

Machine Guessing – I

David Miller

Department of Philosophy
University of Warwick
COVENTRY CV4 7AL UK
e-mail: dwmiller57@yahoo.com

© copyright D. W. Miller 2011–2018

ABSTRACT

According to Karl Popper, the evolution of science, logically, methodologically, and even psychologically, is an involved interplay of acute conjectures and blunt refutations. Like biological evolution, it is an endless round of blind variation and selective retention. But unlike biological evolution, it incorporates, at the stage of selection, the use of reason. Part I of this two-part paper begins by repudiating the common beliefs that Hume’s problem of induction, which compellingly confutes the thesis that science is rational in the way that most people think that it is rational, can be solved by assuming that science is rational, or by assuming that Hume was irrational (that is, by ignoring his argument). The problem of induction can be solved only by a non-authoritarian theory of rationality. It is shown also that because hypotheses cannot be distilled directly from experience, all knowledge is eventually dependent on blind conjecture, and therefore itself conjectural. In particular, the use of rules of inference, or of good or bad rules for generating conjectures, is conjectural. Part II of the paper expounds a form of Popper’s critical rationalism that locates the rationality of science entirely in the deductive processes by which conjectures are criticized and improved. But extreme forms of deductivism are rejected. The paper concludes with a sharp dismissal of the view that work in artificial intelligence, including the JSM method cultivated extensively by Victor Finn, does anything to upset critical rationalism. Machine learning sheds little light on either the problem of induction or on the role that logic plays in the development of scientific knowledge.

0 Introduction

There is an antiquated view about logic, originating perhaps with Aristotle, that reasoning in mathematics is bound by the canons of deduction, while reasoning in the empirical sciences is bound by the quite distinct canons of induction. For all their differences, however, deductive logic and inductive logic are both departments of the science of logic and are held to share three important characteristics: both are *formal* disciplines, in the sense that, for the most part, their principles do not vary with the subject matter, be it mathematical or empirical, to which they are applied; both are *creative* disciplines, in the sense that they furnish means by which we may extend our knowledge from what, if anything, is given in intuition, or in experience; and both are *regulative* disciplines, in the sense that they authorize or legitimize or justify their conclusions, endowing them with a seal of reliability and credibility, if not a guarantee of apodeictic certainty.

A paper presented at the conference J. S. MILL’S IDEAS ON INDUCTION AND LOGIC OF THE HUMANITIES AND SOCIAL SCIENCES IN THE COGNITIVE RESEARCH AND IN ARTIFICIAL INTELLIGENCE SYSTEMS, held at the Russian State University for the Humanities (URSS), Moscow, in June 2011. The central theme of the meeting was the work of Victor Finn in the theory of intelligent systems, which is briefly discussed in §7 below. A Russian translation by Delir Lakhuti has been published in *Voprosy Filosofii* 7, pp. 110–119, and 8, pp. 117–126, in 2012. My thanks are due to Ali Paya, who drew my attention to Searle (1999), and to Alain Boyer for some useful hints.

The doctrine that deductive inferences are peculiar to mathematics is now wholly discredited. What was, in my judgement, Mill's most potent contribution to logic, his argument that the conclusion of a syllogism (and, by extension, any deductive inference) only repeats, partly or wholly, what is stated in the premises, and therefore cannot be better secured by the premises than it is secured by itself (1843, Book II, Chapter III, § 1), shows that deductive logic, though admirably formal, has no creative or regulative force. Even within mathematics, deductive inferences are not demonstrations, proofs of their conclusions, but only derivations, capable of transmitting truth but not of introducing it. The term *non-demonstrative inference* is nonetheless used exclusively for deductively invalid inferences, not only those inferences traditionally called inductive but also those called abductive. It defines an area in which the doctrines mentioned in the previous paragraph have become permanent possibilities of dissension. The difficulties exposed by Goodman (1954) have revealed a high cost in maintaining that inductive logic consists of purely formal rules. While allowing that the conclusion of an inference that generalizes from instances (induction by enumeration) may be credible, indeed compelling, Hume (1738), Book I, Part III, § XII, subverted the idea that this credibility is of other than psychological significance, even if the empirical premises are beyond dispute. Although some philosophers may have freed themselves from Hume's sceptical doubts, they have not identified clear errors in his premises or his argument, and the justification of induction remains an insubstantial dream. Yet there is growing support for the idea that non-demonstrative inferences may be both rational and creative, and that a working logic of discovery provides a way of vanquishing humean scepticism. This prospect has been popularized especially by workers in the domain of artificial intelligence.

The objective of this paper is to instil disenchantment. Non-demonstrative inference is either simple guesswork, untouched by reasoning, or the application of procedures that have themselves been guessed to be congenial to the task. The role of logic in scientific discovery is the purely negative one of abetting the elimination of bad guesses. The development of learning machines, in brief, may constitute an impressive engineering feat, and may provoke metaphysical speculation, but it is devoid of methodological significance. Hume's problem is not solved by circumvention.

My own view is that Hume's logical problem of induction is best solved by renouncing the prejudice that knowledge needs justification, and embracing wholeheartedly, but not slavishly, Popper's non-justificationist *critical rationalism*. In § 1, I shall say a little about why justificationist solutions to the problem flounder, and about why the problem is not wisely disregarded. The aim of § 2 is to put a brake on the popular doctrine, embraced by Hume, that although induction is logically unsound, it is a fact of psychological life, and must be properly acknowledged. In § 3, I note that it is guesswork that is a fact of life. In § 4, I provide a summary of my version of critical rationalism, and in § 5, I distinguish critical rationalism from other positions that might well be called deductivism. For more details, and especially more defence, Miller (1994), Chapters 1–3, (2006), Chapters 3–5, and (2007b) may be consulted. I shall then look more closely at the submissions made on behalf of machine learning. In § 6, I shall show what is mistaken in Gillies's thesis that the successes achieved in the development of artificial intelligence cast a dark shadow on the critical rationalist dismissal of inductive inference; and finally, in § 7, I shall look at Finn's claims on behalf of his JSM method, which is said to offer 'a synthesis of induction, analogy and abduction' (2002). I hope to show why it too is methodologically inert.

1 The problem of induction

What Carnap, writing in 1968 about inductive logic, called 'the grand old problem of justification, discussed so much since Hume's time', is now discussed not so much by philosophers, except in the classroom. There are several pretexts for this inattention. Carnap himself declared that

‘the old problem is not the right way of asking the question’ (p. 258). He proceeded (pp. 259f.):

Many believe that the essence of inductive reasoning is to make inferences of a special kind [, so that] the demand for justification would naturally be directed at any alleged principles of inductive inference. I would not say that is wrong to regard inference making as the aim [b]ut . . . it seems preferable to take as the essential point of inductive reasoning the determination of probability values. . . . Inductive probability is related to degree of belief, as Ramsey explained [I]n inductive logic we are . . . concerned with the *rational* degree of belief. Thus we have to consider the question: ‘how can we find principles which guide us to reasonable degrees of belief?’

Austere personalist Bayesians regard even this question of Carnap’s as asking too much; what matters are only reasonable principles for the coherence of degrees of belief (or credences) and for their modification in the light of evidence. Other personalists, such as Lewis (1980), think that some non-extreme credences can be rationally evaluated. Credence management, which consists entirely of deductive moves within the theory of probability, is often, tendentiously, called *inductive inference*. But since Carnap had little time, and more personalist thinkers have no time, for inductive inferences with factual conclusions, it is fair to say that Hume’s problem has been put to bed unsolved (Miller 1994, Chapter 6, § 5). Despite much talk of the justification of belief, as in the title of Howson (2000), Bayesians have not told us how beliefs can be justified.

Popper too, of course, thought that the traditional philosophical problem of induction, that is to say, the question ‘What is the justification for the belief that the future will be (largely) like the past? Or, perhaps, What is the justification for inductive inferences?’ is ‘wrongly put’ (proem to 1971). His objection to these formulations was that they appear to assume first that the future will be (largely) like the past, which is a big assumption, and secondly that there exist inductive inferences that are in need of justification. Popper’s well known view, here endorsed, is that there exist no such inferences, and hence there is no need to agonize about justifying them. What there are instead are conjectures, or guesses, or hypotheses, which are not generated by any cognizable procedure, and refutations, which are generated deductively. Neither conjectures nor refutations are justified, or justifiable, but rationality is not the same thing as justification.

Some authors who push aside the traditional problem of induction do so even though, unlike Carnap and Popper, they seem not to know where to go next. Papineau (1995), pp. 4f., writes:

It is true that induction presents an abstract philosophical puzzle. Inductive inferences are not logically compelling, and because of this their ultimate authority is an issue of philosophical controversy. But this is a puzzle, not the start of a philosophical system. It is akin to the question, ‘How do I know there is a table in front of me?’ This is a good issue for first-year philosophy students to cut their teeth on. But outside the classroom nobody seriously doubts that we do know about tables.

Similarly scornful of the value of abstract philosophical inquiry, Searle (1999), pp. 2079f., writes:

I believe that epistemic problems, [such as] ‘How is it possible that we can have knowledge at all in the light of the various sceptical paradoxes?’, should be regarded in the same way as other such paradoxes have been regarded in the history of philosophy. Zeno’s paradoxes about space and time, for example, pose interesting puzzles, but no one supposes that we cannot seriously attempt to cross a room until we have first answered Zeno’s scepticism about the possibility of moving through space. Analogously, I believe, we should have the same attitude towards the paradoxes about the possibility of knowledge that were advanced by sceptical philosophers. That is,

these are interesting puzzles, and they provide good five-finger exercises for training young philosophers, but we should not suppose that the possibility of knowledge and understanding rests on our first being able to refute Hume’s scepticism.

It seems to have occurred to neither of these writers that there might be something valuable to be learnt from an attempt to resolve the problem of induction, just as much was learnt from attempts to resolve Zeno’s paradoxes. Both Papineau (‘science can never yield any positive findings’) and Searle (‘the scientist does not arrive at truths about nature’) attribute to Popper a depiction of science as an enterprise with no positive features. Had they interested themselves in his solution to Hume’s problem, they might have learnt how wrong they are: that science consists of both conjectures and refutations, and whereas the latter are negative, the former are affirmative — not in the authoritarian sense of being justified, which both Papineau and Searle evidently see as the only standard of knowledge, but in the sense of being claimants to the truth.

According to Russell, Hume’s ‘destruction of empiricism’ rested ‘entirely on his rejection of the principle of induction’, the principle, in short, that exceptionless regularities observed profusely in the past will probably be abode by on the next occasion (1946, p. 699). Moreover,

[i]f this principle is not true, every attempt to arrive at general scientific laws from particular observations is fallacious, and Hume’s scepticism is inescapable for an empiricist. The principle itself cannot, of course, without circularity, be inferred from observed uniformities, since it is required to justify any such inference. It must therefore be, or be deduced from, an independent principle not based on experience. To this extent Hume has proved that pure empiricism is not a sufficient basis for science.

But Russell was mistaken here. It was not the weakness of empiricism that Hume exposed, but the presumption of justificationism, the idea that the only respectable way to ‘arrive at general scientific laws’, or other general statements, is by inference, rather than by guesswork. True or not, empirical or metaphysical, the principle of induction, and any ‘independent principle’ from which it is deduced, remain unjustified, and cannot provide justification in their turn. We must look elsewhere for the rational element in science, since nothing can give it ‘a sufficient basis’. A non-justificationist theory of rationality is needed, and this is what critical rationalism supplies.

A recent example of the strategy recommended by Russell is Zahar (2007), Chapter II, who introduces a (metalinguistic) principle of induction, admitted to be both ‘unjustifiable and uncriticizable’, that, he says, ‘allows us to assess science’ (p. 19). No doubt fundamentalist religious zealots can suggest alternative principles of assessment. Zahar reports Popper’s position, as others do, through the eyes of someone who does not know that critical rationalism has no use for justification. Discussing Popper’s approach in (1971), § 9, to the problem of how scientific theories are used in practice, but without reference to the improved treatment in (1974b), § 14, Zahar infers from Popper’s declaration that ‘*there can be no good reasons . . . for expecting that it [a theory] will in practice be a successful choice*’ (§ 9) that, according to Popper, ‘non-rationality has to be admitted into the sciences . . . at the level of technology’ (p. 10). It is remarkable that Popper’s appeal to the non-justificationist theory of rationality sketched in Chapter 24 of his (1945), could be admonished in 2007, as it is by Zahar (p. 15), for being ‘somewhat disingenuous’.

2 Why learning by repetition is impossible

What is the source of the general ideas, the universal hypotheses, that individual agents are in possession of? There seem to me to be four possible answers: (i) they are prompted directly by perceptual experience; (ii) they are inherited genetically, or suggested by teachers; (iii) they are

the product of pure guesswork; and (iv) they are inferred from more primitive ideas (perhaps themselves originated by processes (i)–(iii)). It should be evident that (ii) cannot be the whole answer, and I shall not consider it apart from (i). My purpose in this section is to confound (i).

The empiricist doctrine that we learn wholly from experience has to be rejected; not because of its emphasis on experience, about which I shall say something more in a moment, but because of the direction implicit in the phrase ‘from experience’. Experience is not given in experience; it is not another name for interaction with something else. That much is familiar, and is by no means controversial. Kant decided that at least the form of our experience, though not its content, is given prior to experience. A more modern version of this doctrine has it that our sense organs, being largely pre-formed, impose not only some form, but also much content, on what is experienced. The frog, who sees only moving objects, is frequently cited as an example.

A more biologically informed empiricism may therefore hold that although our knowledge is not fully explicable in terms of our encounters with the external world (even supposing that there is a sharp distinction between what is internal and what is external), it is fully explicable by the joint action of our experiences and our inheritance. Inherited knowledge delivers expectations, universal hypotheses, generalizations, statements of regularities (including statistical regularities); experience delivers singular statements, interpretations of experience in accordance with those expectations. Working together, these sources could perhaps provide enough knowledge for survival and reproduction, enough knowledge to maintain life at a low and almost mechanical level. But it is incredible to suppose that all human knowledge of regularities is inherited knowledge. Even when experience is recognized not to be the unstructured starting point once imagined, empiricism is driven to invent some mode of generalization from perceptual experience.

This is induction, still popularly thought to be an essential factor in our learning about the world. It is admitted, of course, that induction, whatever psychological process it is, is a fallible process, and that what is learnt by induction may be wholly incorrect. Logically the situation is fairly clear: induction, whatever it is, is invalid. What is much less clear is what induction is.

The difficulty in making sense of induction as a mode of generalization from experience is a difficulty in working out what exactly it is that is being generalized. An experience may be a singular event in space-time, but that is not something simple; on the contrary, it is so complex that there is no saying what might be a repetition of it (Popper 1957a, § IV). There are any number of ways of generalizing from a particular experience or set of experiences, each yielding different results, and it is therefore absurdly tendentious to represent induction as a determinate method of generalizing from experience. Induction, that is to say, is not a specific process, but at best a label for innumerable unrelated acts of generalization (as the problem of curve fitting, and Goodman *ibidem*, suggest). It is apparent too that even when the matter of its validity is not an issue, the usual problem concerning the status of induction does not entirely go away. Induction cannot be a skill learnt from experience, which means that it is an inherited skill, which means that there must at least exist in most organisms an inherited disposition to generalize, but not to generalize in any particular way (Popper 1974a, § 10). Organisms, that is, inherit not only expectations that express regularities; they inherit also an expectation of unspecified regularity.

If this is so, then induction, thought of as a mode of learning rather than as a mode of inference, is arbitrary not because it moves from the singular to the universal but because it moves from the existential to the singular. Let a be some action that an agent performs that is followed by some advantageous outcome c . How is the agent to learn from this success? If the position presented here (which has something in common with the ‘fourth [metaphysical] stage of the problem of induction’ in Popper 1983, Part I, § 5) is correct, then what is required is a transition from the expectation of regularity (which may be expressed by ‘there exists a regularity here’) and the experience of action a ’s being followed by outcome c , to the expectation

that a 's being followed by c is an instance of a particular regularity; in the formal mode of speech, 'actions of type X [such as a] are invariably followed by outcomes of type Z [such as c]'. The hard part of this process is the categorization of the action a and the outcome c as instances of the types X and Z respectively. What should be obvious is that the appropriate types, if there are any, are not learnt from the experience of a and c . Of course, some anticipation (correct or incorrect) of which types are appropriate may be inherited, and may be applied automatically. In the limiting case, there is only a response to experience. The phenomenon of imprinting, discovered by Lorenz, is a perfect example (though its irreversibility may not be characteristic).

The logical problem of induction is a problem that arises when the events under investigation are correctly (if not appropriately) sorted into types, but there is no premise that implies the existential statement 'there exists a regularity here'. There is no valid move from 'an action of type X here was followed by an outcome of type Z ' to 'actions of type X are always followed by outcomes of type Z '. Induction is logically bankrupt. The psychological (or, better, the biological) problem of induction, for its part, is best seen as a problem that arises when the initial situation is the other way round. The premise 'there exists a regularity here' is in place, and what is missing is an appropriate categorization into types of the events experienced. In situations crucial to survival and reproduction, inheritance may provide clues, but if it does not, then no profitable expectation is formed. Induction is biologically and psychologically bankrupt.

The inference from an existential statement $\exists yFy$ to an instance Fb is of course logically equivalent to the inference from the instance $\neg Fb$ to the universal statement $\forall y\neg Fy$. To avoid misunderstanding, I should stress that the inferences *existential* \rightarrow *singular* and *singular* \rightarrow *universal* considered in the last two paragraphs do not stand in the relation of equivalence to each other.

Having said all this, I must admit that the previous two paragraphs present the shortcomings of induction too ingenuously. For in both ways of looking at the problem, logical and biological, the invalid move or inference in question is not validated by a premise expressing, or an expectation of, the existence of regularities. What is needed is at least a premise to the effect that there exists a regularity connecting events of one of the mutually exclusive and effectively discriminable types X_0, \dots , of which X is one, with outcomes of one of the mutually exclusive and effectively discriminable types Z_0, \dots , of which Z is one. (For a more sensitive discussion of what is the weakest extra premise needed to validate an inductive inference, see my 1995.) The most that can be learnt from experience when the action a belongs to some particular X_i is that the outcome b belongs to some particular Z_k . What is learnt from experience is singular. Of course we learn, and to deny that experience plays a role in our learning would be fantastic. We need only consider how children learn their native language. It is uncontroversial that many of the abilities needed to learn a language are inherited, but that it is the child's early experiences that determine the particular language or languages learnt. Another explanation is therefore required of the role that experience plays in the creation of our general knowledge of the world.

It was this problem that Popper solved with what was initially a somewhat modest proposal concerning the logical problem of induction. Universal hypotheses, said Hume, cannot be derived from reports of experience. They must therefore, said Popper, precede experience, rather than follow it. But experience still has a service to perform. Universal hypotheses can be contradicted by reports of experience even if they cannot be derived from them. The role of experience is accordingly never to suggest hypotheses, which, as we have seen, it has no power to do, but to exclude them. This idea, which is logically trite, is of considerable significance. For by putting the hypotheses ahead of the experiences, we obtain a simple solution, though not effective one, to the biological problem of induction: the aspects of our experiences that are appropriately generalized are those that occur in hypotheses that will survive the later exposure to experience. This is the fundamental idea of the epistemology of trial and error, of conjectures and refutations.

When I say that the role of experience in learning is to exclude, I do not mean that experience can teach an organism to avoid mistakes, but not how to get things right. The contrast is not between the stick and the carrot, between the effectiveness of punishment and the ineffectiveness of reward. All that experience can impart is that a mistake has been made. If the organism was acting under the conjecture that such actions would be universally successful, then that conjecture may be discarded. But that is all that the organism can learn from its experience. If it is lucky, a different conjecture, a better one, may suggest itself. But only if the organism can recognize that the new conjecture is a different conjecture will it be able to improve its knowledge. Otherwise it may repeat the mistaken action, learning nothing. We see the great advantage for a species, if not for the individual, of the copying mechanism of the genetic code.

3 Invention

There is a familiar distinction, often attributed to Reichenbach, but essentially going back to Whewell (Hoyningen-Huene 1987, §II) between the discovery (better: the invention) of a hypothesis and its justification (better: its evaluation). The distinction is made, for example, in Whewell (1847), Book XI, Chapter VI, §7. We can all agree that a hypothesis cannot be evaluated before it has been invented, but it is possible that the two processes, of invention and of evaluation, might be amalgamated in self-regulating rules of inference. Mill's 'methods of experimental inquiry' are sometimes presented as rules of inductive inference of this kind, leading unerringly from carefully assembled data to true causal laws. Mill himself originally described the method of agreement as a 'mode of discovering and proving laws of nature' (1843, Book III, Chapter VIII, §1). Whewell, who disagreed often with Mill, wrote that, on the contrary, there are no such rules of discovery (*ibidem.*, Chapter II, §IV): 'Scientific discovery must ever depend upon some happy thought, of which we cannot trace the origin; — some fortunate cast of intellect, rising above all rules. No maxims can be given which inevitably lead to discovery.'

In Lecture VI.2 of *Lectures on Pragmatism*, Peirce (1903) distinguished three modes of reasoning: deduction, induction (confirmation), abduction (guesswork). In Lecture VI.4 he elaborated:

Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea; for induction does nothing but determine a value, and deduction merely evolves the necessary consequences of a pure hypothesis. . . . Deduction proves that something *must* be. Induction shows that something *actually is* operative. Abduction merely suggests that something *may* be.

He notes in the same section that the 'only justification' for using abduction is that 'from its suggestion deduction can draw a prediction which can be tested by induction', and that it is only by abduction that we shall 'understand phenomena at all'. But this explains only why abduction is useful, not why it is successful when it is successful. 'No reason whatsoever can be given for it . . . , as far as I can discover; and it needs no reason since it merely offers suggestions.'

Popper too thought that 'the act of conceiving or inventing a theory . . . seems . . . neither to call for logical analysis nor to be susceptible to it'. Empirical psychology may be interested in 'the question how it happens that a new idea occurs to a man', but it is 'irrelevant to the logical analysis of scientific knowledge' (1934, §2). He had previously identified 'the logic of scientific discovery' and 'the logic of knowledge' (*ibidem*, proem to Chapter 1), and went on to suggest that its task — 'in contradistinction to [that of] the psychology of knowledge — consists solely in investigating the methods employed in those systematic tests to which every new idea must be subjected . . .'. Despite this aversion to psychologism, in his early work *The Two Fundamental Problems of the Theory of Knowledge*, Popper, influenced by the biologically oriented approach of

Otto Selz (ter Hark 2004, 2006), had tentatively floated what he called ‘a deductivist psychology of knowledge’ (1930–1932, Chapter II, § 4). It was because it is both deductively invalid and superfluous that, like Hume, Popper excluded inductive inference from the logic of knowledge. It was because it is logically impossible (see § 2 above), that, unlike Hume, he excluded inductive learning (learning by repetition) from the psychology of knowledge (Popper, 1971, § 10; 1974a, *ibidem*). But, beginning with Whewell, many philosophers, contrary to both Hume and Popper, have denied any role to induction (generalizing from instances) in the psychology of discovery but have insisted on the indispensability of confirmation by instances in the logic of evaluation.

The view that induction plays a part in the formulation of scientific hypotheses was so outmoded 30 years ago that we read in a text for workers in AI (Mortimer, 1982, pp. 90f.) that

...inductionists do not question the hypothetico-deductive method propagated by Popper, they only express the opinion that induction is the essential element of this method. ... Popper’s arguments against inductionism usually seem to be aimed at this conception of induction, whose adherents would be hard to find today, namely, the view that induction is the process of getting to hypotheses in the heuristic sense. ... This kind of criticism can be considered as relevant only with reference to some grotesque version of inductionism. It would be difficult to quote an example of a contemporary inductionist who could be accused of psychologism. ... All known attempts at creating an inductive logic refer to the problem of justification, to the problem of the criteria for correct acceptance of hypotheses and not to the problem of where the ideas of hypotheses in the mind of a research worker come from.

This passage is an odd way of opening a chapter entitled ‘Popper’s anti-inductionism and anti-probabilism’, since the thesis that an important feature of a hypothesis is its probability relative to the available evidence belongs indubitably (though erroneously) to the stage of its evaluation, not to the moment of its creation. Mortimer’s judgement that attention to the empirical psychology of knowledge is firmly out of favour (except amongst psychologists, such as the contributors to Oaksford & Chater 2008) may, however, still be sound. For it is not to the psychology of science that today’s advocates of abduction and heuristics aspire to contribute, but to its logic.

The principal point of disagreement, I take it, between authors such as Aliseda (2006), Catton (2004), Finn (2002), (2011), Josephson & Josephson (1994), Simon (1973), and Zahar (2007), who set out to formulate some logic of discovery, and critical rationalists, who do not, is whether new hypotheses are free conjectures, or whether they are formed by some process, however dimly articulated, of inference or reasoning. These are the alternatives that were numbered (iii) and (iv) at the beginning of § 2. There are plenty of opportunities for confusion here, as we have seen, since Peirce called abduction a process of reasoning, and later in the same section used the term *guesses* to refer to scientific hypotheses (‘man ... cannot give any exact reason for his best guesses’). Let us therefore adopt the convention that inferences are constrained by rules, and guesses are not. But there is extant another serious ambiguity concerning the term *abduction*, which is sometimes used not for the process of invention, but for a process of selection amongst hypotheses that have already been formulated and scrutinized. Abduction in this sense, more often known as *inference to the best explanation*, is subsequent to empirical evaluation, and does not properly belong to any logic of discovery. Understood as inference, indeed, *abduction is abductively unsound* — just as induction is inductively unsound (Popper 1971, § 6) — since there exists a much better explanation of how our knowledge grows. This is critical rationalism.

References for Part I

- Aliseda, A. (2006). *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Dordrecht: Springer.
- Carnap, R. (1968). ‘Inductive Logic and Inductive Intuition’. In I. Lakatos, editor (1968), pp. 258–267. *The Problem of Inductive Logic*. Amsterdam: North-Holland Publishing Company.
- Catton, P. (2004). ‘Constructive Criticism’. In P. Catton & G Macdonald, editors (2004), pp. 50–57. *Karl Popper. Critical Appraisals*. London & New York: Routledge.
- Finn, V. K. (2002). ‘The Use of Induction in Plausible Reasoning in Intelligent Systems’. Abstracts, Karl Popper 2002, p. 19. Vienna: Karl Popper Institute. Reprinted as Appendix 4 of Finn (2011).
- (2011). *Artificial Intelligence. Methodology, Applications, Philosophy*. Moscow: Krasand/URSS.
- Goodman, N. (1954). *Fact, Fiction, and Forecast*. London: Athlone Press.
- ter Hark (2004). *Popper, Selz and the Rise of Evolutionary Epistemology*. Cambridge: Cambridge University Press.
- (2006). ‘The Historical Roots of Popper’s Theory of the Searchlight’. In I. C. Jarvie, K. M. Milford, & D. W. Miller, editors (2006a), pp. 37–56. *Karl Popper: A Centenary Assessment*. Volume I: *Life and Times, and Values in a World of Facts*. Aldershot & Burlington VT: Ashgate.
- Howson, C. (2000). *Hume’s Problem. Induction and the Justification of Belief*. Oxford: Clarendon Press.
- Hoyningen-Huene P. (1987). ‘Context of Discovery and Context of Justification’. *Studies in the History and Philosophy of Science* **18**, 4, pp. 501–505.
- Hume, D. (1738). *A Treatise of Human Nature*. Book I, *Of the Understanding*. London: John Noon.
- Josephson, J. R. (1994). ‘Conceptual Analysis of Abduction’. In J. R. Josephson & S. G. Josephson, editors (1994), pp. 5–29. *Abductive Inference: Computation, Philosophy, Technology*. Cambridge & elsewhere: Cambridge University Press.
- Lewis, D. K. (1980). ‘A Subjectivist’s Guide to Objective Chance’. In R. C. Jeffrey, editor (1980), pp. 263–293. *Studies in Inductive Logic and Probability*, Volume II. Berkeley & Los Angeles CA: University of California Press. Reprinted as Chapter 19 of Lewis (1987). *Philosophical Papers*, Volume II. Oxford: Oxford University Press.
- Mill, J. S. (1843). *A System of Logic, Ratiocinative and Inductive*. London, New York, & Bombay: Longmans, Green & Co. Ltd.
- Miller, D. W. (1994). *Critical Rationalism. A Restatement and Defence*. Chicago & La Salle IL: Open Court.
- (1995). ‘How Little Uniformity Need an Inductive Inference Presuppose?’. In I. C. Jarvie & N. Laor, editors (1995), pp. 143–148. *Critical Rationalism, Metaphysics and Science. Essays for Joseph Agassi*. Volume I. Dordrecht: Kluwer. Reprinted as Chapter 8 of Miller (2006).
- (2006). *Out of Error. Further Essays on Critical Rationalism*. Aldershot & Burlington VT: Ashgate.
- (2007a). ‘El modo único para aprender’. *Estudios de Filosofía* (Antioquia), 36, agosto 2007, pp. 25–39. On line at <http://www.scielo.org.co/pdf/ef/n36/n36a03.pdf/>.
- (2007b). ‘The Objectives of Science’. *Philosophia Scientiæ* **11**, 1, pp. 21–43. Revised version of §0, §1, §5, and §6 of Chapter 4 of Miller (2006). On line at <http://go.warwick.ac.uk/dwmiller/poincare.pdf/>.
- Mortimer, H. (1982). *Logika indukcji*. Warsaw: Panstwowe Wydawnictwo Naukowe. References are to the English translation 1988. *The Logic of Induction*. Chichester: Ellis Horwood Ltd.
- Oaksford, M. & Chater, N., editors (2008). *The Probabilistic Mind*. Oxford: Oxford University Press.

- Papineau, D. A. (1995). ‘Open Society, Closed Thinker’. *The Times Literary Supplement*, 23.vi.1995, pp. 4f.
- Peirce, C. S. S. (1903). ‘Lectures on Pragmatism’. In C. Hartshorne & P. Weiss, editors (1934–1935), pp. 11–131. *Collected Papers of Charles Sanders Peirce*, Volume V. Cambridge MA: Harvard University Press.
- Popper, K. R. (1930–1932). *Die beiden Grundprobleme der Erkenntnistheorie*. Tübingen: J. C. B. Mohr (Paul Siebeck). Published in 1979. English translation 2009: *The Two Fundamental Problems of the Theory of Knowledge*. Abingdon: Routledge.
- (1934). *Logik der Forschung*. Vienna: Julius Springer Verlag. Expanded English translation 1959. *The Logic of Scientific Discovery*. London: Hutchinson Educational. Repaginated Classics edition 2002. Abingdon: Routledge.
- (1945). *The Open Society and Its Enemies*. London: George Routledge & Sons. 5th edition 1966. London: Routledge & Kegan Paul. Repaginated Classics edition 2003. Abingdon: Routledge.
- (1957a). ‘Philosophy of Science: A Personal Report’. In C. A. Mace, editor (1957), pp. 155–191. *British Philosophy in the Mid-Century*. London: George Allen & Unwin Ltd. Reprinted as Chapter 1 of Popper (1963).
- (1963). *Conjectures and Refutations*. London: Routledge & Kegan Paul. 5th edition 1989. Repaginated Classics edition 2002. Abingdon: Routledge.
- (1971). ‘Conjectural Knowledge. My Solution to the Problem of Induction’. *Revue internationale de philosophie* **25**, 95–96, fasc. 1–2, pp. 167–197. References are to the reprint as Chapter 1 of Popper (1972).
- (1972). *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press. 2nd edition 1979.
- (1974a). ‘Intellectual Autobiography’. In Schilpp (1974), pp. 1–181. Reprinted separately as K. R. Popper (1976). *Unended Quest*. London: Fontana. Repaginated Classics edition 2002. Abingdon: Routledge.
- (1974b). ‘Replies to My Critics’. In Schilpp (1974), pp. 961–1197.
- (1983). *Realism and the Aim of Science*. Edited by W. W. Bartley III. London: Hutchinson Educational. Current edition. Abingdon: Routledge.
- Russell, B. A. W. (1946). *History of Western Philosophy and its Connection with Political and Social Circumstances from the Earliest Times to the Present Day*. London: George Allen & Unwin Ltd. Repaginated Classics edition 2004. Abingdon: Routledge.
- Schilpp, P. A., editor (1974). *The Philosophy of Karl Popper*. THE LIBRARY OF LIVING PHILOSOPHERS, Volume 14. La Salle IL: Open Court.
- Searle, J. R. (1999). ‘The Future of Philosophy’. *Philosophical Transactions: Biological Sciences* **354**, 1392, 29.xii.1999, pp. 2069–2080.
- Simon, H. A. (1973). ‘Does Scientific Discovery Have a Logic?’. *Philosophy of Science* **40**, 4, pp. 471–480. Reprinted as Chapter 11 of H. Keuth, editor (1998), pp. 235–248. *Karl Popper. Logik der Forschung*. Berlin: Akademie Verlag.
- Whewell, W. (1847). *Philosophy of the Inductive Sciences, Founded upon their History*. 2nd edition. London: J. W. Parker.
- Zahar, E. G. (2007). *Why Science Needs Metaphysics*. Chicago IL: Open Court.

Machine Guessing – II

David Miller

Department of Philosophy
University of Warwick
COVENTRY CV4 7AL UK
e-mail: dwmiller57@yahoo.com

© copyright D. W. Miller 2011-2018

ABSTRACT

According to Karl Popper, the evolution of science, logically, methodologically, and even psychologically, is an involved interplay of acute conjectures and blunt refutations. Like biological evolution, it is an endless round of blind variation and selective retention. But unlike biological evolution, it incorporates, at the stage of selection, the use of reason. Part I of this two-part paper begins by repudiating the common beliefs that Hume's problem of induction, which compellingly confutes the thesis that science is rational in the way that most people think that it is rational, can be solved by assuming that science is rational, or by assuming that Hume was irrational (that is, by ignoring his argument). The problem of induction can be solved only by a non-authoritarian theory of rationality. It is shown also that because hypotheses cannot be distilled directly from experience, all knowledge is eventually dependent on blind conjecture, and therefore itself conjectural. In particular, the use of rules of inference, or of good or bad rules for generating conjectures, is conjectural. Part II of the paper expounds a form of Popper's critical rationalism that locates the rationality of science entirely in the deductive processes by which conjectures are criticized and improved. But extreme forms of deductivism are rejected. The paper concludes with a sharp dismissal of the view that work in artificial intelligence, including the JSM method cultivated extensively by Victor Finn, does anything to upset critical rationalism. Machine learning sheds little light on either the problem of induction or on the role that logic plays in the development of scientific knowledge.

4 Critical rationalism

This is the appropriate moment to be more explicit about critical rationalism, and to explain why, despite the prominence it gives to explanatory hypotheses (Popper 1957b, Deutsch 2011, Chapter 1), it has no truck with abduction as a process of inference. Given how often it has been expounded, by Popper himself, by aficionados such as myself, and by antagonists, it may be thought extravagant to devote more space here to a statement of its central ideas. The truth, however, is that critical rationalism has time and again been represented as incorporating justificationist tendencies; tendencies that are unnecessary and unwelcome, despite their having some support in some of Popper's expositions (for example, in his discussions of verisimilitude in 1972, Chapter 2, § 7, and in 1983, Part I, § 2, which are discussed briefly on p. 126 of my 2006). In Chapter 2 of my (1994), I took to task many of the criticisms

A paper presented at the conference J. S. MILL'S IDEAS ON INDUCTION AND LOGIC OF THE HUMANITIES AND SOCIAL SCIENCES IN THE COGNITIVE RESEARCH AND IN ARTIFICIAL INTELLIGENCE SYSTEMS, held at the Russian State University for the Humanities (URSS), Moscow, in June 2011. The central theme of the meeting was the work of Victor Finn in the theory of intelligent systems, which is briefly discussed in § 7 below. A Russian translation by Delir Lakhuti has been published in *Voprosy Filosofii* 7, pp. 110–119, and 8, pp. 117–126, in 2012. My thanks are due to Ali Paya, who drew my attention to Searle (1999), and to Alain Boyer for some useful hints.

that, in its application to empirical science, critical rationalism (or falsificationism) cannot escape an appeal to inductive inference or to a principle of induction, and must be pronounced a failure. On the other side, some writers influenced by Lakatos have been encouraged by Popper's more conciliatory statements of his position to concoct half-baked versions of critical rationalism with rich (but to my taste, unsavoury) justificationist flavours. In equal measure opponents and proponents 'are, I feel, unconscious witnesses to ... [Popper's] originality; for they fail now ... to grasp his main point', as Popper himself wrote (1945, Chapter 2, note 2) about some of those who belittled the novelty of Heraclitus' doctrine of universal flux.

There are important secondary features, such as the recognition that good hypotheses solve genuine problems, but the spirit of critical rationalism may be condensed into three principles:

- (a) Hypotheses may be freely conjectured (as Whewell said). Even if there is some procedure for inferring hypotheses from facts, such inferences are not necessary.
- (b) Every attempt must be made to depose a conjecture by deducing from it consequences that may clash with the facts (or otherwise reveal it to be inadequate).
- (c) A conjecture that is not rejected remains accepted. No further action is needed (except further tests). Even if 'confirmation' is possible, it serves no function.

According to this position, science really does consist of not much more than positive conjectures controlled by negative refutations. Induction and abduction play no part in the feverish dialectic. The import of principle (a) is that, in the realm of invention, *induction and abduction are not needed*. The import of (b) is that, in the realm of evaluation, *deduction is needed*. The import of (c) is that, in the realm of evaluation again, *induction and abduction are still not needed*. Nothing happens to a conjecture that survives severe testing; it is neither proved nor improved nor supported. But reasoning plays a grand part — in the demolition of conjectures, not in their construction. Critical rationalism fully deserves both parts of its name: *critical* and *rationalism*.

Although the critical rationalist hopes to identify, among the available explanatory hypotheses, the one that accounts best for the phenomena in the domain under investigation, he will not infer it non-demonstratively from the empirical evidence, but will attempt to bring it into the limelight by disqualifying its competitors. The more wedded he is to empiricism, the more these disqualifications will be underwritten by empirical refutations, rather than by metaphysical objections or other critical broadsides. (Nothing in this paper should be taken as recommending an unaffected empiricism.) Should more than one explanatory hypothesis escape refutation, the critical rationalist will regard his project as unfinished. From this point of view, resorting to a non-demonstrative inference to complete the job looks like an unempirical swindle. As Russell saw, empiricism is threatened by an impatient demand for finality. What he did not clearly comprehend was that by jettisoning justification, both empiricism and rationality may be saved.

Many authors have thought that leaving the origination of conjectures to the unbridled and undisciplined imagination leaves too much to chance. Peirce, for example, asked (1903, *ibidem*):

But how is it that all this truth [in science] has ever been lit up by a process in which there is no compulsiveness nor tendency towards compulsiveness? ... Is it by chance? Consider the multitude of theories that might have been suggested. ... A physicist comes across some new phenomenon in his laboratory. ... Think of what trillions of trillions of hypotheses might be made of which one only is true. And yet after two or three or at the very most a dozen guesses, the physicist hits pretty nearly on the correct hypothesis. By chance he would not have been likely to do so in the whole time that has elapsed since the earth was solidified. You may produce this or that excellent psychological account of the matter. But let me tell

you that all the psychology in the world will leave the logical problem just where it was. . . . You may say that evolution accounts for the thing. I don't doubt that it is evolution. But as for explaining evolution by chance, there has not been time enough.

It is true enough that guesses are not random, or completely at the mercy of chance (but neither is evolution). Guesses are blind. Even this thesis, which says little more than that what is known implies nothing about what is not known, is too much for many knowledgeable people, who are often displeased to be told that the bright ideas that they have made such efforts to produce are no more than blind (though perhaps brilliant) guesses. Inexplicably they value the weight of their entrenched learning more than the sparkle of their unconstrained imagination. (Einstein 1931, p. 97, thought differently, though the passage is open to other interpretations.) New hypotheses may well be guesses, the objectors say, but they are not *mere* guesses; they are *informed* or *educated* guesses. Quine put it this way: 'Jumping to conclusions is our way of daily life. Jumping to reasonable conclusions is the busy scientist's way . . .' (1986, p. 332). But informed guesses may also be blind. What informs informed guesses are earlier guesses, no less blind, that have survived the critical investigation to which they have been exposed. That is, they are guesses that are informed by what is conjecturally known, but blind to what is unknown (that is, blind to what goes beyond earlier guesses). As Campbell declared wisely (1974, p. 422): 'In going beyond what is already known, one cannot but go blindly. If one can go wisely, this indicates already achieved wisdom of some general sort.' Some guesses may indeed turn out, on examination, to be implied by what is already known, and accordingly not to advance our objective knowledge. Even here it is a case of the blind leading the blind (Miller 2005b, p. 79).

Guesswork has no rational component. As Peirce stated boldly in the passage quoted in §3 above (p. 7): 'No reason whatsoever can be given for it[,] . . . and it needs no reason since it merely offers suggestions.' Discovery, whether scientific or unscientific, is rational only to the extent that it uses reasoning to expose the errors in these suggestions. Popper, at that time knowing little about Peirce's philosophy, said much the same in 'Back to the Presocratics' (1958), § XII: 'There is only one element of rationality in our attempts to know the world: it is the critical examination of our theories. These theories themselves are guesswork.' This fact too engenders much disquiet. Bernays (1974), § 14, for example, having quoted the same words, added that 'it seems that . . . [in these words] there is the hidden assumption that rationality must be *knowing*', and called for an extension of the role of reason beyond the critical process. In Bernays's paper I can find no substantial argument against the idea that guesswork involves no reasoning, and this is surely what Popper had in mind, Popper's reply was respectful, but he repudiated the assumption attributed to him, and conceded only that conjectures anticipate rational criticism — which has a purpose only if the target is held conjecturally (1974b, § 28.IV).

That conjectures and refutations are intricately intermingled is not, I suppose, in dispute. Where the critical rationalist position, that pure guesswork contains no element of reasoning, differs from the popular view may be only with regard to the level at which succeeding conjectures are analysed and severalized. Unconscious conjectures, condemned in summary justice, and eliminated from contention without ever being fully formulated, will be invisible at the stage at which pencil is put to paper. On this point, see Popper's remark about the ease with which Russell evidently composed his manuscripts (1978, p. 245; 1994, p. 21), his anecdotal report about Einstein's method of working (*ibidem*), and the claim that it was the 'consciously critical attitude [that] . . . made it possible for Einstein to reject, quickly, hundreds of hypotheses as inadequate before examining one or another hypothesis more carefully, if it appeared to be able to stand up to more serious criticism' (1966, § XXI). A neglect of what is never written down is the only excuse (apart from an unfamiliarity with mathematical work) for the failure to recognize

that in ‘new mathematical proofs there is first the intuitive and imaginative idea of how the proof might run; and then comes the critical testing of the various steps of the proof — a critical revision which, more often than not, reveals that the proof is not valid’ (Popper 1974b, §28.IV).

5 Exaggerated deductivism

Critical rationalism is sometimes known more briefly as *deductivism*, plainly not a bad name for a doctrine that holds that all genuine inferences are deductive. I have, however, shunned the term here, since several authors have extended the hegemony of deductive logic even further, and have construed as products of deductive reasoning even the guesses from which, according to critical rationalists, the growth of knowledge must begin. According to Worrall (1995, p. 91):

The fact is that scientists don’t simply guess their theories; they don’t make ‘bold’ Popperian conjectures. Instead they arrive at their theories in a way that, while it no doubt involves intuition and creativity, can nonetheless be reconstructed as a systematic and logical argument based on previous successes in science and parts of ‘background knowledge’ taken as premises. ... [But] even in the absence of a general analysis of theory-construction, it is not difficult to show by examining the *details of particular historical episodes* that such an analysis must exist.

Zahar (2007), Chapter VI, §(H), and Musgrave (2009), pp. 218–222, have also proposed what are represented as deductivist *heuristics* or logics of discovery. There seems to be agreement that only at the psychological level can discoveries be made in this way, since the deductive consequences of a hypothesis, viewed objectively, are already present within it. Deduction is not creative in the sense of §0 above. In his exposure of the pretensions of syllogistic reasoning Mill rightly dismissed the attempt ‘to attach any serious scientific value to such a mere salvo as the distinction drawn between being involved *by implication* in the premises, and being directly asserted in them’ (1843, Book II, Chapter III, §2). But we need not concede much to the advocates of deductive heuristics. It is unusual, to say the least, for a surprising conclusion to be uncovered by the process of deduction. What is more common is that the conclusion (or its negation) is first guessed to be a theorem, and then proved to be a theorem. A conclusion that typically turns up unexpectedly in the course of a deduction or calculation is the identity $0 = 0$.

It is a mystery to me, indeed, what the import of Worrall’s thesis is. There can be no doubt that previously unformulated hypotheses can sometimes be incorporated into a deductive network composed of ‘previous successes in science and parts of “background knowledge”’. This happens more than sometimes in modern mathematics (most of which can be assimilated into axiomatic set theory). But it cannot be the norm, since the deductive network (or one of its ancestors) must have been one of those ‘“bold” Popperian conjectures’ that have escaped Worrall’s attention. There can be no doubt either, as we have just seen, that the work needed to fit a hypothesis into the ‘structure of scientific doctrines [that] is already in existence’ (Popper 1934, preface) is more ‘trial and error, experimentation, guesswork’ (Halmos 1985, p. 321) than ‘systematic and logical argument’. The earlier doctrines control the later ones much as the banks of a river control the river’s flow. They do not direct it, but they do redirect it. I do not deny the historical and scientific value of sorting out connections in this manner, but I resist the conclusion that the deductive organization of the finished article reveals the logic of its discovery.

The educational cult known as *critical thinking*, or *informal logic*, is another activity where an excessive effort is often made to impose a deductive structure on diffuse segments of conversation and writing. The main task allotted to students of ‘critical thinking’ (I feel that the name must be permanently confined within deprecatory scare quotes) is the identification and repair

of what are regarded as bad arguments, which often means the identification of suppressed or omitted premises that are able to restore deductive validity. It is rarely recognized that the suspect passages may not be, objectively considered, arguments at all, even bad ones, despite their containing such markers as *if*, *so*, and *because*. Since the supposed conclusions are admitted to be badly disconnected from the supposed premises, it must be better to regard these conclusions as simple conjectures, and the argumentative flavour, if any, to be given instead by such words as *unless*, *but*, and *despite*, which indicate that there are reservations afoot and negative considerations at play. From an objective point of view, ‘critical thinking’ is therefore triply ill conceived (Miller 2005a). In the first place, it wrongly construes conjectures as the conclusions of enthymemes; in the second place, it directs critical attention away from what a conjecture says towards the spurious argument that is alleged to be mustered in its support; and in the third place it creates the impression that if only that argument were made deductively valid, all would be well. In brief, ‘critical thinking’ mistakes both the purpose and the power of criticism. After all, the only function of argumentation (except, conceivably, in mathematics) is to help us to work out what is true (or worthy or wise or prudent) and what is not. An invalid argument that seeks to justify its conclusion, however provocative, deserves to be ignored, since it tells us nothing about what we want to know, the truth value of the conclusion. Even if the other premises are not in dispute, the ‘missing premise’ that enforces validity gives no guidance, since the truth value of the conclusion is within our grasp only if this new premise is apprehended as true. Musgrave is surely right that ‘[n]ot every [invalid] argument should be counted as having a missing premise’ (*ibidem*, p.206). But only if the fruitless task of criticizing arguments is conflated with the fruitful task of criticizing their conclusions is he right that ‘[d]eductive logic will be deprived of any critical function’ if all invalid arguments are interpreted as enthymemes.

6 Enter machines

In §3 above I asked that we should use the term *inference* only for those transitions to new hypotheses that are constrained in some way by rules, and acknowledge that other transitions are best regarded as guesswork (p.8). I have been careful not to say that non-demonstrative (inductive or abductive) rules of inference cannot be formulated, since it is undeniable that they can be. Equally easily formulated are *counterinductive* rules, which permit us to infer from a report that all encountered instances of *X* have been followed by *Z* to the conclusion that no future instances of *X* will be followed by *Z*. An agent could of course allow himself to be guided by such inductive rules, and a machine could be designed to work according to them. We need to inquire into the question whether the feasibility of such induction machines, as they may be called, gives the lie to Popper’s view that induction exists neither logically nor psychologically. For according to Gillies (1996, p.53), reporting on ‘advances in machine learning’, this view ‘can no longer be maintained in the light of programs such as ID3 or GOLEM which do make inductive inferences based on many observations and have become a part of scientific procedure’.

Without denying the ingenuity that doubtless went into the design of these programs, and others, we ought, I think, to adopt a cautious view of the methodological import of such claims. In the first place, as Tamburrini has argued in his careful study of this question (2006), p.268:

AI investigations on learning systems do not compel one to relinquish Popper’s radical scepticism towards induction. A proper understanding of both learning-theoretic and machine learning results does not require any appeal to alleged principles of induction, which are supposed to provide partial justification for hypotheses that are effectively generated on the basis of available data by computational agents.

He goes on to say that ‘the intelligent behaviour exhibited by learning systems can be properly accounted for in terms of trial and error-correction processes’. If something that may be called induction is being used, therefore, it is not regulative in the sense of §0 above. Secondly we should notice an unavoidable feature of all approaches that introduce rules for generating hypotheses, or inferring conclusions: that is, that the hypothesis that the rule is appropriate to the situation in which it is adopted does not fall within the scope of the rule, and therefore, if not blindly guessed, must be generated or inferred by some other rule. I have noted elsewhere that the adoption of a rule for making decisions does not insulate the decision to be made from the uncertainty inherent in the quandary that makes the decision necessary. The *outcome* of the decision remains uncertain, of course, but that is not what I mean. For no less uncertain is the advisability of abiding by the rule. It is almost an axiom of Bayesian philosophy, for example, that good decision making is decision making according to Bernoulli’s rule of maximizing expected utility. Yet Bernoulli’s rule is not an incontestable ‘rule of rationality’, especially if the probabilities involved are subjective, and the decision to be guided by it can be underwritten only by the wholly conjectural judgement that it is an appropriate rule to use (Miller 1998, §1).

Quite a bit of trial and error activity is involved along the way, but the results delivered by ID3, which is ‘concerned to induce classification rules’ (Gillies, *ibidem*, p. 33), and GOLEM, which is designed to induce generalizations (pp. 41–44), are not deductively implied by the input data. These programs (and others) do indeed make inductive inferences. Each program incorporates a rule of inference, and it is reported that the inferences that have been licensed by the rule are good inferences in some agreed sense. This is not to say that the conclusions have all been true, or accurate, but (I suppose) only that they have for the most part been successful. So far there is no real problem. But if there is a hypothesis that the rule itself is a good rule that will continue, in the future and in new areas, to license good inferences, it has not been obtained by using the rule. Hume’s argument is not so easily bypassed. If there is no such hypothesis, then the successes chalked up by the rules are an interesting curiosity of history, nothing more.

The possibility of induction machines that are successful when the conditions are right was explained by Popper himself at the end of § v of (1957a), and in more detail in § 13, of Part II of (1983). Popper’s machine generates a data stream (a succession of balls of different materials and colours) and uses what he calls the *simple inductive rule* (sometimes called the *straight rule*) to infer, from the observed frequencies of various events, the probabilities of their occurrence, and universal generalizations if there are any valid ones. For example: ‘It may discover, in this way, that the probability of a steel–copper pair to be followed by a copper event, or by a steel–copper pair, or by any triplet except a steel–steel–copper triplet, is zero [T]he machine has discovered the “law” that steel events tend to occur . . . in a succession of exactly three.’ Popper notes that it is assumed in the machine’s construction that the world to be investigated is such that ‘*the simple inductive rule can be successfully applied to it*’. As we know from the game *Red or Blue* (explained in *ibidem*, Part II, § 8), this condition is not a trivial one. Popper concludes:

I never said that . . . we cannot successfully use the simple inductive rule — if the objective conditions are appropriate. But I do assert that we cannot find out by induction whether the objective conditions are appropriate for using the simple inductive rule. . . . the architects of the machine . . . must decide what constitutes its ‘world’; what things are to be its individual events; *what constitutes a repetition*.

It is of course possible to build a machine with the function of learning about the world but without any idea of the conditions under which it is supposed to work. It is impossible to evaluate the effectiveness of such a machine, unless it happens to be successful under all conditions. Induction machines are known not to be like that. If the claim to have a successful

machine is to be worth considering, it is therefore essential that the designer of the machine specify its range of application. Popper's point is that the induction machine itself cannot help us to discover the conditions under which it works (though it might help us to discover conditions under which it crashes). Behind every successful induction machine or rule of inductive inference there is a wild conjecture. The more precise we are about the necessary and sufficient conditions for successful operation, the closer we are to a deductive explanation of what the machine does.

Tamburrini made much the same diagnosis of ID3: if the '*presuppositions* ... or biases embedded in ID3 proper (which determine both the language for expressing concepts and the construction of decision trees) ... can be suitably stated in declarative form, then a concept learning algorithm such as ID3 can be redescribed as a theorem prover' (*ibidem*, §4). To this Gillies had two responses (2009, p.107). The first response, voicing unspecific doubt that ID3 can in practice be so redescribed, may be ignored. The second response was to assert that ID3,

redescribed as a theorem prover, would be much more complicated than the original ID3 presented as an inductive learning system. Why should we introduce all this unnecessary complication, which would never be adopted in practice? [If we] allow the introduction of inductive rules of inference, we get simple computer induction systems which are successful in practice. ... [If we] allow only deductive rules ... we are forced to try to transform these systems into equivalent theorem provers ... This is a difficult, probably hopeless, task which adds complexity with no practical gain.

Such a response might be expected from an engineer, concerned only with a machine's efficiency, or from an accountant, concerned only with costs and benefits. Simon (1973) too seems interested much less in epistemological issues than in the relative efficiency of various methods of processing data. A more philosophical response would be to ask why ID3 (or any other induction machine) is successful on the occasions on which it is successful. Is there something objective about the world that explains the machine's success, or is the success an artefact of the machine's design and of the way in which it is applied? It is undeniable that the phenomenal world is partly regular and partly irregular, and that one of the main tasks of science is to identify regularities lying below the surface. If workers in AI have discovered even deeper and more general regularities that explain the success attributed to learning machines, they ought to let the rest of us know.

7 The JSM method

JSM reasoning is a method of moving automatically from more or less disorganized data to an appropriate causal explanatory theory. In my brief remarks here I shall rely on the excellent summary of Burch (2000), which describes the JSM method as a method of *data fusion*, and to some extent on the more technical material in Finn (2011) and in Anshakov, Finn, & Vinogradov (2005). A detailed exposition is not needed for the simple philosophical points I wish to make.

The use of the JSM method begins with ill-defined and incomplete data about a well-defined field of entities with respect to a number of (monadic) properties. Typically, for a property P there is information in the database that some of the entities under study possess P and that others do not possess P . For other entities there is no information, and for still other entities there is conflicting information: the database says that they possess both P and $\neg P$. The first task, as I understand it, is to amplify and to correct the initial information, by extrapolating any limited generalizations that hold, and by exploiting analogies. The method proceeds to bring order to the database by what seems to be a method of trial and error, gradually plugging gaps and negotiating gluts. At some point, the procedure comes to a halt, and nothing new emerges. This marks the beginning of what Finn calls the abductive stage of the automated inquiry. The

plan is to find, for any P of interest, what should be conjectured to be a cause of P , using a method akin to Mill's method of agreement: all possible causal hypotheses are considered, and the one that fits best the description of the *maximal cause* is accepted as the explanation of P .

In the Conclusion to his (2000), Burch says of the JSM method that it 'can produce useful results on databases that contain only small amounts of information'. In Finn's words (2002, pp. 407f.), the method 'is capable of solving problems of a definite type which demand the use of automated reasoning ... [and] cannot be solved by human persons in real time. ... [It is] a synthesis of cognitive procedures — induction, analogy and abduction ... in the sense of C.S. Peirce — ... [that has been] tested experimentally in different subject domains (pharmacology, biochemistry, technical diagnostics, and sociology)'. As I understand the matter, some of these successful tests have used the JSM method in order to determine causal connections in areas that are independently well understood. But Popper's celebrated thesis that confirmation, if uncritically sought, is usually easy to find, should have warned us not to be too impressed by successes of this kind. I should like to learn more about domains where the method has been unable to unearth any causal connections, and to learn by what independent criteria it is decided whether such a result is a success or a failure. I should especially like to learn about the performance of the JSM method in unpromising cases, such as the astrological data assembled by Gauquelin and others that appear to show dependencies between the times of birth of eminent people and the fields in which they achieved eminence (Gauquelin 1988; Eysenck & Nias 1982).

Above all I should like to learn what are the conditions under which the JSM method claims to be effective. Finn lists a number of syntactic '[c]onditions of applicability for the JSM Method of AHG [automatic hypothesis generation]' (*ibidem*, p. 413). But, as explained in the previous section, if the worth of the JSM method is to be empirically evaluated, we do need rather more.

Finn, like Simon (1973) and Michalski (1983), pp. 87f. (quoted by Tamburrini *ibidem* p. 268), and many others, thinks that the philosophical problem of induction needs a radical facelift: 'the problem of induction as related to universal theories ... should be replaced by the problem of adequacy of formalized heuristics in IS [Intelligent Systems] for the corresponding classes of problems' (2002, p. 408). If this replacement problem is treated as one of pure engineering, then we may allow that some steps have been made towards its solution. But if 'the problem of adequacy of formalized heuristics in IS' is a philosophical problem about how we can come to obtain knowledge of the world, then I do not see that any advance has been achieved. Nor do I understand why Hume's problem, which is not solved by being replaced, should be replaced; that is, why it should be allowed to drop out of sight unsolved, as has been recommended even by philosophers (as we saw in §1). One may of course live without a solution, that is, without a theory of rationality. But I do not know what those who have abandoned the search for a theory of rationality mean by such logical expressions as 'plausible reasoning' (Finn 2011, Appendix 4).

Portraits of Mill, Peirce, and Popper, the heroes of the JSM campaign, decorate the front board of Finn's book (2011). I have tried to indicate in what way Peirce's and Popper's insights have not been fully taken advantage of. As for Mill, this passage from Chapter II of *On Liberty* (1859) offers a wiser philosophy of knowledge than does the whole of *A System of Logic* (1843).

The beliefs which we have most warrant for, have no safeguard to rest on, but a standing invitation to the whole world to prove them unfounded. If the challenge is not accepted, or is accepted and the attempt fails, we are far enough from certainty still; but we have done the best that the existing state of human reason admits of; we have neglected nothing that could give the truth a chance of reaching us: if the lists are kept open, we may hope that if there be a better truth, it will be found when the human mind is capable of receiving it; and in the meantime we may rely on having attained such approach to truth, as is possible in our own day. This is the amount of certainty attainable by a fallible being, and this the sole way of attaining it.

References for Part II

- Anshakov, O. M., Finn, V. K., & Vinogradov, D. W. (2005). ‘Logical Means For Plausible Reasoning of JSM-Type’. On line at <http://www.ifispan.waw.pl/studialogica/AnshakovFinnVinogradov.pdf/>. Accessed 24.v.2011.
- Bernays, P. I. (1974). ‘Concerning Rationality’. In Schilpp (1974), pp. 597–605.
- Burch, R. W. (2000). ‘Semeiotic Data Fusion’. *Proceedings of the Third International Conference on Data Fusion*. Volume II, pp. 11–16. Paris: ISIF. Reprinted as Appendix 5, pp. 437–444, of Finn (2011). On line at <http://www.student.nada.kth.se/~tessy/RobertBurch.pdf/>. Accessed 24.v.2011.
- Campbell, D. T. (1974). ‘Evolutionary Epistemology’. In Schilpp (1974), pp. 413–463.
- Deutsch, D. (2011). *The Beginning of Infinity. Explanations that Transform the World*. London: Allen Lane (The Penguin Press).
- Einstein, A. (1931). *Einstein on Cosmic Religion and Other Opinions and Aphorisms*. New York: Covici-Friede.
- Eysenck, H. J. & Nias, D. K. B. (1982). *Astrology: Science or Superstition?*. London: Maurice Temple Smith.
- Finn, V. K. (2002). ‘The Use of Induction in Plausible Reasoning in Intelligent Systems’. Abstracts, Karl Popper 2002, p. 19. Vienna: Karl Popper Institute. References are to the reprint as Appendix 4 of Finn (2011).
- (2011). *Artificial Intelligence. Methodology, Applications, Philosophy*. Moscow: Krasand/URSS.
- Gauquelin, M. (1988). *Written in the Stars*. Wellingborough: Aquarian Press.
- Gillies, D. A. (1996). *Artificial Intelligence and Scientific Method*. Oxford: Oxford University Press.
- (2009). ‘Problem-Solving and the Problem of Induction’. In Z. Parusniková & R. S. Cohen, editors (2009), pp. 103–115. *Rethinking Popper*. Springer Science + Business Media B. V.
- Halmos, P. R. (1985). *I Want to be a Mathematician. An Automathography*. New York: Springer-Verlag.
- Keuth, H., editor (1998). *Karl Popper. Logik der Forschung*. KLASSIKER AUSLEGEN 12. Berlin: Akademie Verlag.
- Michalski, R. S. (1983). ‘A Theory and Methodology of Inductive Learning’. In R. S. Michalski, J. Carbonell, & T. M. Mitchell, editors (1983), pp. 83–134. *Machine Learning, An Artificial Intelligence Approach*. Berlin: Springer Verlag.
- Mill, J. S. (1843). *A System of Logic, Ratiocinative and Inductive*. London, New York, & Bombay: Longmans, Green & Co. Ltd.
- (1859). *On Liberty*. London: J. W. Parker & Son.
- Miller, D. W. (1994). *Critical Rationalism. A Restatement and Defence*. Chicago & La Salle IL: Open Court.
- (1998). ‘On Methodological Proposals’. In Keuth (1998), pp. 67–81.
- (2005a). ‘Do We Reason When We Think We Reason, or Do We Think?’. *Learning for Democracy* 1, 3, October 2005, pp. 57–71. On line at <http://go.warwick.ac.uk/dwmiller/1fd-.pdf>.
- (2005b). ‘A Reply to Trudy Govier and Philip Adey’. *Ibidem*, pp. 78–80.
- (2006). *Out of Error. Further Essays on Critical Rationalism*. Aldershot & Burlington VT: Ashgate.
- (2007a). ‘El modo único para aprender’. *Estudios de Filosofía* (Antioquia), 36, agosto 2007, pp. 25–39. On line at <http://www.scielo.org.co/pdf/ef/n36/n36a03.pdf/>.

Musgrave, A. E. (2009). ‘Popper and Hypothetico-Deductivism’. In S. Hartmann, editor (2009), pp. 205–234. *Handbook of the History of Logic*. Volume 10: *Inductive Logic*. Amsterdam: Elsevier B. V. On line at http://stephanhartmann.org/HHL10_Musgrave.pdf/. Accessed 5.vi.2011.

Peirce, C. S. S. (1903). ‘Lectures on Pragmatism’. In C. Hartshorne & P. Weiss, editors (1934–1935), pp. 11–131. *Collected Papers of Charles Sanders Peirce*, Volume V. Cambridge MA: Harvard University Press.

Popper, K. R. (1934). *Logik der Forschung*. Vienna: Julius Springer Verlag. Expanded English translation 1959. *The Logic of Scientific Discovery*. London: Hutchinson Educational. Repaginated Classics edition 2002. Abingdon: Routledge.

——— (1945). *The Open Society and Its Enemies*. London: George Routledge & Sons. 5th edition 1966. London: Routledge & Kegan Paul. Repaginated Classics edition 2003. Abingdon: Routledge.

——— (1957a). ‘Philosophy of Science: A Personal Report’. In C. A. Mace, editor (1957), pp. 155–191. *British Philosophy in the Mid-Century*. London: George Allen & Unwin Ltd. Reprinted as Chapter 1 of Popper (1963).

——— (1957b). ‘The Aim of Science’. *Ratio* **1**, 1, pp. 24–35. Reprinted also as Chapter 5 of Popper (1972), and as § 15 of Part I of Popper (1983).

——— (1963). *Conjectures and Refutations*. London: Routledge & Kegan Paul. 5th edition 1989. Repaginated Classics edition 2002. Abingdon: Routledge.

——— (1966) *Of Clouds and Clocks: An Approach to the Problem of Rationality and the Freedom of Man*. St Louis MO: Washington University Press. Reprinted as Chapter 6 of Popper (1972).

——— (1972). *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press. 2nd edition 1979.

——— (1974b). ‘Replies to My Critics’. In Schilpp (1974), pp. 961–1197.

——— (1978). ‘Natural Selection and the Emergence of Mind’. *Dialectica* **32**, 3–4, pp. 339–355. References are to the abridged version in D. W. Miller, editor (1985). *Popper Selections*. Princeton NJ: Princeton University Press.

——— (1983). *Realism and the Aim of Science*. Edited by W. W. Bartley III. London: Hutchinson Educational. Current edition. Abingdon: Routledge.

——— (1994). *Knowledge and the Body-Mind Problem*. Edited by M. A. Notturmo. London & New York: Routledge.

——— (1998). *The World of Parmenides. Essays on the Presocratic Enlightenment*. Edited by A. F. Petersen. London & New York: Routledge.

Quine, W. V. (1986). ‘The Sensory Support of Science’. *Discursos* (Granada). Reprinted as Chapter 25 of W. V. Quine (2008). *Confessions of a Confirmed Extensionalist*. Edited by D. Føllesdal & D. B. Quine. Cambridge MA: Harvard University Press.

Schilpp, P. A., editor (1974). *The Philosophy of Karl Popper*. THE LIBRARY OF LIVING PHILOSOPHERS, Volume 14. La Salle IL: Open Court Publishing Co.

Simon, H. A. (1973). ‘Does Scientific Discovery Have a Logic?’. *Philosophy of Science* **40**, 4, pp. 471–480. Reprinted as Chapter 11 of Keuth (1998).

Tamburrini, G. (2006). ‘Artificial Intelligence and Popper’s Solution to the Problem of Induction’. In I. C. Jarvie, K. M. Milford, & D. W. Miller, editors (2006b), pp. 265–284. *Karl Popper: A Centenary Assessment*. Volume II: *Metaphysics and Epistemology*. Aldershot & Burlington VT: Ashgate.

Worrall, J. (1995). ‘“Revolution in Permanence”: Popper on Theory-Change in Science’. In A. O’Hear, editor (1995). *Karl Popper: Philosophy and Problems*, pp. 75–102. ROYAL INSTITUTE OF PHILOSOPHY SUPPLEMENT **39**. Cambridge & elsewhere: Cambridge University Press.

Zahar, E. G. (2007). *Why Science Needs Metaphysics*. Chicago IL: Open Court.