The paper was published in *Synthese* 130, 49 - 68 (2002).

METHODOLOGICAL PATTERNS IN A STRUCTURALIST SETTING

WOLFGANG BALZER

ABSTRACT. A new approach to analyze scientific methods as patterns of state transitions is proposed and exemplified by the two most important, general methods: induction and deduction. Though only 'local' states of science are considered in this paper, including hypotheses, data, approximation and degree of fit, the approach can easily be extended to more comprehensive kinds of states. Two 'pure' forms of induction are distinguished, enumerative and hypothesis construction induction. A combination of these two forms is proposed to yield a more adequate picture of induction. While the pure forms of induction are clearly distinct from the deductive pattern, the pattern of the combined form of induction is very similar to the latter. The present account of scientific methods not only points out the differences between different methods but - in contrast to usual discussions of methodology - also clarifies what they have in common.

1. FROM RULE TO PATTERN

There are many things that are called scientific methods and it seems unlikely that all these things can be subsumed under one common notion. I want to deal here with a certain subclass of methods which has mostly attracted the attention of philosophers of science, the class of most abstract and general scientific methods containing induction, deduction, hermeneutics and abduction as its most prominent members. Other kinds of scientific methods, like mathematical, statistical, or logical methods, general methods of measurement (e.g. panel studies), specific methods of measurement (e.g. triangulation in geometry), or concrete procedures (e.g. how to grow a certain strain of bacteria in a vat), will not be discussed here.

Etymologically, and in most of its uses, 'method' refers to a process or procedure, a certain way of doing something. This procedural aspect has been used to argue against the kind of analysis typical for the structuralist approach in the philosophy of science in which (set theoretic) structures occupy a central position.[1] The alleged argument is that structural (and thus structuralist) analysis has a static flavor and therefore is not the best way to approach procedural issues, such as methods. This 'argument' is clearly wrong, as should be obvious at least to anybody with some basic knowledge of set theory. All descriptions of procedures which have been given so far in science, not to speak of all mathematical theories treating dynamical systems, can be, and usually are, cast in set theoretic terminology, i.e. in the vocabulary mistakenly seen as 'static'. It is true, though, that methodological issues have been somewhat neglected in the structuralist approach.

The standard account of methods is in terms of rules. A scientific method is described as a rule the following of which counts as doing science in a respectable manner. On this account a method is given by a *rule description*. There are three main forms such rules can take. First, there is the unrestricted form: *Do b*, where *b* is (a sentence describing) an action. For instance, 'temporarily accept an hypothesis only if it has passed a severe test'. The second, conditional, form is: *If condition c obtains, do b.* Here *c* is a sentence or a set of sentences describing a certain state of affairs. For example, 'whenever the preliminary acceptance of hypotheses is on the agenda, only those hypotheses should be accepted which have passed a severe test'. The third form is this: *in order to achieve goal g, whenever condition c obtains, do b*, where g is a sentence describing a goal. In the Popperian example, we might say 'if you want to do good science, whenever the acceptance of hypotheses is on the agenda, only those hypotheses should be temporarily accepted which have passed a severe test'.

Rules are often formulated in a prescriptive way. In my view, the prescription, addressed to the members of a population, to do a certain kind of action in certain situations is best understood in terms of the fact that that kind of action in the population is performed in the pertinent situations without too many exceptions, and in terms of certain social 'mechanisms' furthering such performances. In this sense, the prescription can be understood in descriptive terms, and for this reason I prefer the descriptive account of rules and rule

[1] The basic reference for the structuralist approach is (Balzer et al., 1987). Compare also (Stegmüller 1979, 1986), (Balzer et al., 1993), (Balzer and Moulines, 1996) and (Balzer et al., 2000).

following as being more basic. The members of a population follow a rule if, in a statistical reading, they perform the pertinent action whenever appropriate.

Clearly, an unrestricted rule *do b* cannot be understood as a permanent prescription to do b; there are other things besides b which people have to do. This weakness is removed in the conditional form. However, this form still has a drawback because in order to distinguish between events of successful rule following and failure one has to refer to some standard external to the description of the rule. In order to state that a rule has been followed successfully, we must refer to the goal which usually is reached applying the rule. In the Popperian example, the goal is to do good science. Applying the Popperian rule therefore leads to success iff it leads to doing good science, and whatever that means, there is no obvious way leading from the observation that the rule has been followed to the conclusion that good science has been done. No wonder that this rule aroused an extensive discussion. One way of introducing more content here is to refer to truth. Instead of doing good science, the goal then becomes to find true hypotheses. The difficulty with the goal and success in this example is generic for the whole domain of scientific methods under discussion, and the shift from 'good science' to 'truth' does not seem to provide much additional clarity.

The explicit inclusion of a goal is not a matter of necessity. The goal may be taken to be part of the action,[2] in which case it need not be mentioned as an extra component of the rule. Whether the goal is made explicit or not thus largely is a matter of convenience. If the action regularly is successful, the goal becomes less important. If the action often fails to reach its goal, success may become a major issue, and in order to check for success the goal must be clearly stated.

All this taken for granted, there is one big problem with the rule description account of scientific methods. The account is firmly tied to natural language and does not allow to impose theoretical structure on the description of a rule unless this structure can be formulated in terms of a natural language. Any rule description therefore is bound to the fuzzyness of natural language. As long as natural language remains recalcitrant to formal analysis there is little hope to achieve a precise, theoretical understanding of scientific methods along the rule account. More precisely, this *problem of natural fuzzyness*, as I will call it, arises as follows.

In natural languages, actions are described in terms of verbs or verbal phras-

[2]This is the standard approach in action theory, see e.g. (Tuomela, 1984).

es. Most natural languages contain very many verbs (from 30 to 50 thousand and more),[3] and provide rules to generate an unlimited number of verbal phrases. However, the internal structure of this huge system of verbs is still largely unknown. As long as no such structure is available, any action description by means of a verb will provide a rather isolated picture. The point is that a description using some definite structure can exploit the many other concepts which are used in order to define that structure, and their interrelations. A description in terms of just one verb lacks these possibilities. This point is further clarified by looking at the way of scientific, theoretical descriptions. These are stated by means of a whole *set* of axioms and a *set* of concepts which describe the theoretical models. The description of a system or situation in terms of such a theoretical model thus exploits more resources than a description in terms of one single concept or term. It can be much more detailed and at the same time more precise than a description in terms of one notion. A description by means of one verb from a natural language does not reveal any interesting internal structure comparable to the internal structure contained in a theoretical description of a scientific model.

This discussion points to a way in which the problem of natural fuzzyness can be overcome. It can be overcome by leaving the domain of verbs and verbal phrases as found in natural language and by describing the procedures and actions associated with scientific rules and methods in a theoretical way which does not depend on the system of verbs of a natural language.

There are already two modest attempts in this direction. A first proposal was made in the context of measurement.[4] According to this proposal, a method of measurement is identified in terms of theoretical models, called *measuring models*, which can be defined in a formally precise way. Each measuring model 'describes' a system, whether real or only possible, in which a measurement is, or could be, performed. A method of measurement therefore can be identified with the class of all corresponding measuring models.[5] Each measuring model has a rich internal structure, which is obtained from some established scientific theory or several such theories. The method thus is described in a way which explicitly uses a complex structure, the structure of models of established scientific theories. This structure describes (among other things) the transition from an initial state before the method was applied to a state obtaining after its application. As descriptions of scientific method involve many state tran-

[3] See (Ballmer and Brennenstuhl, 1981) for a compilation of English verbs.

[4] See, for instance, (Balzer, 1985, 1992).

[5] It can be left open here whether success should be explicitly included in the definition.

sitions the account just described is not general enough to capture scientific methods.

A second, very simple account of rules is found in computer science, where rules are construed as pairs (*condition, action*). Here, the actions are just changes of the internal state of a machine, the states are completely specified by finite descriptions, and the changes themselves also are completely specified as state transitions, i.e., by two lists of symbols, describing respectively the state before and after the change. The transition or 'action' thus is completely specified by two state descriptions (before and after) which in turn are completely specified in terms of strings of abstract symbols, not in terms of verbal phrases. The problem with this approach is that it achieves precision by means of using very simple states, and even though it can be proved that this simple approach is very powerful indeed, such proof is a piece of theoretical knowledge which does not help in applications where transitions involve human action and states are of the complexity of scientific theories.[6]

The approach presented here takes over this general view about modelling of methods as state transitions but uses more complex kinds of states. Moreover, the picture is extended from single state transitions to *patterns of state transitions*. The scientific methods under discussion are complex. Each single application of such a method takes place, and generates, a real process, i.e., a system running through different states over time. Thus in one application of a method *several* state transitions of different kinds will occur. The sequence of these transitions follows certain patterns which are different for different methods, and typical for each method. It therefore seems possible to characterize each method by 'its' pattern of state transitions. Of course, 'characterization' here must not be understood in the sense of grasping the full meaning of a scientific method. It means that the patterns of state transitions (a) can be used to distinguish between the different methods, and (b) are necessary conditions which any process must satisfy in order to pass for an application of the corresponding method. In the description of methods, three levels can be distinguished. On a first, most basic level, the 'deep structure' of the processes arising from application of the method is revealed as a pattern of state transitions. On a second level, special constraints on single state transitions are described which are typical for a particular method. On a third level, informal components are added to the description which at the present stage cannot be

[6]Maybe this state of affairs is historically contingent and simply due to the lack of interest of computer scientists in social theories so far.

made formally precise.

I will now propose a general, structuralistically inspired frame in which a large set of transition patterns can be precisely described, and I will have a closer look at the two examples of induction and deduction, leaving hermeneutics[7] and abduction for separate treatment.

2. A FRAME FOR SCIENCE DYNAMICS

The states I will conside[8] roughly can be taken as corresponding to states of structuralist *theory-element*, i.e., smallest units which can pass as empirical theories. States consist of four components: a class M of *models*, a set D of *data structures*, an *approximation apparatus* U and a *degree of fit* F.

Readers not familiar with the structuralist meta-theory may best think of M as represented by a scientific hypothesis,[9] and of D as a finite set of data, i.e. atomic sentences formulated in the vocabulary of M. The hypothesis may fit with the data up to a certain degree. Think of the data as points in a 2-dimensional coordinate system, and of the hypothesis as a curve in this coordinate system. The degree of fit then corresponds to some kind of 'mean distance' of the points to the curve. In general, there is no single distinguished notion of distance that can be used in all cases of determination of fit between a hypothesis and given data. Rather, the notion of distance usually is specified only in the context of a concrete, given theory, and *for* that theory. The specification of some notion of distance[10] therefore is presupposed if one wants to talk about the fit of a hypothesis and data. The approximation apparatus can be regarded as yielding (among other things) such a specification. The degree of fit in general may be represented by a non-negative, real number (infinity included), but in the present paper I will use very coarse, qualitative degrees.

More formally, the models are conceived of as set theoretic structures of

[7]See (Balzer, 1997, Chap. 4) for a first analysis of hermeneutics along the lines presented here.

[8]The following draws from, and emends, ideas first formulated in (Balzer, 1997, Chap. 4). The standard structuralist notions (models, approximation apparatus, theory-elements etc.) are used in the sense of (Balzer et al. 1987).

[9]In fact, I will make use of this reading in the following, and often speak of M as a 'hypothesis'.

[10]Besides the well known numerical distances, Euclidean, maximum, minimum, supremum etc., there are many other formal definitions of distance functions which satisfy the topological axioms for a metric, for instance the number of elements of the symmetry difference of two sets. Compare (Balzer and Zoubek, 1994) for some standard examples.

a given type τ in the sense of formal logic, and the data structures as finite substructures of structures of type τ, where the notion of a substructure is used in the following, generalized sense. If $x = \langle u_1, ..., u_k \rangle$ is a structure of type τ, then y is a *substructure of x* iff $y = \langle v_1, ..., v_k \rangle$, where, for all $i \leq k, v_i \subseteq u_i$, and the non-empty components of y satisfy all typification requirements pertinent to them. For instance, if $x = \langle D_1, D_2, R_1, R_2 \rangle$ with $R_1 \subseteq D_1 \times D_2$ and $R_2 \subseteq D_1$ then $y = \langle D_1', D_2', R_1', \emptyset \rangle$ is a substructure of x if $\emptyset \neq D_1' \subseteq D_1$, $\emptyset \neq D_2' \subseteq D_2$ and $\emptyset \neq R_1' \subseteq R_1$. This allows for substructures to have empty sets as components, and thus for data structures representing data formulated only with a subset of the notions that occur in the full models of a theory.

The approximation apparatus of a theory among other things specifies either a topological (e.g. a metrical) or a uniform space on the set of structures of type τ. With the help of such spaces we can express that a model and a data structure are close to each other (with a certain degree) and in this sense fit with each other in the special sense given by the particular space. The approximation apparatus moreover contains *admissible* blurs or degrees of fit or closeness. If a model and a data structure are close to each other up to a degree admissible for the theory under consideration, this means that, from the point of view of that theory, the data are 'good enough' in the light of the model and conversely, the model is 'good enough' in the light of the data. So both parts corroborate each other. These admissible blurs usually are tied to similarities or distances in the data. For instance, the standard deviation of values obtained from repeated measurements of a function for 'the same' argument yields a kind of very local 'admissible blur' which can be used only for the particular argument under investigation. If such 'local' blurs could systematically be combined to comprise all the parts of full structures we would obtain definitions of a theory's admissible blurs in terms of distances of measured values.[11]

In the structuralist model a theory has more inner structure than it has in the statement view (i.e., when conceived as a set of statements). There is a *class* of models (corresponding to one hypothesis or to several hypotheses in the statement view), and there is a set of data structures (corresponding to one homogenous, unstructured set of statement view observation sentences). There is a set of *intended systems*, i.e. real systems to which the adherents of the theory intend to apply it. The data which have been obtained from one such system, when put together in the right way, form a data structure as introduced above. A theory therefore has many different data structures, depending on

[11] Compare (Balzer, 1997, p. 223) for more details.

its different intended systems.[12] This means that the undifferentiated set of statement view observation sentences on the structuralist account is structured according to the different real systems from which the data come.

The approximation apparatus initially provides notions of closeness between single models and single data structures, i.e., expressions of the form *fit(x, z, ε)* where x is a model, z is a data structure, and ε a positive real number. Though the following analysis in principle can be carried out even at this fine grained level of single structures, reasons of simplicity suggest to begin at the aggregate level of *sets* of models and data structures. Accordingly, in this paper I will consider statements of fit of the following form $\mathit{fit}(M, D, U, F)$ where M is a model class, D a set of data structures, U an approximation apparatus, and F a degree. Such a statement is read as follows: 'Relative to the approximation apparatus U, the degree of fit between M and D is F'.

For the purpose of this paper it is not necessary to specify how exactly the fit of M with a set D of data structures is defined in terms of the fit of single members M and D. In general, there are many possibilities, such as minimum or maximum of the individual values of fit (which also works for qualitative degrees, where e.g. the minimum corresponds to an intersection), or mean values of individual values of fit (in cases where quantitative measures of fit are available).[13] Assuming any such definition, we can express that, with respect to an approximation apparatus U, a set D of data structures fits with a class of models M to a given degree F.

I will consider only three degrees of fit in this paper: $+$ (= good fit), $-$ (= bad fit), and $\circ$ (= fit is unknown), thus F will take only one of the three values $+, -, \circ$. Of course, what counts as good or bad depends on the particular situation in which fit is investigated and on the topological notion of approximation or closeness given by the approximation apparatus. As already pointed out, the admissible blurs are tied to distances and similarities in the data at least in a pragmatical way. The degree of fit being unknown usually means that it has not been investigated.

In science dynamics we want to describe (and ultimately also to explain) transitions of states of science, and in a first approximation we may consider states of the above form $\langle M, D, U, F\rangle$ to represent states of science (though only very locally and on the 'pure' level of knowledge representation). We can conceptually distinguish transitions in which only one

[12] It should not be ruled out that one intended system gives rise to different data structures.

[13] Compare (Balzer and Zoubek, 1994) for concrete definitions.

$$
\begin{array}{llcl}
\text{(A1)} & \langle M, D, U, \circ \rangle & \longrightarrow & \langle M, D, U, + \rangle \\
\text{(A2)} & \langle M, D, U, \circ \rangle & \longrightarrow & \langle M, D, U, - \rangle \\
\text{(B)} & \langle M, D, U, X \rangle & \longrightarrow & \langle M', D, U, \circ \rangle \\
\text{(C)} & \langle M, D, U, X \rangle & \longrightarrow & \langle M, D', U, \circ \rangle \\
\text{(D)} & \langle M, D, U, X \rangle & \longrightarrow & \langle M, D, U', \circ \rangle \\
\text{(E)} & \langle X, D, Y, \circ \rangle & \longrightarrow & \langle M, D, Y, \circ \rangle \\
\text{(F)} & \langle M, X, Y, \circ \rangle & \longrightarrow & \langle M, D, Y, \circ \rangle \\
\text{(G)} & \langle M, D, Y, \circ \rangle & \longrightarrow & \langle M, D, U, \circ \rangle
\end{array}
$$

Figure 1

component, two components, or more than two components are changed. This yields a longish list of formal possibilities. In Figure 1 some of the more interesting cases are summarized.

In cases (Al) and (A2) in the initial state the degree of fit is unknown, while after the transition it is good or bad. Such transitions capture the actions of investigating and stating the degree of fit between M and D with respect to U. In each of the cases (B), (C), (D) one of the M, D, U components is changed. Irrespective of the degree of fit in the initial state (denoted by the variable X), this results in a new state in which the degree of fit involving the new component has not yet been investigated. Cases (E), (F), (G) cover transitions at very early stages of a scientific development (a theory-evolution or a research program) in which not all the three components M, D, U have been introduced. In case (E) initially there are just ‘data’. This represents a state in which a new, reproducible phenomenon has been observed and stated but has not yet been explained by a hypothesis. In the final state of this transition a corresponding hypothesis has been introduced, but the degree of fit has not yet been investigated. In the initial state of case (F) a hypothesis M ‘looks for’ data. This can occur when new models are transferred from one domain of application, where they had been successful, to a new domain which initially is just given in terms of some intended systems. No data are yet available for the models, e.g. because they use a new vocabulary, and data in the new domain previously had been formulated in a different vocabulary. After the transition, data are present, but it has not yet been investigated. The final case describes the first introduction of an approximation apparatus for a new pair $\langle M, D \rangle$. Of course, in the real process, some of these transition types may be realized at the same time.

Sequences of state transitions are best represented by means of flow chart diagrams. Figure 2 shows such a diagram which captures the most essential transitions. Each box contains a description of a state, and the arrows indicate possibilities of state transitions. If only one arrow leaves a box this means that the state noted in that box is followed by another state

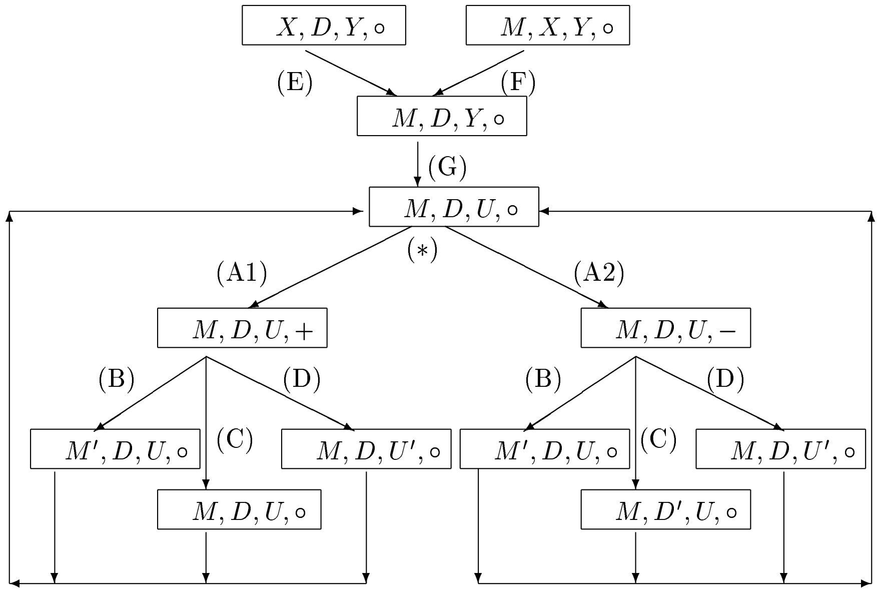

Figure 2

of the form written in the box to which the arrow points. If two or more arrows leave a box, then external influences determine along which arrow the transition will proceed. The arrows pointing upwards denote loops (or recoupling). When the process runs along such an arrow, the content (state) of the box at the beginning of the arrow is written in the box at the end of the arrow, whose original content is overwritten. The use of loops allows the depiction of possibly infinite sequences of state transitions in a rather compact form in which patterns can be recognized relatively easily.

The flow chart in Figure 2 has no end. Beginning with data or a hypothesis, it will reach the state $\langle M, D, U, \circ\rangle$ in the center at the top. In this state, models, data structures and an approximation apparatus are given but fit has not been determined. The two arrows leaving this box represent the two possible outcomes of an investigation of fit: good fit (+, left hand arrow) or bad fit (−, right hand arrow). In each of the two successor states three possibilities arise, represented by three arrows leaving that state: the models can be replaced by a different set M', and similarly for D and U. Each of these replacements leads to a state in which fit again is unknown (the six boxes at the bottom). From each of these boxes there is a recoupling arrow back to the central state $\langle M, D, U, \circ\rangle$. Running along any of these arrows the content of the box at the bottom is written into the box $\langle M, D, U, \circ\rangle$ where the next round is started. As there is no exit, the process thus depicted would run forever.

In general, more complex flow charts can be constructed from the 'basic' transitions (arrows) by means of two devices. First, certain small, local patterns may be 'put together' in a certain order. Second, constraints may be introduced on the choice of new model classes M', new sets of data structures D' and new approximation apparatuses U'. I will now exemplify these possibilities in the discussion of deduction and induction.

3. EXAMPLES: DEDUCTION AND INDUCTION

The core of the deductive method can be described as follows. In order to test a new hypothesis, the hypothesis should be submitted to a severe test. This test consists in the deduction of a new observation sentence (datum, prediction) from the hypothesis plus necessary initial conditions and/or background knowledge. If the observation sentence can be verified the hypothesis can be temporarily accepted, if not it must be rejected. In the latter case a new hypothesis should be considered and treated in the way just described. The choice of a new hypothesis is constrained by two conditions. First, the new hypothesis should be rather improbable, and second, it should not be *ad hoc* (i.e. obtained from the 'old' hypothesis by means of a minimal change to 'save' the new observation). A third constraint is imposed on the agreement about the new observation sentence. This should proceed in terms of rules about which the persons involved have agreed beforehand, i.e., before they engage in applying these rules in a particular case.[14]

I will not try to incorporate the constraints, but will concentrate on the basic sequence of introducing a new observation sentence, testing it, and eventually choosing a new hypothesis. This yields a flow chart as depicted in Figure 3. In

[14]See for instance, (Popper, 1959) or (Lakatos, 1978).

some initial state at the top, models (hypothesis), data structures (data) and an approximation apparatus are given, and fit has not been investigated. That is, the hypothesis has not even been evaluated in the light of the available data.

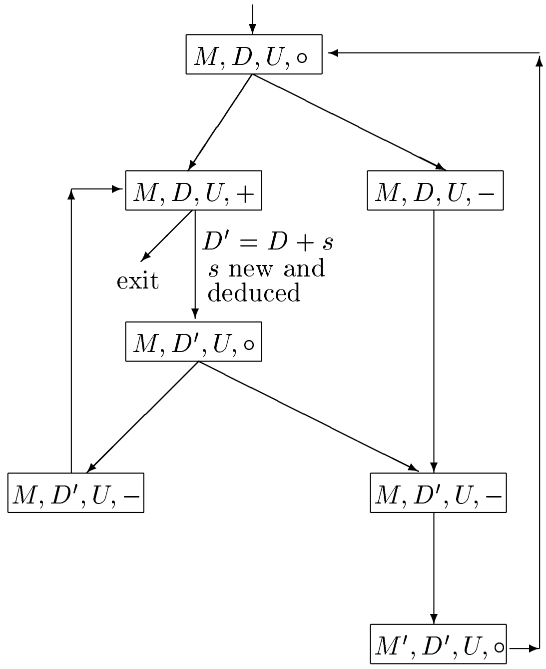

Figure 3

The two branches below that box show the two possible outcomes of an investigation of whether the hypothesis fits the given data. If it fits, we reach (on the left) the initial state for further deductive development, $\langle M, D, U, + \rangle$: the hypothesis fits the available data. This initial condition of fit is not stressed in the literature, but it is necessary. If a new hypothesis is to be considered at all, it first of all has to fit the data already at hand. A hypothesis not satisfying this condition will not be seriously considered at all, and will not be communicated to the scientific public. Now the hypothesis (represented by M) is severely tested. First, a new observation sentence s is derived. This leads to a new state $\langle M, D', U, o \rangle$ in which the initial data D are extended by s. Next the validity of the new observation sentence s is checked. In the picture this is represented by the two arrows leaving that box. If the state at the left is reached this means that the new sentence, the 'prediction', was validated, because fit with the other data was already investigated in previous steps with positive result. If the new sentence turns out true, then by recoupling the resulting state $\langle M, D', U, +\rangle$ is written into the upper left box, where the process of testing the hypothesis starts anew. In case this left hand loop has been run through several times, which means that the initial, given hypothesis has been successfully tested several times, a provision is made to stop the process by entering the exit arrow. As a matter of empirical, historical fact, even hard headed deductivists do not repeat this loop many times.

If the prediction turns out to be false, we follow the right hand arrow leaving $\langle M, D', U, \circ\rangle$, reaching $\langle M, D', U, -\rangle$. In this state, the hypothesis (represented by M) has been falsified. It is abandoned and replaced by a new one, M', satisfying the constraint that it has not yet been tested. $\mathbf{M}^*$ denotes the set of all model classes which have previously been considered

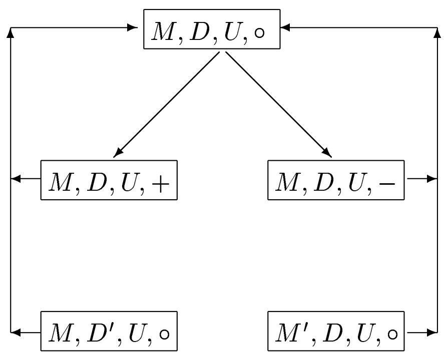

Figure 4

in the process. The new state $\langle M', D', U, \circ\rangle$ is then written in the box on top where the process starts again.

When in the first, top branching in Figure 3 the right hand state is reached, this means that a hypothesis (class of models) which is considered does not fit with the known data. In this case no test of the hypothesis is necessary. It is abandoned, and replaced by another one. In the flow chart, an intermediate step is inserted to achieve greater simplicity: $\langle M, D, U, -\rangle$ first is replaced in a trivial way by $\langle M, D', U, -\rangle$ where $D' = D$, i.e., nothing has changed. Only after that step, M is replaced by M'.

This flow chart in Figure 3 contains two main loops. The first includes the three states $\langle M, D, U, +\rangle$, $\langle M, D', U, \circ\rangle$, $\langle M, D', U, +\rangle$ on the left hand side. This loop may be run through several times, which means that a fixed hypothesis successfully passes several tests (confrontations with new predictions). The second main loop is the one including $\langle M, D, U, \circ\rangle$, $\langle M, D, U, +\rangle$, $\langle M, D', U, \circ\rangle$, $\langle M, D', U, -\rangle$, $\langle M', D', U, \circ\rangle$. This loop may be run through with or without interceptions of the previous, left hand 'success' loop. Neglecting these possible successes, the second loop may be described as follows. The hypothesis fits the available data and thus is considered seriously. It is then confronted with a severe test (a prediction s) which fails. The hypothesis is abandoned, replaced by a new one, and the loop is repeated.

Readers used to read flow charts will immediately see that this flow chart has two redundant boxes, namely $\langle M, D', U, +\rangle$ and $\langle M, D', U, -\rangle$. Removing these and adjusting the arrows we obtain the following simpler, equivalent flowchart (Figure 4).

Turning to induction, the situation becomes more complex. There is no authoritative source (like Popper for deduction) nor a school or a research program investigating 'the' inductive method. Even worse, two quite different approaches can be found in the literature which use the label 'induction'. The first is typically found in philosophical writings, and sometimes labelled 'enumerative induction'. The second approach has spread from philosophy of science to AI and computer science, where it is found in two different versions, called 'machine discovery'[15] and 'inductive inference in the limit',[16] respectively. I will summarize both approaches under the label of 'inductive hypothesis construction'. I will look at these variants and then suggest to extend the use of the term 'induction' to a more general method in which they are *aufgehoben*.

The first form of induction may be understood as a method of testing a given hypothesis. It proceeds by systematically testing all instances of the hypothesis. For example, if the hypothesis is a universal sentence 'for all x, $A(x)$', one tries to test all instances $A(b)$ where b varies in the set of all objects for which the hypothesis is claimed to hold true.

The flow chart of this very simple method is shown in Figure 5. It begins with a hypothesis (class of models), an approximation apparatus, and a set of data which initially may be empty. All these data, as well as those to be considered in the process, come from one system in which the validity of the

[15]Compare e.g. (Langley et al., 1987).

[16]Compare e.g. (Lauth, 1996). All these kinds of induction as scientific methods are related to the so called problem of induction, but I will not attempt to spell out this relation here.

hypothesis is under investigation. In Figure 5 the initial state is represented in the top box. If the hypothesis is tested and fits the given data (following the arrow to the left), a *new* instance b is considered, and it is checked whether the hypothesis is true for b. This leads to the lower left box, from which by recoupling we get back to the original state. The loop consisting of these three states represents the core of enumerative induction. The hypothesis is inductively confirmed by repeated check of different instances. In the standard situation where a fixed system is considered, b is just a new datum from that system, i.e., an observation sentence which is obtained from the system and which has not yet been investigated in previous loops. D^{tested} denotes the set of all data which have been considered previously. The situation gets a bit more complicated when there is more than one data structure (which is the normal case). In this case the loop just considered is restricted to data from one system, and a second loop over the different systems is added 'on the top' of the loop just described.

When things go wrong, i.e. when an instance is found which does not fit the hypothesis, the arrow from the top box to the right is reached. Here the process of enumerative confirmation stops. It also stops when the set of data exhausts the given system, i.e. when no new data can be found in

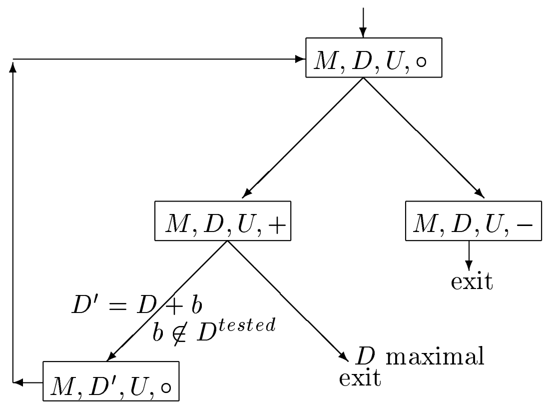

Figure 5

the system which have not already been checked, so that the set of data is maximal in this sense. Though this case ideally can occur only in finite systems it is very relevant for real-life experimental science where some numerically infinite domain is usually coarsened by identifying values which cannot be distinguished experimentally. A typical example is the ideal gas law which is tested by finite series of measurements over real intervals (containing uncountably many numbers).[17]

The second pattern, of inductive hypothesis construction, is obtained from the previous one by interchanging data and hypothesis. Now the data are given and a hypothesis fitting these data is searched in a space of hypotheses. The resulting flow chart is depicted in Figure 6.

When a chosen hypothesis M does not fit the data (in state $\langle M, D, U, -\rangle$) a new hypothesis M' is chosen and by recoupling we get back to the initial, upper box. $\mathbf{M}^*$ denotes the space of all those hypotheses which have not yet been investigated in previous loops. Usually this is a very large space. If $\mathbf{M}^*$ is empty this means that all possible hypotheses from the space have been considered. In this case the process terminates at the lower exit. In case of good fit of M and D the process terminates because the goal of finding a hypothesis fitting the data has been reached.

Both these patterns represent respectable scientific methods. They share the feature of systematically searching a space of possibilities - instances in the enumerative case and possible hypotheses in case of hypothesis construction - a feature which is completely absent from the deductive approach.

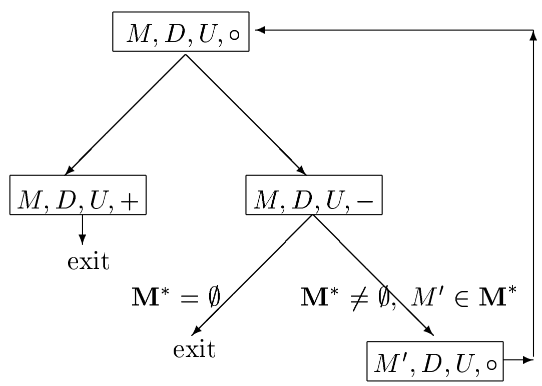

Figure 6

[17]Compare (Balzer, 1997, Sec. 4.5) for a detailed account of this example.

Given that both patterns are well entrenched in scientific practice it would not be fair to reserve the label ‘induction’ to just one of them. Therefore the possibility of combining the two patterns should be seriously considered. In reality both patterns often get mixed. Scientists repeat the confirmation loop until they feel sure enough that further investigation will not lead to significant change of fit, they then try another, sharper hypothesis, run through a number of confirmation loops until this hypothesis also is sure enough, then try another, sharper hypothesis, and so on. Instead of pointing out case studies of historical developments let me just mention the BACON programs[18] which exemplify this mix in a very pure and clear way. As both ‘component’ methods are called inductive, why not also call their combination ‘induction’? The combined pattern is much more systematic than the deductive pattern, and thus certainly deserves recognition as a scientific method. I suggest to use the label ‘induction’ also, and deliberately, for a combined pattern of enumerative and hypothesis constructive induction, and I will use the term in this sense in the following. What is the precise structure of this combined pattern?

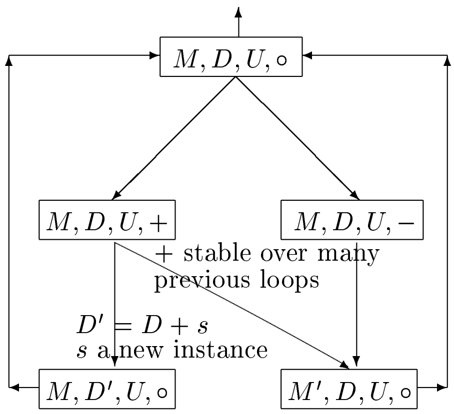

Figure 7

If we simply put together the two flow charts from Figures 5 and 6 in the order giving priority to confirmation loops, the result is a flow chart which is identical with the reduced flow chart of *deduction*. This is somewhat surprising and reveals a closer relationship between the two methods, induction and deduction,

[18](Langley et al., 1987).

than one would expect from the discussions found in the literature. On the other hand there are clear differences between the two methods, and they are at each of the three levels introduced above.

At the level of the deep structure of a pattern there is the following difference. In induction, there is no sharp criterion for leaving the (left hand) confirmation loop. In deduction, this loop is left if and only if the new datum is incompatible with the hypothesis. However, in induction it can also be left in case the datum is compatible. As already noted, if the confirmation loop is successfully repeated many times the hypothesis may become validated in a statistical sense so that there is not much gain in repeating the loop any further. In such cases the loop may be left even if no contradiction has occurred. In deduction, by contrast, the choice of a new hypothesis is triggered only by failure of the test of the previous hypothesis. A decisive difference in the patterns for deduction and induction therefore is that the latter pattern includes an arrow from a state of good fit to a state in which the models have been changed. Figure 7 shows the flow chart for induction obtained by putting together enumerative and hypothesis constructive induction, and adding the possibility of leaving the confirmation loop even in cases where no contradiction has occurred.

Comparing Figure 7 with the reduced scheme of Figure 4 for deduction we see that a further arrow has been added to the deductive scheme, leaving the $\langle M, D, U, +\rangle$ box and pointing to a state $\langle M', D, U, \circ\rangle$ in which a new hypothesis M' is considered. On the level of patterns of state change the combined inductive pattern therefore is properly more general than the deductive one.

At the level of constraints there is a first difference in the nature of the new datum which is added to D in order to obtain D'. While in induction this datum must be an instance of the hypothesis, in deduction the new datum usually is related to the hypothesis in a more complex way, governed by logical deduction. In induction this leads to repetitions of the left hand loop with data of the same form, while in deduction very different kind of data and predictions will occur in the loop - as long as it continues. This difference is not explicit in the flow chart, but we can add a constraint on induction expressing that the new datum is an instance of the hypothesis represented by M.[19] Note that, with respect to this feature, deduction is more general than induction, for the prediction s

[19]Although the straightforward way to formulate such a constraint is to introduce syntax, is should be noted that a purely structuralist formulation avoiding reference to syntax also is possible. I do not work out the details of this constraint which require lengthy definitions and are not really pertinent to the present paper.

in the deductive method may well be just an instance of the hypothesis - even though it must be deducible from the hypothesis and background assumptions.

A second difference on the constraints level is that in deduction the new hypothesis should be a bold or improbable one - at least for Popperian deductivists. This requirement could be stated as a constraint on deductive method. However, the definition of a probability space which allows to assign probabilities to hypotheses is problematic, and presentlv far from being applicable to concrete theories.[20] On the inductive side this condition is replaced by the requirement that the new hypothesis is found by means of some clever heuristic algorithm to search the space of all possible hypotheses. Without entering into a detailed analysis of the two notions of boldness and clever search algorithm it seems clear that requiring boldness yields a more special approach. Among the clever search algorithms one would expect to find some which in each step (or most steps) find the respectively boldest hypothesis.

At the informal level, one difference is that in deduction the new hypothesis should not be *ad hoc*. As no workable criterion of simplicity is available for theories, and as ad hocness is tied to simplicity, this requirement presently remains an informal one.

The result of this first comparison of deduction and the combined notion of induction may be summarized as follows. While at the pattern level and in the rules for the choice of new hypotheses the inductive model is more general, with respect to the choice of new data or predictions it is the other way round. Therefore neither method is more general than the other, though induction (in the combined sense) has some advantage.

4. CONCLUSIONS

I suggested to study scientific methods in terms of patterns of state transitions. The states which were considered in this paper are of the most simple form, consisting of a class of models (a hypothesis), data, approximation apparatus and degrees of fit. Social aspects of scientific developments are left out of the picture. Even at this very coarse and idealized level the approach yields a rich domain of possible patterns, two of which were considered in detail: deduction and induction. In the domain of induction, two different accounts can be distinguished, the enumerative and the hypothesis constructive account. I proposed to use the label 'induction' also for a combination of these two accounts. It

[20] See (Lauth, 1996) for a general construction of such a space via axiomatic set theory.

turned out that at the level of flow chart patterns, deduction and induction are rather similar, in fact deduction is a special case of the combined inductive pattern, though differences could be seen even at that level.

The notion of states used here can be, and should be, generalized to comprise structuralist theory holons[21] which are the most adequate representations of scientific states at the level of knowledge representation. When this is done, the constraints on single state transitions which were described only informally in this paper, can be formally stated. Using more general kinds of states also will allow to pinpoint those areas of scientific change in which social factors are important.

Finally, I claim that other scientific methods, in particular that of hermeneutics, also can be satisfactorily described in the present frame as patterns of state transitions.

REFERENCES

Ballmer, T. and W. Brennenstuhl 1981: *Speech Act Classification*, Springer, Berlin.

Balzer, W., C. U. Moulines, and J. D. Sneed 1987: *An Architectonic for Science*, Reidel, Dordrecht.

Balzer, W., B. Lauth, and G. Zoubek 1993: 'A Model for Science Kinematics', *Studia Logica 5*2, 519-548.

Balzer, W. and G. Zoubek 1994: 'Structuralist Aspects of Idealization', in M. Kuokkanen (ed.), *Idealization VII: Structuralism, Idealization and Approximation*, Poznan Studies in the Philosophy ofthe Sciences and the Humanities, Vol. 42, Rodopi, Amsterdam, pp. 57-79.

Balzer, W. and C. U. Moulines (eds) 1996: *Structuralist Theory of Science: Focal Issues, New Results*, de Gruyter, Berlin.

Balzer, W., J. D. Sneed, and C. U. Moulines (eds) 2000: *Structuralist Knowledge Representation: Paradigmatic Examples*, Poznan Studies in the Philosophy of the Sciences and the Humanities, Vol. 75, Rodopi, Amsterdam.

Balzer, W. 1985: *Theorie und Messung*, Springer, Berlin.

Balzer, W. 1992: 'The Structuralist View of Measurement: An Extension of Received Measurement Theories', in C. Wade Savage and P. Ehrlich (eds),

[21] See e.g. (Balzer et al., 1987).

Philosophical and Foundational Issues in Measurement Theory, Lawrence Erlbaum, Hillsdale, NJ, pp. 93-117.

Balzer, W. 1997: *Die Wissenschaft und ihre Methoden*, Alber, Freiburg.

Lakatos, I. 1978: *The Methodology of Scientific Research Programmes, Philosophical Papers*, Vol. 1, edited by J. Worrall and G. Currie, Cambridge University Press, Cambridge.

Langley, P., H. A. Simon, G. L. Bradshaw, and J. M. Zytkow 1987: *Scientific Discovery*, MIT Press, Cambridge, MA.

Lauth, B.: 1996, 'Inductive Inference in the Limit of Empirically Adequate Theories', *Journal of Philosophical Logic* 25, 525-548.

Lauth, B. 1996: 'Probability, Confirmation and Testing from a Structuralist Perspective', in W. Balzer and C. U. Moulines (eds), *Structuralist Theory of Science*, de Gruyter, Berlin, pp. 115-137.

Popper, K. R. 1959: *The Logic of Scientific Discovery*, Hutchinson, London.

Stegmüller, W. 1979: *The Structuralist View of Theories*, Springer, Berlin.

Stegmüller, W. 1986: *Theorie und Erfahrung*, Band II, Dritter Teilband, Die Entwicklung des neuen Strukturalismus seit 1973, Springer, Berlin.

Tuomela, R. 1984: *A Theory of Social Action*, Reidel, Dordrecht.