Reports on Philosophy
№ 21/20003

JAN WOLEŃSKI

Metalogical Observations About Underdetermination Of Theories By Empirical Data

The thesis that scientific theories are undetermined by empirical data (hereafter the thesis **TU**) belongs to the most important elements of the contemporary philosophy of science. In particular, **TU** is usually considered as a very serious argument for an essential dependence of experience data on theories. **TU** is formulated in two versions[1] strong (**TU1**) and weak (**TU2**):

(**TU1**) For every theory **TH** which is accepted on the basis of an experience **E** there is at least one theory **TH'** such that
a) **TH'** is acceptable;
b) **TH** and **TH'** are empirically equivalent;
c) **TH** and **TH'** are mutually incoherent;
(**TU2**) There are always alternative theories which
(a) have an empirical confirmation;
(b) have essentially different worlds as their models, that is, such that
mutually contradictory statements are valid in them.

I will treat theories and empirical data as sets of sentences. If a set **TH** belongs to theories, **TH** = *Cn***TH**. It means that theories are closed under the consequence operation (*Cn* denotes here the consequence operation associated with classical logic). Moreover, I assume that our theories are axiomatizable, that is, for any theory

[1] See M. Hesse, "The Hunt for Scientific Reason", in: *PSA 1980*, v. II, ed. by P.D. Asquith, R. Giere, East Lansing: Philosophy of Science Association, 1981, p. 5, 8.

TH, there is a set $\mathbf{AX}^{\mathbf{TH}}$ such that $\mathbf{TH} = Cn\ \mathbf{AX}^{\mathbf{TH}}$; symbol $\mathbf{KAX}^{\mathbf{TH}}$ will denote the conjunction of axioms belonging to $\mathbf{AX}^{\mathbf{TH}}$. That a theory **TH** is acceptable on the base **E** (consisting of sentences assumed to be true) means that if $A \in \mathbf{E}$, then $A \in Cn\mathbf{KAX}^{\mathbf{TH}}$. It is equivalent to the condition: $\mathbf{E} \subseteq Cn\mathbf{AX}^{\mathbf{TH}}$ or $\mathbf{E} \subseteq Cn\mathbf{TH}$. If **TH** is strictly universal (I will assume that), $\mathbf{E} \subset Cn\mathbf{KAX}^{\mathbf{TH}}$. Furthermore, since $\mathbf{KAX}^{\mathbf{TH}} \cap \mathbf{E} = \varnothing$, we have that for any $A \in \mathbf{KAX}^{\mathbf{TH}}$, $A \notin Cn\mathbf{E}$. Both (**TU1**) and (**TU2**) are existential statements which assert the existence of sets satisfying the prescribed conditions. I will not enter into a controversial problem whether the history of science confirms (**TU1**) or even (**TU2**). Hence, I regard both theses as asserting that such theories are possible. Hence, the assertion that there are (exist) sets falling under (**TU1**) or (**TU2**) should be understood in the same way as any existential statement in set theory.

Assume that the content of (**TU1**) is represented by

(1) For any theory **TH** acceptable on the basis of **E**, there is another theory **TH′** such that
(a) $\mathbf{E} \subset Cn\mathbf{KAX}^{\mathbf{TH}'}$;
(b) for every **E**, $\mathbf{E} \subset Cn\mathbf{KAX}^{\mathbf{TH}}$ if and only if $\mathbf{E} \subset Cn\mathbf{KAX}^{\mathbf{TH}'}$;
(c) **TH**, **TH′** and **E** are internally consistent;
(d) $\mathbf{TH} \cup \mathbf{E}$ and $\mathbf{TH}' \cup \mathbf{E}$ are consistent;
(e) **TH** and **TH′** are mutually inconsistent.

In order to explain some problems concerning (1) we can use the concept of branchable sets of sentences[2]. Here is the definition (and a basic fact about branchable sets).

(2) (a) A consistent set **X** of sentences is branchable at a sentence A if and only if the sets $\mathbf{X} \cup \{A\}$ and $\mathbf{X} \cup \{\neg A\}$ are consistent;

X is branchable if and only if there is at least one sentence such that **X** is branchable at A;

X is branchable if and only if it is incomplete.

It might seem that (1) suggests that **E** is branchable at $\mathbf{KAX}^{\mathbf{TH}}$, because **E** is consistent and incomplete and the sets $\mathbf{E} \cup \{\mathbf{KAX}^{\mathbf{TH}}\}$ and $\mathbf{E} \cup \{\neg\ \mathbf{KAX}^{\mathbf{TH}}\}$ are consistent; of course, **E** is branchable at $\mathbf{KAX}^{\mathbf{TH}'}$ for similar reasons. However, it leads to a consequence

[2] See G. Asse, *Einführung in die mathematische Logik*, Teil II, Leipzig: Teubner, 1972, p. 168.

which is difficult to be accepted. Consider an arbitrary A belonging to **E**. By the assumptions about **E**, **TH** and **TH′**, we have that $A \in Cn\{\mathbf{KAX}^{\mathbf{TH}}\}$ and $A \in Cn\{\neg\ \mathbf{KAX}^{\mathbf{TH'}}\}$. It entails, by axioms for Cn, that $A \in Cn\emptyset$, that is to logic. However, it means that A is a theorem of logic and cannot represent any piece of synthetic empirical data. At first sight, two possible solutions come to mind. The first is to admit that A has one meaning in the context of **TH**, but a different one with respect to **TH′**. Yet it raises the problem how **TH** and **TH′** can be empirically equivalent in this situation, because they work with different data. Another way out consists in taking the set $\mathbf{E} \cup \{\mathbf{KAX}^{\mathbf{TH}}\}$ as branchable, but it is unclear for which sentences, interesting from the point of view of **TH**, it holds.

The relative Lindenbaum theorem on maximalization[3] is another metalogical device which can be used in analysis of **TU**. The theorem is this:

(3) If **X** is a consistent set of sentences and $A \notin \mathbf{X}$, then there is a set **Y** such that

(a) $\mathbf{Y} = Cn\mathbf{Y}$;

(b) $Cn\mathbf{X} \subseteq \mathbf{Y}$;

(c) $A \notin \mathbf{Y}$;

(d) for every B, $A \in Cn(\mathbf{Y} \cup \{B\})$.

This theorem asserts that for every consistent set of sentences **X** and every sentence A which does not belong to consequences of **X**, there is a maximally consistent set **Y** (an oversystem of **X**) relatively to A. The extension of $\boldsymbol{X}$ to **Y** is not unique. There are many (in fact, at least countable infinitely) oversystems to a given set **X**. Now (3) applies to the considered problem, because **E** is consistent and $\mathbf{KAX}^{\mathbf{TH}} \notin \mathbf{E}$.

Thus we can treat theories as maximally consistent extensions of empirical data with respect to a given axiomatics. It does not imply that theories are simple inductive generalizations of data. They are only reconstructed as extensions of data; in fact, (3) is not an effective theorem, because it asserts that something (a maximal oversystem) exists, but without saying how it is to be constructed. Returning to (**TU1**), if there is a theory **TH** which implies **E** (in this case, **TH** and **E** are mutually consistent), (3) guarantees the existence of an alternative theory **TH′** entailing **E**, but inconsistent with

[3] W.A. Pogorzelski, *Notions and Theorems of Elementary Formal Logic*, Bialystok: Warsaw University – Bialystok Branch, 1994, p. 210–219.

TH. However, not all constraints stated by (1) are satisfied. In particular, the condition (1b) fails. Clearly, since **TH** is based on $\mathbf{AX}^{\mathbf{TH}}$ but **TH′** implies $\neg\mathbf{KAX}^{\mathbf{TH}}$, both considered theories have to have different sets of logical consequences. Assume that **E** = **E′**. If so, differences between mutually inconsistent theoretical parts of **TH** and **TH′** are simply redundant or artificial; otherwise, both theories are not empirically equivalent in any reasonable sense. It entails that we can regard **TH** and **TH′** as empirically equivalent only modulo the already given **E**. Although **TH** and **TH′** are acceptable, they are so (or at least, can be) not to the same degree. Moreover, nothing blocks unnatural extensions by branchability or taking oversystems. Thus, the defender of (**TU1**) should prescribe additional conditions constraining natural alternatives for given theories. However, it also does not guarantee that the condition (2b) will be preserved. The above argumentation suggests that (**TU**) in its strong formulation fails.

(**TU2**), i. e., the weak thesis, omits (2b). Instead, it requires that alternative theories have an empirical confirmation, which is not the same. They must imply the same already acquired **E**, but not arbitrary data. More formally, the content of (**TU2**) is represented by

(4) For any theory **TH** acceptable on the basis of **E**, there is another theory **TH′** and data **E′** such that

(a) **E** and **E′** at least partly overlap;

(b) $\mathbf{E} \subset Cn\mathbf{KAX}^{\mathbf{TH'}}$;

(c) $\mathbf{E'} \subset Cn\mathbf{KAX}^{\mathbf{TH'}}$;

(d) **TH**, **TH′** and **E** are internally consistent;

(e) $\mathbf{TH} \cup \mathbf{E}$, $\mathbf{TH'} \cup \mathbf{E}$ and $\mathbf{E} \cup \mathbf{E'}$ are consistent;

(f) **TH** and **TH′** have different models in the following sense: there is a sentence $A \in \mathbf{TH} \cap \mathbf{TH'}$, but it is true in one model and false in the second.

The last condition is equivalent to the clause that **TH** and **TH′** are mutually inconsistent, but, according to (**TU2**) in the version given above, it has a semantic formulation. Although the original version particularly stresses that alternative theories have different models or generate different world-pictures, the main point of difference consists in rejecting the condition (1b), that is, the requirement that **TH** and **TH′** are empirically equivalent. Now, they are confirmed by data which at least partly overlap (4a), but are also mutually consistent (4e). The case in which **E** and **E′** form an inconsistent set seems uninteresting for the considered problem. However, the novelties do not help very much. In particular, branchability and (3) are applicable to (**TU2**) and license various extensions of **E** and **E′**. In particular, we will have trouble with (4e). I take this condition in a radical sense, that is, I interpret difference via inconsistency. There are of course other ways. One can understand the difference in question as generating models which are mutual restrictions or extensions. Nevertheless, I think that such modest interpretations are not especially interesting in the context of (**TU**) and it motivates (4e). Since, by (4d), **E** and **E′** are mutually consistent, the inconsistency prescribed by (4e) lays on the theoretical level. It means that **TH** and **TH′** are mutually inconsistent in their theoretical parts. The only interesting situation arises when **E** and **E′** are different (but not inconsistent), because otherwise we come to the argument that empirical sentences are tautologies, because any such sentence belongs to consequences of assumptions which are inconsistent. On the other hand, if **E** and **E′** are different, it means only that either they force a radical theory change, that is, replacing a theoretical postulate by its negation, or a supplementation of the old theory by a new axiom. In the first case, it is difficult to say that both theories generate acceptable world-pictures, but in the second case, we have no reason to apply (4e), because the new theory is an extension of the old one. Once more we can appeal to (**IT**). If this thesis is accepted in its radical form (no theories are commensurable), (**TU2**) receives a strong argument. On the other hand, (**IT**) raises several objections (see Pearce 1986) and by no means is obvious. Independently of the acceptance of (**IT**) it is perhaps interesting to observe, that (**TU**) in any of its forms rather assumes (**IT**) than gives a justification to it. If (**TU**) is separated from (**IT**) and taken in its weak form, it only notes a trivial fact that no set of empirical data uniquely forces a theory which explains them. Metalogic shows that different completions of data by theories are a normal thing, but also makes clear that applications of such terms as “equivalence” or “model” cannot be arbitrary and governed only by rough intuitions.

(**TU**) in its both formulations has obvious consequences for the debate between realism and anti-realism. Let us agree that realism maintains something like that:

(5) Theories refer to the real world.

Now, (**TU1**) as based on (**IT**) is explicitly anti-realistic. However, (**TU2**) is always considered (see Hesse 1980) as dangerous for realism, because it equalizes different "real" worlds. It seems that this conclusion is not justified. In the light of the above given arguments, (**TU2**) just does not regard different models of theories as representations of the real world to the same degree. Several pragmatic factors (see Sintonen 1984) select arguments pro and contra that such and such representation is better (or worse) than its rivals. For example, the realist can maintain that if $\mathbf{E} \subset \mathbf{E'}$, then we have collected new data for regarding the model of **TH'** as a better approximation of the real world than the model of **TH**. Of course, this argument may be erroneous, but there are no sound reasons in order to claim that the realist is infallible.

Bibliography

Hesse M., "The Hunt for Scientific Reason", in: *PSA 1980*, v. II, ed. P.D. Asquith, R. Giere, East Lansing: Philosophy of Science Association, 1981, p. 3-19.

Asser G., *Einführung in die mathematische Logik*, Teil II, Leipzig: Teubner, 1972.

Pearce D., *Road to Commensurability*, Dordrecht: D. Reidel, 1987.

Pogorzelski W.A., *Notions and Theorems of Elementary Formal Logic*, Białystok: Warsaw University – Białystok Branch, 1994.

Sintonen M., *The Pragmatics of Scientific Explanation*, Helsinki: Societas Philosophica Fennica, 1984.