the detriment of happiness — but requires the simultaneous and harmonious fulfilment of them all, including the value of happiness as consequent on fulfilment of the other values. Yet in cases such as these, where the subject's order of preferences is composed so that none of them predominates over the others but all are taken into account, what will be the consequences of commitment to the basic principle of practical rationality SP?

In my view, the principle states that one fundamental necessity of rational will is to pursue *globally* sensate projects. Indeed, one may ask, what is a globally sensate project if not a project perceived by the subject as not only worth pursuing ideally by virtue of the values that it represents, but also as one compatible with operational conditions and therefore susceptible to fulfilment in respect of all the subject's expectations, subjective ones as well? In other words, a project is rationally willed only if it is deemed sensate, and it is deemed sensate only if it corresponds to the subject's expectations

as a whole, which comprise, besides the intentioned values, also the belief that these values can be fulfilled together with satisfaction of the paramount exigency of happiness.

This contextualization of the Kantian principle has numerous implications for analysis of the meaning of rational action and its presuppositions. If an at least necessary condition for rational action is the principle that the project pursued is to be deemed globally sensible — that is, able to fulfil all the subject's expectations — then it is not rational to pursue a project in which moral values predominate over the exigency of happiness to the extent that the subject must sacrifice his life — in favour of other subjects, for example — in the name of these moral values. It is not rational because a project of this kind is nonsensical and thus conceived cannot become the real object of intention. It could be objected, perhaps, that such a radical opposition between the two instances could never come about in reality, so that the occasion to behave irrationally will never arise. Unfortunately, however, experience shows that cases of this radical opposition do occur, and only one of them suffices to pose the problem of principle.

In effect, the purpose of the Kantian postulates of practical reason is to avert this problematic situation. The solution is achieved by requiring that the absolute fulfilment of moral duty — i.e. the realization of a project in which moral values predominate — is made sensible by the intervention of God, who guarantees justice and therefore rewards personal merits in the life hereafter. Of course, in Kantian theory the existence of God as guarantor of the moral order and as overseer of the real order is simply postulated, which highlights the cleavage in Kant's thought between practical and theoretical reason. Nevertheless, besides the necessary direction of the outcome — that is to say, the shift to transcendence — the problematic root that generated the whole of Kant's argument seems undeniable. Rational will must, by necessity, pursue the moral good only in the conviction that in the long run the expectation of happiness will not be frustrated.

Università Cattolica del Sacro Cuore,  
Largo Gemelli, 1

I-20123 Milano, Italy

E-mail: sergio.galvan@unicatt.it

#### REFERENCES

- Allen, C., Bekoff, M. and Lauder, G. (eds.) (1998): *Nature's Purposes. Analyses of Function and Design in Biology*, Cambridge, Massachusetts: The MIT Press.

- Beckner, M. (1968): *The Biological Way of Thought*, Berkeley: University of California Press.
- Bedau, M. (1998): 'Where's the Good in Teleology?', Allen, C., Bekoff, M. and Lauder, G. (eds.) (1998): 261–291.
- Braithwaite, R.B. (1953), *Scientific Explanation*, Cambridge: Cambridge University Press.
- Carter, B. (1974): 'Large Number Coincidences and the Anthropic Principle in Cosmology', Longair, M.S. (ed.), *Confrontation of Cosmological Theories with Observation*, Dordrecht: Kluwer Academic Publishers.
- Chellas, B.F. (1980): *Modal Logic. An Introduction*, Cambridge: Cambridge University Press.
- Galvan, S. (1988): 'Underivability Results in Mixed Systems of Monadic Deontic Logic', *Logique et Analyse* 121–122: 45–68.
- Galvan, S. (1992): 'Inferenza pratica e teoria del volere', Galvan, S. (ed.), *Forme di razionalità pratica*, Milan: Franco Angeli: 189–211.
- Hempel, C.G. (1965): 'The Logic of Functional Analysis', Hempel, C.G. (1965): *Aspects of Scientific Explanation*, New York: Free Press.
- Leslie, J. (1989): *Universes*, London: Routledge.
- Leslie, J. (1990) (ed.): *Physical Cosmology and Philosophy*, New York: Macmillan Publishing Company.
- Mackie, J.L. (1974), *The Cement of the Universe. A Study of Causation*, Oxford: Oxford University Press.
- Millikan, R.G. (1984): *Language, Thought, and Other Biological Categories*, Cambridge, Massachussets: The MIT Press.
- Millikan, R.G. (1998): 'In Defence of Proper Functions', Allen, C., Bekoff, M. and Lauder, G. (eds.) (1998): 295–312.
- Nagel, E. (1961): *The Structure of Science*, New York: Harcourt, Brace and World.
- Nagel, E. (1998): 'Teleology Revisited', Allen, C., Bekoff, M. and Lauder, G. (eds.) (1998): 197–240.
- Neander, K. (1998): 'Functions as Selected Effects: The Conceptual Analyst's Defence', Allen, C., Bekoff, M. and Lauder, G. (eds.) (1998): 313–333.
- Schurz, G. (1997): *The Is-Ought Problem. An Investigation in Philosophical Logic*, Dordrecht/Boston/London: Kluwer Academic Publishers.
- Smith, Q. (1994): 'John Leslie, *Universes* (London: Routledge, 1989), vii + 228 and John Leslie (ed.), *Physical Cosmology and Philosophy* (New York: Macmillan Publishing Co., 1990), viii + 227', *Noûs* 28: 262–269.
- Wright, L. (1976), *Teleological Explanation*, Berkeley: University of California Press.
- Wright, L. (1998): 'Functions', Allen, C., Bekoff, M. and Lauder, G. (eds.) (1998): 51–78.