ABSTRACT: This paper consists in an analysis of two explanatory models of human behaviour which play a prominent role in the contemporary literature on human action. The first model – the causalist – aims at explaining action in terms of causes and general laws. The second model – the intentionalist – explains human action in terms of intentions and practical syllogism. The difficulties of both models are presented and in the last part of the essay we propose one alternative model, based on the notion of retroduction.

KEY-WORDS: Explanation; covering-law model; human action; causalism; intentionalism; retroduction.

As we all know human behaviour can be described and classified in a several ways. We speak about activities, actions, achievements, habitual and automated behaviour, reflexes and so on. It is plausible to think that our explanations of behaviour are strongly linked with our ways of describing behaviour. It could also be argued that one’s general views concerning the nature of man and the sources of man’s activity strongly affect one’s way of conceptualizing and explaining behaviour.

In this article we shall concentrate on behaviour as action and on some of the several ways one can explain actions. Our strategy will be threefold: in the first part we will try to sketch an approach to the explanation of action which we will call “the causal approach to the explanation of action” or “causalism”. We shall illustrate this approach by focusing on a specific explanatory model, the so-called nomological-deductive model. The version of this model we wish to study in detail is Hempel’s one. Its importance in the philosophical literature on explanation is almost self-evident and the attempt to extend this model to the explanation of items of human behaviour would certainly lead us to review a lot of questions which are embodied in the explanatory task, such as causation, determinism, laws, etc.

In the second part of this paper we shall present another explanatory model, which we will call “the intentionalist approach to the explanation of human action” or “intentionalism”. A version of the intentionalist model of explanation can be found in G. H. von Wright’s, Explanation and Understanding. The reason to present the intentionalist view according to von Wright’s approach resides in its fruitfulness to raise a set of questions with which we are particularly in agreement here. We are referring to von Wright’s criticism of the Logical Connection Argument. The Aristotelian idea of “practical syllogism”, re-stated by G. Anscombe and Von Wright will also prove fruitful to the explanatory scheme to which we are going to devote the third part of this paper.

* Departamento de Filosofia – Faculdade de Educação, Filosofia, Ciências Sociais e da Documentação – UNESP – 17500 – Marília – SP.
Causalist and intentionalist models of explanation are presented as opposite. The reason of this opposition can be grasped if we look backwards retracing the different philosophical background from which they emerge. For reasons of space we shall not attempt to characterize those historical roots.

The third part is devoted to the design of an alternative explanatory frame for human action. It emerges from the discussion of some theoretical flaws which we have identified in the exposition of the former explanatory models, and constitutes our own contribution to the issue. The alternative model we are going to present is based on the idea that causalism and intentionalism can be assembled to form a wider explanatory model, which uses the notion of "retroduction" in a very particular way. In this sense, the confrontation of intentionalist/causalist models will prove useful to the drawing of the new approach.

There are some key concepts which are related to the discussion of explanatory models of action. Among them there is the concept of action itself, and in order to clarify it one would have to accomplish the task of characterizing what is to be understood as an action, its difference from bodily movements etc. We shall not discuss those preliminary concepts, taking the risk of setting up our discussion on the basis of the common sense idea of characterizing action primarily in terms of antecedents and consequences. Another concept which would deserve some attention is the idea of explanation itself. Its detailed discussion would lead us far from our objectives. Anyway, we will take for granted that explanation can be understood as a partial or complete answer to queries, and, in special, to why-queries.

1 – A short characterization of Hempel’s nomological-deductive model would run as follows: There are some events or phenomena which are called "E" and are described by a set of statements called explanandum. The explanandum is deduced from a set of other statements named explanans. This latter is a premise cluster formed by general laws and statements describing particular events or conditions which are antecedent or simultaneous to the event "E" to be explained.

The following scheme represents the nomological-deductive model:

<table>
<thead>
<tr>
<th>Deduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Explanans C1C2Cn – statements describing antecedent/simultaneous</td>
</tr>
<tr>
<td>conditions.</td>
</tr>
<tr>
<td>L1L2Ln – General Laws</td>
</tr>
<tr>
<td>(2) Explanandum E – statement describing the event/phenomenon to be</td>
</tr>
<tr>
<td>explained.</td>
</tr>
</tbody>
</table>

The model is well-known throughout the philosophical literature, and, since that there is no lack of good expositions and critical examinations we shall not portray its details any longer. The features which are worth emphasizing to our purposes here are: 1) The laws L1L2Ln forming the explanans should be universal statements, otherwise it would be impossible to frame the model as a deduction, and, 2) According to Hempel’s model prediction is just the "reverse side" of explanation, which means, that explaining "E" corresponds to predicting "E" after its occurrence.
Furthermore, Hempel claimed that the nomological-deductive model (ND or HD) could be viewed as a version of causal explanation itself. He says:

"Causal explanation is a special type of deductive nomological explanation; for a certain event or set of events can be said to have caused a specified "effect" only if there are general laws connecting the former with the latter in such a way that, given a description of the antecedent events, the occurrence of the effect can be deduced with the help of laws". (4, p. 300-301)

The same aspect is stressed by von Wright, who emphasized that,

"The covering law model was originally thought of as a generalization of ideas associated with causal explanation. The specific problems about causation seemed to many to have lost their urgency because of this widening of the conceptual horizon - just as Russell had thought that causation had become philosophically uninteresting because it might be subsumed under the broader category of functional relationship. But this is a mistake" (5, p. 37)

The critical examination provided by von Wright in his Explanation and Understanding approximates the DN-model to the idea of Human causation, which is tantamount to saying that causality is to be understood in terms of regularity and contingency.

The idea of extending this model to the explanation of human action is not explicitly stated by Hempel. Nevertheless, there are several passages in Aspects of Scientific Explanation in support of the view that the model is conceived as encompassing also human behaviour. Those passages criticize motivational accounts of human action, and attempt to prove that any teleological feature of agency can be framed into a classical nomological-deductive explanation (See 4, p. 469 onwards). The criticism aims at demonstrating that motives can function as causes and that there is no peculiarity in human action that would require a different explanatory account. The intent of formation is casually linked to behaviour production, and it is possible to provide an explanation of action with the help of general laws, even though we have to refer to motives, desires, intentions etc.

D. Davidson, in his article "Hempel on Explaining Action" (Note A) refers to a paper delivered by Hempel at the Eastern Division of the American Philosophical Association in the early sixties. The title was "Rational Action", and, in a nutshell, it shows that Hempel's position has not changed ever since.

Moreover, the account of historical explanation provided in "The Function of General Laws in History" also seems to stress the idea that the model has a very wide range of applications, including the action of groups in the past, which should be understood in terms of regularities. On the basis of this account, Hempel draws a dichotomy between pseudo and genuine explanations.

The way to criticize Hempel's approach would run as follows: 1) It is very questionable whether general laws can be understood on the basis of the ideas of regularity and contingency. Besides that, it can be argued that Hempel did not provide criteria to distinguish genuine laws from accidental generalizations (Note B). 2) The second aspect which may receive criticism is the idea that human action can be explained causally. One could argue that explanation of behaviour is to be searched in the light of motives and intention formation. These two criticisms are the guidelines of von Wright's account of the explanation of human action which we shall focus on in the next pages, in a very oversimplified presentation.
16

2 – Von Wright's Explanation and Understanding starts by reviewing the notion of causation and its relationship to the idea of general laws. The approximation of the DN-model to the idea of Humean causation leads von Wright to a critical survey of the criteria of lawlikeness.

Causation cannot be viewed in terms of regularity and contingency, and, unless we introduce the idea of a "natural necessity" it would be almost impossible to conceive causation. What confers on observed regularities the character of causal or nomic connections is the possibility of subjecting cause-factors to experimental test by interfering with the "natural" course of events provided by "natural necessity". In an important sense, the causal relation can be said to be dependent upon the concept of human action. This dependence has to do with the way causal relations are established and distinguished from accidental regularities.

The criticism of the conception of general laws adopted in the DN-model, which is in principle indistinguishable from an accidental regularity, leads von Wright to question the applicability of causal categories to the explanation of human action. Explanation of action is often and typically, even if not exclusively, given in terms of intentions, motives and reasons. We also say that in acting an agent is aiming at something, an end of action, and we explain his conduct in terms of his aims and ends. Such explanations are teleological.

The conceptual scheme developed by von Wright to explain human action is based on the idea of practical syllogism. The practical syllogism basically consists of two premises and a conclusion. The first premise of this pattern of thought is a statement about an agent's intention to achieve a certain end. The second premise is a statement about what he believes to be required of him to do in order to achieve this aim. The conclusion is a statement which, roughly, says that the agent does or proceeds to do what is required of him in the second premise.

The basic idea can be represented in the following scheme of von Wright's:

P1 – A intends to bring about a certain end E.
P2 – A considers that unless he does action "x" he cannot bring about E.

C – A sets himself to do "x".

This scheme represents the basic form of the practical syllogism. Its first premise speaks about A's intention to do something which has as its consequence the fact that E is realized or brought about. We could also have said, alternatively and more specifically, that A intends to do "x" such that the result of the action "x" is the realization of the event (or state of affairs) E.

The view that practical syllogisms, such as those represented above, are logically conclusive has been discussed extensively by von Wright. He has also emphasized the pivotal role of the practical syllogism in "understanding" behaviour as intentional action. Von Wright claimed that the "tie" between the premises and the conclusion of a practical syllogism is logical.

The basic idea of the argument is this, if one can show that it is not possible to verify or falsify the premises and the conclusion independently of each other, then the tie is logical rather than empirical or causal. Since the relation between premises and conclusion is not one between logically independent terms one cannot say that they hold a causal relation. In a causal relation, cause and effect are logically independent.

This "oblique" demonstration allows von Wright to re-state the so-called Logical Connection Argument. According to the Logical Connection Argument nothing can be considered a motive...
unless logically connected with what is wanted. But the way in which the earlier versions of the Logical Connection Argument was presented is considered by von Wright as faulty or else unconvincing. For intention formation does not necessarily mean that an agent will certainly act accordingly. The reason may "drop out" before the action has taken place, i.e., the agent may change his mind. Or the reason may continue to be there for him to act — and the agent may try but fail to accomplish the action. Or something may happen which prevents him.

When both the action and the reason for it, i.e., the reason why an agent acted are there, then their connection is conceptually, logical and not causal. But the necessity embodied in the practical syllogism is not the necessity which is immune to failure. The failure to materialize an action makes the practical syllogism an instrument to explain human action "ex post actu". The earlier version of the Logical Connection Argument are defective according to this view for they regarded the conceptual, connection between motives and actions as deterministic. 3

— So far we have presented the two explanatory models. To sum up, it would be necessary to say that the explanation of a sequence of behaviours, according to the causalist view would consist in its subsumption under a set of general laws. The idea embodied in the model is that those laws express a causal regularity.

According to the intentionalist view, however, the explanatory task would consist of establishing a set of premises from which items of intentional behaviour could be derived. Those premises make up a practical inference conceptually sound but they do not entail the effectiveness of the agent's doing. So, the agent can be prevented from executing his action although we can ascribe him some process of intention formation. Thus, the practical syllogism is a reconstruction, but, according to von Wright's account, an "ex post actu" reconstruction. Another point to be stressed is that we can proceed to such reconstruction in several ways, matching different premises whose conclusion would be the performed action. In this case, the criteria for reconstruction is related to patterns of current intentional behaviour, whose performance allows the external observer to associate them to a lot of intentions. The weight of habits and of cultural contexts plays an important role in the task of reconstructing the agent's internal states and of setting up the premises as well.

The causative approach has received several criticisms. Besides the difficulty of establishing clear-cut criteria of lawlikeness (stressed by von Wright in the first part of Explanation and Understanding) there are those who have noted to the fact that causalism rules an important aspect out of human action: its teleological feature, i.e., its reference to intentions and beliefs. The causalist, however, could simply argue that the explanatory model is functioning as a kind of ontological arbiter, and one could also claim that there really is not such a peculiarity in human behaviour. Hence no description of the agent's ends would be needed.

Such a view of human behaviour is indeed possible but it does not seem that detecting regularities could exhaust the explanatory task. As long as we restrict ourselves to observing regular sequences of behaviour it cannot be said what the agent is performing. The difficulty is like the one the behaviourist has encountered: while it is clear that an experimenter can predict rate of learning from the initial condition of his mazes and the experience history of his animals, how does he specify just what is learned? (Note C) The detection of a regularity does not explain human action as well unless by including a reference to the ordinary intentional vocabulary where terms such as "believe", "know", "want" play a prominent role. The description of some goal directedness to human agency is the guarantee of an intelligible explanation and the distinctive feature of action from skeletal motion.
The argument seems to be sound, but it could still be said that this is an argument from the intentionalist point of view. The trouble with causalism has to be found out in a different way. What is troublesome is the existence of anomalous instances of behaviour that cannot be framed into regularities, and so cannot be explained in terms of series or laws.

It could be said that any model or theory has to get along with exceptions, and that they do not constitute a sufficient reason to abandon a theory or explanatory model. It can be argued that a theory is not necessarily superseded when we discover exceptions. They could also bear such exceptions which can be statistically understood sometimes. Anyway we would have to raise the question why an agent changes the direction of his agency and this would lead us to the conclusion that no account can be given unless by reference to a change in the agent’s internal states. If the causalist is to recognize the existence of such anomalies he has also to recognize that it is not possible to hold that there is a causal link between intention and action. At least this causal link cannot be understood as a universal law.

Accounting for anomalies in terms of a probabilistic connection between intentions and actions could complicate such an approach: there is no way to assess the truth or falsehood of probabilistic laws. A low-rate probability can be crucial for human agency, specially in the case of anomalous action, and so it turns up to be difficult to derive possible universal laws which would be required to preserve the deductive feature of the model.

The trouble with anomalies seems to be a problematic feature shared by both intentionalism and causalism. A causalist point of view seems to arise from the idea that human action exhibits some kind of regularity which would, in principle, expand to some kind of law or invariance. The intentionalist, on the other hand, does not seem to have means to account for this latter feature of human agency. According to a model such as von Wright’s, regularity can only be explained as a result of a coincidental intention formation conceptually linked to a certain set of behaviours. It is also possible to suppose that both models emerge from a common data basis where behaviour is normally couched, exhibiting either regularity or anomalousness. From that basis it is possible either to infer the existence of a causal connection, or the presence of intentions conceptually connected to actions. However, the data obtained do not support the ontological claim that there is a causal link tying items of behaviour, nor the claim that there is not such a tie and that action stands for intention formation conceptually tied to behaviour.

Intentionalism has also received several criticisms by authors such as D. Davidson, R. Tuomela and others which have emphasized its troubles, specially the problems of von Wright’s version of motivational explanation. However, we shall center here on the difficulties the model presents in coping with anomalous instances of action.

Let us suppose we are to explain on the intentionalist account, the following action: I have the intention to open that door and leave the room. The intentionalist would say that the action of opening the door is not independent of my intention formation, and, so, they cannot be causally explained. However, it can come about that I have the intention to open the door, but, as a matter of fact, I did not open the door. The question that would arise immediately would be: why did I not open that door, if I had the intention to? The causalist could immediately argue: if you had the intention, you would certainly have to perform the corresponding action. If you did not, you have to suppose the presence of some intervenient factor. The explanation of this disjunction between intention and action cannot be searched far from the idea of causation. However, as we have pointed above, the question why the intention lacked its causal efficacy remains equally unanswered, and the causalist explanation would also fail.

The Wrighean intentionalist would argue that such an example does not constitute a failure of his model. It could be seen as a failure of the traditional version of the Logical Connection Argument whose defectiveness has been already stressed by von Wright. The idea, as we have already described, is that we have to dissociate the conclusiveness of practical inference from the actualization of the agent’s doing. The latter can only be recovered and associated to intention formation “ex post actu”. But is that not question-begging since it obscures the necessity to suppose a causal intervenient factor to reach a suitable explanation? The appeal to the idea of intervenient factors of other sort would not help the Wrighean intentionalist (5, Chapter 3) for in the majority of cases, the intervenient factor alluded to by von Wright allows one to explain only why such action was postponed and so on. That does not explain why an agent changed the course of his action, unless by causal means.

The disadvantage of the intentionalist account seems to become more evident when we consider the “reverse side” of the question. As we have noticed above, intentionalism seems unable to explain the recurrence of actions unless by stating the coincidence of the same intention formation. But why does an agent keep on forming the same intentions and acting up to them? The causalist approach seems more adequate to deal with such an issue: at least in the domain of causal explanation we can state that a particular event “a” is followed by another event “b” by the invocation of some more general causal law to the effect that all events of type A (which includes a) are followed by events of type B (which includes b) and this law may be explained in turn by being subsumed under or deduced from still wider laws. In the intentionalist approach, on the other hand, the sequences of events are so characterized that the occurrence of a particular consequent action is explained by the occurrence of a particular antecedent, say, a perception, belief or intention; and there is no room for the question of why this consequent should follow this antecedent, and hence no room for any general law “explaining” the sequence.

The assessment of the models we have begun here would deserve a much more detailed account. The common difficulty shared by the two models seems to reside in accounting for the anomalies. We have presupposed that an anomaly is just an exception to a series. That means that it corresponds to some action that was expected but not performed. The expectancy can be based on either a process of intention formation or on the expectation we develop toward a regular sequence.

We shall suggest next how to develop a solution for such a problem-bundle. An alternative to the two explanatory models can emerge from the assessment of those theoretical issues we have considered so far. Intentionalism and causalism have both virtues and difficulties that may be reconciled by devising an alternative explanatory model. To begin with, we shall review the status of laws/predictions in the explanation of action and re-examine the role of probabilistic explanation which we have purposively set aside so far. The causal efficacy of motives and intentions, paradoxically entailed by the intentionalist view will be used as an indicator to a new way of conceiving the role of laws in the explanation of behaviour. Furthermore, we shall attempt to re-state the role of intentions in the description of action as well as in coping with anomalies.

4 Aristotle once said that “since scientific knowledge involves demonstration, but there is no demonstration of things whose first principles are variable, and since it is impossible to deliberate about things that are of necessity practical wisdom cannot be scientific knowledge nor art” (1, 1140a–1140b). That is to say that action theory cannot be expected to develop in ways exactly parallel to the physical sciences. But this is not to say that no explanatory account can be developed.
The trouble with anomalies does not constitute a serious handicap to causalism if some of its thesis are re-arranged. Nor does it affect the view according to which intentions can exhibit a causal efficacy which has proved its importance to cope with the regularity of action performance. The trouble with anomalies bears on the attempt to expand some universal laws governing the connection between intentions and actions. As a consequence, however, the symmetry between explanation and prediction held by Hempel would have to be abandoned. It represents a failure to bring action as described in the vocabulary of action theory under strict deterministic laws. Behaviour would certainly resist the incorporation into such a closed deterministic system.

Nevertheless there is still another problem to be dealt with: if there is no such a universal tie connecting intentions and actions, how to derive universal empirical generalizations to provide the DN-model with its deductive feature? We will answer this latter question by assuming that intention exhibits a causal efficacy as a preliminary step to develop our alternative explanatory model.

The assumption that causal dependence can be reconciled with the teleological feature of behaviour has already been stressed by D. Davidson (Note D) and formerly by Kant. The idea, as developed by Davidson in the domain of the mental, does not correspond exactly to the one we want to emphasize here. We will just state that anomalousness does not necessarily entail a rupture with a nomological net. Anomaly, understood as a change in the intention formation process, can be regarded as an interruption or substitution of a causal chain by another one. Moreover, new sequences can be voluntarily set up by the agents, for instance, as a result of decision making. However, in neither case we have to abandon the idea that intention exhibits a causal efficacy. That means that an anomalous instance of behaviour has not necessarily to correspond to the "failure" of the causal linkage between intention and action, but rather to its substitution by another intention which causes a different manifest behaviour. What is to be abandoned is not the causalism, but the monocausalist perspective. Prominent intentions have to be statistically understood, maybe as a function of changes in environmental circumstances. However, in this account there is no complicating factor in deriving universal laws as in the classical version of a nomological "probabilistic" model.

Since we have stated that intentions and motivations play a causal role in behaviour production, we have got the tools to suggest an alternative explanatory model. The question is now how to match the intentional feature of behaviour, causation and anomalousness altogether. Our suggestion is that it can be done by conjoint two available models: the nomological deductive one (Hempel’s version) and the "retroductive".

The DN-model is rather to be understood as explaining sequences of actions. The "explananda" of this model are to be understood as items of intentional behaviour. The model bears on the links between intention formation and behaviour that can be approached as a causal result. At this stage, the regularity exhibited by sequences of behaviour can be accounted for in terms of general laws. The generality of laws can support the deductive character of the model at this stage. On the other hand, the intentional description of behaviour is preserved.

Anomalies challenge the DN-model, and request another type of reasoning. Anomalies can only be identified post festum and the DN approach can also prove fruitful at this stage. With the help of the DN account one can set out the "logical expectations" of a given sequence and hence highlight any deviation from these expectations.

The second stage will consist of using retroductive reasoning. The anomalous character of
behaviour is related to some change in intention formation. So, the task of retroductive reasoning will be reconstructing the way between the observed behaviour and the projective intention which has caused the deviation.

It has been argued that there is no such a thing as retroduction. One could argue that the difference between the deductive model and the retroductive one is just psychological. We disagree with such an approach. However, discussion concerning the nature of retroductive reasoning has raised an extended controversy which we shall not reproduce here. We will define retroduction by comparing it with the task of tying pearls in order to form a necklace. This is quite different from a deductive argument. The "working backwards" activity may be viewed as a traveler's puzzle when he asks "here am I, river to the left, mountains to the right, canyon ahead, where do I go from there?"

The task of retroductive reasoning is to form a hypothesis cluster which accounts for the anomaly. Since the anomaly can be conjoined with some hypothesis cluster the latter turns up to function as a premise or "explanans" of the deductive explanatory account. The DN procedure and the retroductive procedure are therefore intimately linked. They are two stages of the three-stage explanatory task which we are suggesting here, consisting in: a) accounting for the regularities in terms of causal laws and deduction. Regularities here are the product of the causal efficacy of intentions, allowing the identification of, the anomalies which do not square with the facts; b) raising some new hypotheses concerning the change. New hypotheses turn up to be premises, entailing the anomaly; c) the new premises can be understood in terms of the orthodox DN account.

In the specific realm of the explanation of action the second stage corresponds to ascribing some intention formation which could be obtained by "working backwards" from an anomalous instance of behaviour. Since the behaviour can be "couched" in this new premise cluster we can account for the change of the agent's internal state. This latter turns up to be the explanans of an orthodox nomological deductive model of explanation, and, a new type of regularity can therefore be explained.

The advantage of this alternative model resides in preserving both intentional feature of behaviour and the idea that intention exhibits a causal efficacy. Besides that, the strict peculiarity of the retroductive approach helps to understand the anomalies. According to the DN account, from premises A, B, C, H, and K, if consistent, it is only possible to entail compatible conclusions D₁, D₂, D₃. But an anomaly, in the retroductive account might be explained not only by different premise sets, but also by incompatible premises sets. A single anomaly can follow from two mutually incompatible premise sets. This is particularly useful for the purposes of explaining behaviour: it is possible to match different (and incompatible) premises following from the anomaly to be explained.

This is of course just the bare core of our alternative explanatory model. Needless to say that, specifying more details such as its applicability, the nature and "logic" of retroduction as well as a more accurate account of the way causal explanations can be conjoined with teleological ones would demand another article, which is now in preparation. This is the dilemma of getting from matter and motion to content and purpose – and back. Moreover, the discussion we have focused on here centers on acts of single agents, rather than the relationship between acts of several agents. Nevertheless, the expansion of the model that has been presented here may provide a framework for dealing with these questions in new ways.

RESUMO: O trabalho consiste numa análise de dois modelos explicativos do comportamento humano considerados fundamentais na literatura contemporânea sobre teoria da ação. O primeiro modelo, o causalista, tenta explicar as ações em termos de causas e leis gerais. O segundo explica a ação em termos de intenções e sínxis de práticas. As dificuldades e problemas de ambos modelos são apresentadas e na última parte do ensaio propomos um modelo alternativo baseado na noção de retrodução.

UNI TERMOS: Explicação; modelo explicativo por leis gerais; ação humana; causalismo; intencionalismo; retrodução.

NOTES
A – This essay is in Davidson, D. (2, p. 261/275).
B – This problem is focused on an essay written by Hempel and P. Oppenheim, “Studies in the Logic of Explanation” in Hempel, op. cit. However, it is doubtful whether or not they have reached a clearcut criterion of lawfulness.
C – This aspect is also stressed by D. Dennett (3, p. 33).
D – The reference is to the essay “Mental Events” (2, p. 207 to 225).

REFERENCES

BIBLIOGRAPHY