Induction Confirmation Paradoxes and Reinforcement Propagation

José Hernández-Orallo
Dep. de Sistemes Informàtics i Computació, Universitat Politècnica de València
Camí de Vera s/n, 46022 València (Spain), E-mail: jorallo@dsic.upv.es

Abstract

In this paper many of the classical paradoxes of induction confirmation are briefly discussed. Most of them are originated by qualitative views of confirmation. Some proposals to solve them are recalled, especially those which are based on the notion of converse coverage or entailment, jointly with a selection criterion. However, these solutions introduce new problems in the context of confirmation. The constructive reinforcement propagation method introduced in [4][7] is proposed as a solution to the problem of confirmation and analysed in front of all of these paradoxes. It is shown that this new measure is robust to all of them.

1 Introduction

The problem of confirmation has parallelled to the problem of induction. In order to find a justification for induction, different criteria of confirmation have been introduced. A pristine criterion is just a converse coverage condition, known as "converse entailment condition" (C1): e confirms H if H entails e, where e is the evidence (observations) and H the hypothesis. Although this condition is more restrictive than a simple converse consistency condition, there is still a problem of selecting among all the many possible H's that entail e. This generates the hypothesis selection problem.

Instead of applying selection criteria some authors have opted for qualitative confirmation criteria. In the way of finding refutable hypotheses, Jean Nicod introduced a criterion N1: the observation of an A which is a B confirms the hypothesis that all A's are B. This criterion is useful in many situations, but it cannot be used in general.

In a more general way, Hempel introduced several adequacy conditions and studied their combination (from [3] and [1]):

- H1. Entailment Condition: If e entails H then e confirms H. A special case of H1 is H1.1: e confirms e.
- H2. Consequence Condition: If e confirms each of set of sentences, K, then e confirms every logical consequence of K. A special case of H2 is

H2.2: Hempel's equivalence condition for hypotheses, which states that if e confirms H then e confirms any hypothesis which is equivalent to H.

- H3. Consistency Condition: e is consistent with the class of all hypotheses that e confirms.
- H4. Equivalence condition for observations: if e confirms H then any observation equivalent to e confirms H.
- H5. Converse consequence condition: if e confirms H then e confirms any other hypothesis that entails H.

Obviously, the five conditions all together are incompatible. Properly, it is H5 which represents inductive confirmation (bottom-up), while H1-H3 are deductive confirmation criteria (top-down). However, Hempel decided precisely to withdraw H5 by introducing the notion of 'development' [3]. This represents a new criterion: e confirms H iff e entails the 'development' of H wrt. the objects mentioned in e. The 'development' of a formula wrt. some objects is understood as what the formula affirms wrt. these objects exclusively.

On the contrary, Flach [1] has preserved and augmented H5 (separated into a "converse entailment condition" and "converse consequence condition") and has removed accordingly some other adequacy conditions (H1 to H3)). This constitutes an explanatory view of inductive confirmation (although e explains e is a result of his framework, which is rather counterintuitive).

However, most of these approaches are still vulnerable to some paradoxes.

2 Confirmation Paradoxes

Let us recall the most important confirmation paradoxes before discussing which solutions have previously been presented and the concrete solution, constructive reinforcement, which will be analysed in sections 3 and 4.

1. Hempel's Raven Paradox: the proposition "All ravens are black" is 'confirmed' by observations of ravens that are black. But the statement "All ravens are black" is formally equivalent to the statement "All nonblack things are nonravens". The latter (and hence the first) proposition would be confirmed by a white shoe. This paradox originates from Nicod's criterion (N1) and Hempel's equivalence condition for hypotheses (H2.2).
2. Goodman's Grue Paradox [2]: Define 'grue' as "green before time t and blue otherwise". Then observing a green emerald seems to confirm equally well both "All emeralds are green" and "All emeralds are grue" (assuming t is still in the future). It has been argued that 'grue' is not a 'projectible' concept but no formalisations have been given for this justification.
3. The Invented Useful Concept Paradox: given any theory T, a new theory T' can be constructed where each rule has been added an extra condition P and this P is added as a fact. This fact is highly confirmed because it is necessary for each rule of the theory under the converse entailment condition (C1). This is the same problem given why a confirmation criterion H5 gives that $A \vee B$ is at least as confirmed as A.

4. The Everything Confirms Everything Paradox: for any A and B, by the converse entailment condition C1, $A \wedge B$ is confirmed by A since $A \wedge B \models A$, but, by the entailment condition H1, $A \wedge B$ confirms B since $A \wedge B \models B$. Hence, A confirms B. This is why H1 and H2 are incompatible with H5.
5. The Representation Counts Paradox: by the assumption of the equivalence condition for hypotheses (H2.2), given three facts $p(a)$, $p(b)$ and $p(c)$, the theories $H_1 = p(X)$ and $H_2 = p(a) \wedge p(b) \wedge p(c) \wedge (p(X) \leftarrow (X \neq a \wedge X \neq b \wedge X \neq c))$ should have the same confirmation degree.

According to the everything confirms everything paradox, which is the gravest one, most authors [3][1] have avoided the joint use of H2 and H5, as we have commented on in the previous section. If H5 is not considered, the propagation of inductive confirmation cannot be done by the other adequacy conditions unless a special procedure is added, such as Hempel's 'development'. But Hempel's 'development' criterion leads to say that $p(a)$ confirms $\forall X : \neg p(X) \rightarrow q$ for whatever q, under the 'development' of the only object a, because $p(a)$ entails $\neg p(a) \rightarrow q$, closely related to Hempel's raven paradox. But 'development' *also* originates the grue paradox because any observation of an object with property p confirms whatever you like about objects without property p. In other words, why the observation of things before time x does not confirm anything you like about observations after time x? Other criteria such as Nicod's or a converse consistency criterion also lead to Hempel's Raven Paradox.

The use of converse covering or entailment condition avoids Hempel paradox, because $H \equiv \forall x(raven(x) \rightarrow black(x))$ does not entail anything about nonravens. However it does not entail $black(a) \wedge raven(a)$ either. This problem can be minimised by *arbitrary* selecting $B \equiv raven(a)$ as a given fact and $F \equiv black(a)$ as the fact to be predicted, since $H \cup B \models F$. However this choice seems arbitrary and paradox 1 remains unsolved[1].

Occam's Razor clearly avoids paradoxes 2 and 3. The use of a covering criterion alone (the converse entailment condition C1) avoids paradox 4. Finally, a modified Occam's Razor that assigns the plausibility of a hypothesis as the length of the shortest equivalent hypothesis avoids the fifth paradox. However, new paradoxes appear with Occam's razor:

- An increasing number of confirming observations may decrease the plausibility of a hypothesis. For instance, the sequence of 2^n a's is more compressible than a sequence of $2^n + 1$ a's.
- Occam's Razor gives a single value for the whole of the theory but there is no information about which parts are more confirmed than others, or, if exceptions are found, which parts of the theory are to be blamed.
- When used jointly with deductive inference, we have counterintuitive results. The probability of the formula $A \cup B$ is at least equal than the probability of the formula A. However, the latter is shorter and Occam's Razor gives it more plausibility.

[1] Another approach is to consider a different kind of material or default implication instead of the logical implication. However, this solution seems more ad-hoc.

- The plausibility of a hypothesis wrt. an evidence is independent to the evidence (just depends on the length of the hypothesis), thus ignoring when a hypothesis has been generated with an evidence and must be evaluated wrt. another evidence.
- When the observations are incompressible, H becomes the quoting of the observations, which can be interpreted as E is the explanation for E.

The first problem has been solved by monotone variants [9] of Kolmogorov complexity. The other four cannot be solved by an exclusively syntactic criterion.

3 Confirmation as Reinforcement

We must contemplate inductive confirmation (bottom-up) by the converse consequence condition H5 and/or the converse entailment condition C1. However, this would be incompatible to any deductive confirmation. The idea is to consider a quantitative view of confirmation propagation, as Carnap suggested, but not probabilistic. Such a quantitative propagation has been developed in [4][7], known as constructive reinforcement, where parts of the theory are reinforced according to their contribution to covering the evidence. This allows us to determine which parts of the theory are justified by the evidence, and more importantly, which parts of the evidence are well explained by the theory. More concretely, the conditions for such a propagation are:

1. A part of a theory is *reinforced* by an observation if that part is necessary for *entailing* the observation (bottom-up confirmation propagation).
2. An observation is *explained* as much as the parts that are necessary for entailing it are more reinforced (top-down confirmation propagation).
3. The better the theory explains an evidence the better the theory is.

Different instances can be developed obeying the previous conditions. The first thing to determine is the notion of part or unit of a theory. In the most granular case, a unit would be a symbol or even a bit in the descriptional language used for hypotheses. In [7] an instance of the previous reinforcement conditions has been formalised by using model-based languages where the unit is a rule[2].

A unit is 'necessary' if $H \models e$ but the removal of the unit would make that $H \not\models e$. The reinforcement of each unit is computed following the formula 2^{-E_n} where E_n represents the number of observations in the evidence for which the unit is necessary. The explanatory degree or *course* for each observation is computed as the product of the reinforcement of all the rules which are necessary for that observation. Since reinforcement is always ¡ 1 then this product penalises the use of many and non-reinforced rules. If multiple proofs (explanations) are given for one observation, the greatest one is selected.

[2]Choosing a gross granularity at the level of rule produces a tricky problem of joining rules by using if-then-else constructions. Although a specific solution is also presented in [7], it is preferable to use thinner granularities. In a similar way, in order to cope with joint examples or non-factual examples, observations are separated in parts such that the parts entail the whole and viceversa, and no part is tautological. This makes it possible to split observations such as $a \vee b$ or $a \wedge b \wedge \leftarrow \neg d$.

4 Reinforcement vs. Confirmation Paradoxes

These simple constructs are sufficient for assigning degrees of reinforcement and explanation to theories wrt. evidences. Let us show now that the framework is free from the paradoxes of Section 2, independently of the representational language. For instance, in the case of Goodman's Paradox, the part of the concept grue after moment x is not used for any evidence before moment x and hence it is not reinforced. Moreover, the Invented Useful Concept Paradox is also avoided, since the inclusion of invented concepts can increase the mean reinforcement of the theory but they usually diminish the mean explanatory value (*course*) of the evidence. The Everything Confirms Everything Paradox is solved by the quantitative way in which inductive and deductive confirmation are propagated. The Representation Counts Paradox is avoided by considering the plausibility of a hypothesis as the plausibility of the best consistent hypothesis, in the same way as it could be done for the MDL principle[3].

Finally, the Hempel's Raven Paradox seems more difficult to address. Holland et al [8] affirm that "nonravens are simply not a coherent kind of thing", but this is not a solution in general. Moreover, not every instance of Hempel's Paradox are paradoxical. Consider for instance a bag full of cards which can be either red or black on one side and can have either a triangle or a circle on the other side. The first side to be observed each time is random, thus precluding any arbitrary assumption of the figure as being cause of the colour or vice-versa. In this example, the observation of a black card is exactly the same as a nonred card, and the same happens with the figures. Therefore, nonred and nontriangle are coherent concepts. That is to say, we have the following background knowledge $B \equiv \{red(X) \leftrightarrow \neg black(X), triangle(X) \leftrightarrow \neg circle(X)\}$. Consider the following evidence and hypothesis:

$$E \equiv \{triangle(a), black(a), black(b), triangle(b), black(c), circle(c), circle(d), red(d)\}$$
$$H \equiv \{triangle(a), triangle(X) \rightarrow black(X), triangle(b), black(c), circle(c), red(d)\}$$

In this case, the theory of reinforcement gives the following values of reinforcement of the six rules of H and the courses of the eight observations:

$$\rho_1 = 0.75, \rho_2 = 0.875, \rho_3 = 0.75, \rho_4 = 0.5, \rho_5 = 0.5, \rho_6 = 0.75$$
$$\chi_1 = 0.75, \chi_2 = 0.656, \chi_3 = 0.656, \chi_4 = 0.75, \chi_5 = 0.5, \chi_6 = 0.5, \chi_7 = 0.656, \chi_8 = 0.75$$

which show that the rule $triangle(X) \rightarrow black(X)$ is reinforced without any information of which is the cause (the 'arbitrariness' would not be needed either in the black ravens case). In this case red circles confirm black triangles, since the background knowledge makes $triangle(X) \rightarrow black(X)$ equal to $red(X) \rightarrow circle(X)$. This would not happen, however, for the black raven paradox.

Finally, some other problems presented when using the MDL principle are avoided. The method of confirmation propagation is monotone for increasing confirmative evidence and provides detailed information about the units or parts of a theory. In the same way, reinforcement allows to compute how well explained or justified the predictions of the theory are, and give different values

[3] Nonetheless, one can consider that simplification processes are important and a better representation or formulation of the same theory makes it more reliable.

for a theory if it is evaluated with different evidences. Also, extensional theories such as E for E are lowly reinforced. Finally, hypotheses such as $A \cup B$ for observation A_i are clearly handled by giving all the reinforcement to the A unit, and no reinforcement at all to the unit B. The result is that an observation A_j is equally justified for A as for $A \cup B$, because B is not used.

Moreover, other characteristics of the whole theory can be established, such as its unifying power, the uniformity of the compression degree between the theory and the evidence, and many other features (see [5], [6]).

5 Conclusions

Confirmation can be propagated bottom-up by hypothetical inference processes from evidence to theory. In a similar way, confirmation must also be propagated top-down by non-hypothetical inference processes from parts of theory to other parts and finally to the evidence in order to see how well explained it is. In my opinion, a quantitative (but not probabilistic) way is the only way to include both H2 (top-down) and H5 (bottom-up) senses of confirmation, and the propagation conditions of Section 3 are a good settlement to work on.

Theory revision can take advantage of this framework because it is easier to discover weak (non-reinforced) parts of the theory to be revised first. In this sense, we have left out of this paper the corresponding blame assignment for negative confirmation (see [7]). Confirmation propagation must also be considered with a quantitative blame assignment in the presence of noise.

References

[1] P.A. Flach, *Conjectures. An Inquiry Concerning The Logic Of Induction.* Thesis Dissertation, Proefschrift Katholieke Universiteit Brabant, Tilburg, 1995.

[2] Goodman, N., *Fact, fiction and forecast* (2nd Edition), Bobbs-Merrill, 1965.

[3] Hempel, C.G., Studies in the logic of confirmation *Mind* 54 (213): 1-25; 54(214): 97-121, 1945.

[4] Hernández-Orallo, J., Reinforcement Learning in Constructive Languages, *Proceedings of CCIA'98*, pp. 264-272, Tarragona, 21 - 23, October, 1998

[5] Hernández-Orallo, J., A Computational Definition of 'Consilience', *Philosophica*, 61:901-920, 1998

[6] Hernández-Orallo, J., Computational Measures of Information Gain and Reinforcement in Inference Processes, *Thesis Diss.*, 1999, Dep. of Logic and Philosophy of Science, Univ. de Valencia. `http://www.dsic.upv.es/~jorallo/tesi/`.

[7] Hernández-Orallo, J., Constructive Reinforcement Learning, *International Journal of Intelligent Systems*, Vol. 15(3), pages: 241-264., 2000.

[8] Holland, J.H.; Holyoak, K.J.; Nisbett, R.E. and Thagard, P.R., *Induction. Processes of Inference, Learning, and Discovery* The MIT Press 1989.

[9] Li, M.; Vitányi, P., *Introduction to Kolmogorov Complexity*, Springer, 1997.