Induction in Philosophy and AI

John Bell [1]

1 Introduction

In this position paper I recall some of the philosphical background of induction. I also discuss the relationship between induction and non-monotonic logics, and provide a general model-theoretic framework for induction and abduction.

2 Induction in Philosophy

This section provides a summary of some of the philosophical background of induction. It is largely based on Swinburne's introduction to his excellent collection of essays on the philosophy of induction [12].

2.1 Inductive arguments

In the introduction Swinburne makes several important points about inductive reasoning. He begins by proposing the following definition: "An inductive argument is an argument which is not deductively valid but one in which, it is claimed, the premisses 'make it reasonable' for us to accept the conclusion. I shall say that a correct inductive argument is one in which the premisses do 'make it reasonable' for us to accept the conclusion, as claimed; and that an incorrect one is one in which they do not, but it is falsely claimed that they do" [12], p.2. Most of our everyday commonsense reasoning is based on this form of inference; for example, "When I argue from the premise that the 8.30 a.m. bus to town has seldom been late in the past, and that the time is now 8.35 a.m., to the conclusion that today's bus has gone, my argument is not deductive. I do not contradict myself if I admit the premisses but deny the conclusion. Maybe today the bus has been held up at an earlier stage of its journey, and so is late. Or maybe the driver is on strike today and the bus will not run at all. Nevertheless it would generally be supposed that in the absence of further evidence the first premise shows that these latter are unjustified assumptions. (It shows that the bus is not usually held up or its driver usually on strike.) We judge that the premisses make it reasonable for us to accept the conclusion, even though no contradiction is involved in asserting the premisses and denying the conclusion", p.1. Russell [11] explains that inferences of this kind are based on past regularities in our experience which lead us to associate certain outcomes with experiences rather than others, and he notes that this kind of association is also very strong in animals; for example domestic animals expect food when they see the person who usually feeds them. Swinburne continues: "[i]t is also agreed that the arguments of the scientist from what has been observed in the past to what are the true laws of nature are inductive. When, for example, Kepler argued from the many observations of the past positions of the planets made by his predecessor Tycho Brahe to the claim that the planets (always) move in ellipses with the sun at one focus, his argument was inductive. The conclusion goes beyond the premisses. It claims more–that the planets moved in ellipses when Tycho Brahe was not observing them and will continue to do so. Yet the premisses, it is claimed, make it reasonable for us to accept the conclusion" pp.2-3. Swinburne points out that he is using the term 'inductive argument' in a wide sense. It includes arguments from a set of particular observations to a general conclusion–for example the argument that all swans are white because each swan observed thus far is (has been) white–which are sometimes referred to as enumerative induction, or induction by simple enumeration, or rule induction. It also includes arguments such as those which support Newton's laws of gravitation: "Newton argued to his theory of gravitation …from …observations of the positions of the planets and moons, and the rates at which bodies rolled down inclined planes. There is more to most scientific [and everyday] reasoning than merely generalizing observations", p.5.

2.2 Correct inductive arguments

A correct inductive argument, according to Swinburne, is one in which the "premisses …bear a favourable evidential relation to the conclusion", the "premisses make it reasonable for us to accept the conclusion in the sense, very roughly, that they report evidence favouring the truth of the conclusion, making it rational to believe the conclusion, making it likely or adding to the likelihood of its truth", p.3. He sharpens this to "an argument …[is] a correct inductive argument if its premisses make its conclusion more likely to be true that any equally detailed rival", p.4. Swinburne also notes two further important properties of correct inductive arguments. Unlike a valid deductive argument, "[a] correct inductive argument yields only probable knowledge or reasonable belief", p.6. For example (Russell) "[t]he man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that more refined views as to the uniformity of nature would have been useful to the chicken", p.21. Furthermore, "the argument will only yield probable knowledge if we know nothing else which affects the probability of the conclusion, apart from what is stated by the premisses", p.6. For example, if, while waiting for the bus at 8.30, we learn of serious traffic congestion at an earlier part of its route, then it is no longer reasonable to conclude that we have missed the 8.30 bus. So "the addition of new knowledge to the old knowledge may mean that I cannot make the same inductive inferences from the old knowledge. Hence I can only use an inductive argument to give me probable knowledge or reasonable belief, if the premisses contain all of my relevant knowledge", pp.6-7.

When it comes to the question of giving criteria for the correctness of inductive arguments, Swinburne rejects the traditional view that inductive arguments are correct if and only if they are enumerative. He argues that this is not a necessary condition; as there are correct

[1] Applied Logic Group, Computer Science Department, Queen Mary and Westfield College, University of London. E-mail: jb@dcs.qmw.ac.uk

inductive arguments which are not of this form. Neither is it sufficient; as Nelson Goodman's famous paradox [5] shows. (Swinburne states this as follows: "Suppose that all emeralds observed so far have been green. We can conclude by enumerative induction that all emeralds (future as well as past) are green. But now we introduce the new term 'grue', which is defined as follows. An object at a time t is grue if (and only if) it is green and t is before A.D. 2000, or it is blue and t is after A.D. 2000. We now, living before A.D. 2000, record our observations using this new predicate. All the emeralds which we have observed so far have been grue. So if all arguments of the enumerative pattern were correct inductive arguments, we could conclude that all emeralds (future as well as past) are grue. But this means that emeralds existing after A.D. 2000 will be blue. For to be grue after A.D. 2000 is to be blue.") Nevertheless Swinburne argues that we generally agree on which criteria to use when assessing inductive inferences; for example, that "in certain respects (e.g. greenness as opposed to grueness) things outside the spatio-temporal region which we have observed continue to behave as they have been observed to behave within that region is basic to inductive inference", p.8.

2.3 The justification of induction

For all is a woven web of guesses. Xenophanes

An issue which is of central interest to philosophers is the justification of induction. This issue was first raised by Hume [7] pp.36-38, in his famous skeptical argument aimed at enumerative induction. Essentially this runs as follows: such arguments arise because we observe uniformities in nature–e.g. that all observed emeralds have been green–and as such they are based on experience. However, we have no grounds for assuming that nature will continue to behave uniformly, other than the appeal to experience. But "[i]f there be any suspicion that the course of nature may change, and that the past may be no rule for the future, all experiences become useless, and can give no rise to inference or conclusion. It is impossible, therefore, that any argument from experience can prove this resemblance of the past to the future; since all these arguments are founded on the supposition of that resemblance"; pp. 37-8. As things do not always behave as we expect, "[w]hy may [this] not happen always, and with regard to all objects? What logic, what process of argument secures you against this suspicion?", p.38. Swinburne continues "Hume's answer is that we have no justification for believing that things will continue to behave as they have behaved. We do believe this and act on this suspicion as a matter of animal habit, but there is no justification for our doing so", p.9. As Russell puts it: "[t]he mere fact that something has happened a certain number of times causes animals and men to believe that it will happen again. Thus our instincts certainly cause us to believe that the sun will rise tomorrow, but we may be in no better position than the chicken which unexpectedly has its neck wrung", p.21.

Various justifications for induction have been proposed, see for example the essays in [12]. The deductive or "pragmatic" justification seeks to show that if any method of inference whatever fulfils the knowledge-extending function, then the method of induction will do so. The claim is not that induction will succeed in establishing true conclusions on the basis of true premisses, nor is it to say that induction is the only method which will. The thesis is that induction will succeed if any method will, and this is more than can be said for any other method. This approach is attractive because it appeals to the idea of correct reasoning methods. One drawback is that it is difficult to carry the argument through; see, for example, the articles in Chapters V and VI of [12]. Another drawback seems to be that the form of an inductive argument is often insufficient justification for its success; for example (anticipating the next section) two conflicting default rules may have the same form. The inductive or "predictionist" justification of induction seeks to justify inductive arguments on the basis of their success in the past. Care needs to be taken to avoid circularity – see, for example, Chapter VIII of [12] – however, it seems that this can be done by distinguishing between object-level inductive arguments and their meta-level justifications. The attraction of this approach is the appeal to the usefulness or otherwise of an argument. My feeling is that it is possible to give a justification of induction by means of a combination of these two accounts: an inductive argument is justified because it has proved successful in the past, and this will only be partly due to its form. One can adopt a Darwinian perspective on this view. In the evolution of natural (and artificial) intelligence, the point of inference is to extend knowledge. Agents which reason successfully survive, while those which don't do not; c.f. evolutionary algorithms in AI. Note that such inferences need not be conscious. Indeed, Nietzsche suggests that consciousness may be a hindrance: "[c]onsciousness is the last and latest development of the organic and consequently the most unfinished. From consciousness there proceed countless errors which cause an animal, a man, to perish earlier than necessary …If the preservative combination of the instincts were not incomparably stronger, if it did not in general act as a regulator, mankind must have perished through its perverse judgements and waking phantasies, its superficiality and credulity, in short through its consciousness" [6].

3 Induction in Artificial Intelligence

While philosophers have largely been concerned with the justification of inductive reasoning, AI researchers are largely concerned with its use. Apart from specific applications, such as learning, there is the more general goal of representing commonsense reasoning of the kind discussed in the previous section. This has lead to the formalisation of various forms of inductive reasoning in nonmonotonic logics.

3.1 Induction and nonmonotonic logic

Nonmnonotonic logics, such as Circumscription [9] and Default Logic [10], have been widely used to formalise inductive arguments; for example the Tweety argument and the commonsense laws of change and inertia.

At first sight it seems that Circumscription can also be used to formalise enumerative induction. If in some axiom A the objects which satisfy some predicate P also satisfy Q, then we can infer from the circumscription of P in A, that all P's are Q's; for example, if A is $P(a) \wedge P(b) \wedge Q(a) \wedge Q(b)$, then from the circumscription of P in A we can conclude that $\forall x(P(x) \supset Q(x))$. The drawback with this approach is that the conclusion that all P's are Q's only follows because the circumsciption of P in A yields the conclusion that the objects which satisfy P in A are the only objects which do so; thus in the example the circumscription gives $\forall x(P(x) \supset x = a \vee x = b)$. But this conclusion defeats the point of using enumerative induction – that P's observed in the future are likely to be Q's because the P's observed to date have been – by ruling out the possiblity of there being other P's. From a technical point of view the intended effect can be achieved by combining circumscription with belief revision [4]. For example, if T is the set of sentences resulting from circumscribing P in the above example, then in computing $T * P(c) \wedge c \neq a \wedge c \neq b$, $\forall x(P(x) \supset x = a \vee x = b)$ is removed, $\forall x(P(x) \supset Q(x))$ is retained, and so $P(c)$ is inferred. Thus circumscription produces the rule, and belief revision uses it as long as it is consistent to do so. However it may be objected that this formalisation of enumerative induction is unacceptable from a psychological point of view as the circumscription step involves a suspension of belief in there being further P's.

3.2 A general model-theoretic framework

It is possible to use the idea of pragmatic entailment [1, 2] to give a general model-theoretic account of induction and abduction. Let $\mathcal{L}$ be a language and I be the class of all models of $\mathcal{L}$. A semantic meaning function is a function $[\![]\!]$ from $\mathcal{L}$ to $\wp I$. Intuitively, for sentence $\phi \in \mathcal{L}$, $[\![\phi]\!]$ represents the semantic, or literal, or context-independent meaning of ϕ. Then for each $\Theta \subseteq \mathcal{L}$, $[\![\Theta]\!] = \bigcap\{[\![\phi]\!] | \phi \in \Theta\}$. The semantic consequence relation induced by the (semantic) meaning function $[\![]\!]$ can then be defined in the usual way. For any $\Theta \subseteq \mathcal{L}$ and $\phi \in \mathcal{L}$, $\Theta \models \phi$ ("Θ (semantically) entails ϕ") iff $[\![\Theta]\!] \subseteq [\![\phi]\!]$. So semantic consequence is just the standard Tarkian notion. The analysis is then extended to pragmatic reasoning. A pragmatic meaning function is a function $[]$ from $\wp L$ to $\wp I$; if ϕ is a sentence of $\mathcal{L}$ then $[\{\phi\}]$ is conveniently abbreviated to $[\phi]$. Intuitively for each $\Theta \subseteq \mathcal{L}$, $[\Theta]$ represents the pragmatic, or intended, or context-dependent meaning of Θ. Then $[]$ and $[\![]\!]$ induce a pragmatic consequence relation: for any $\Theta \subseteq \mathcal{L}$ and $\phi \in \mathcal{L}$, $\Theta \approx \phi$ ("Θ pragmatically entails ϕ") iff $[\Theta] \subseteq [\![\phi]\!]$. Thus Θ pragmatically entails ϕ iff ϕ follows from the intended interpretation of Θ.

In the case of induction, we wish to capture the idea that the premisses Θ represent a partial epistemic context and that induction is the process of extending Θ in an appropriate way and then reasoning semantically. Thus Θ induces ϕ if there is a reasonable expansion $\Theta \cup \Theta'$ of Θ which (semantically) entails ϕ. Intuitively Θ' contains premisses to the effect that Θ contains all of the relevant information, and premisses reflecting the expectations which arise given Θ. In order to formalise this, we require that the pragmatic function is veridical (for each $\Theta \subseteq \mathcal{L}$, $[\Theta] \subseteq [\![\Theta]\!]$), proper (for each $\Theta \subseteq \mathcal{L}$, $[\![\Theta]\!] \not\subseteq [\Theta]$) and that it preserves consistency (for each $\Theta \subseteq \mathcal{L}$, $[\Theta] = \emptyset$ only if $[\![\Theta]\!] = \emptyset$). These restrictions ensure that $\Theta \approx \phi$ if $\Theta \not\models \phi$ and there is some consistent extension $\Theta \cup \Theta'$ of Θ such that $\Theta \cup \Theta' \models \phi$. What counts as a reasonable extension can be further constrained by imposing additional conditions on the pragmatic function, with corresponding rationality postulates on the inductive consequence relation. For example, we can require that if $[\Theta] \subseteq [\![\phi]\!]$ and $[\Theta \cup \{\phi\}] \subseteq [\![\psi]\!]$ then $[\Theta] \subseteq [\![\psi]\!]$. This ensures that the the inductive consequence relation satisfies the following property: If $\Theta \approx \phi$ and $\Theta \cup \{\phi\} \approx \psi$ then $\Theta \approx \psi$. Rationality postulates of this kind are considered by Flach [3]. Deciding on an appropriate set of postulates represents the logical problem of induction. Picking a particular consequence relation which satisfies the postulates – picking a particular pragmatic function which satisfies the corresponding conditions – represents the pragmatic problem of induction.

In the case of abduction, we wish to capture the idea that the premisses Θ represent a partial epistemic context on the basis of which we are required to explain ϕ, and that abduction is the process of extending Θ in an appropriate way in order to do so. Abduction can thus be thought of as the process demonstrating that Θ induces ϕ by finding a particular Θ' such that $\Theta \cup \Theta'$ is consistent and $\Theta \cup \Theta' \models \phi$; where, intuitively, Θ' represents the explanation (or explanations) for ϕ given Θ. Lobo and Uzcátegui [8] investigate the properties of abductive consequence relations. Given an abductive task $\Theta \approx \phi$ and explanation(s) Θ' for ϕ given Θ (that is, $\Theta \cup \Theta'$ is consistent and $\Theta \cup \Theta' \models \phi$), they define ψ to be an abductive consequence of ϕ, $\phi \approx_{\Theta} \psi$, iff $\Theta \cup \Theta' \models \psi$.

Acknowledgements

I am grateful to Murray Shanahan and to the anonymous referees for helpful comments.

REFERENCES

[1] J. Bell. Pragmatic Logics. Proceedings of KR '91, Morgan Kaufmann, 1991, pp. 50-60. Revised and extended in 1992.

[2] J. Bell. Pragmatic Reasoning; a Model-Based Theory. In M. Masuch and L. Polos (eds.), *Applied Logic, How, What and Why?* A selection of papers from the Applied Logic Conference, Amsterdam, December 1992. Kluwer Academic Publishers, Amsterdam, 1995, pp. 1-28.

[3] P. Flach. Rationality Postulates for Induction. Proceedings of the Sixth Conference on Rationality and Knowledge, TARK'96, Y. Shoham (ed.), pp. 267-281.

[4] P. Gärdenfors. *Knowledge in Flux*, MIT Press, Cambridge Mass. 1988.

[5] N. Goodman. *Fact, Fiction and Forecast* (first edition). London 1955.

[6] R.J. Hollingdale. *A Nietzsche Reader*. Penguin, London, 1977.

[7] D. Hume. *An Enquiry Concerning Human Understanding* L.A. Selby-Brigge (ed.), Oxford, 1902. First published 1751.

[8] J. Lobo and C. Uzcátegui (1996) Abductive Consequence Relations. Proceedings of Common Sense-96, T. Costello and S. Buvac (eds.), pp. 104-113.

[9] J. McCarthy. Circumscription – a form of nonmonotonic reasoning. *Artificial Intelligence* 13 1980, pp. 27-39.

[10] R. Reiter. A logic for default reasoning. *Artificial Intelligence* 13 1980, pp. 81-132.

[11] B. Russell. On Induction. In [12], pp. 19-26. First published as Ch. 6 of *The Problems of Philosophy* 1912.

[12] R. Swinburne (ed.) *The Justification of Induction*. Oxford Readings in Philosophy series, Oxford University Press, Oxford, 1974.