AN EXPLORER UPON UNTRODDEN GROUND:
PEIRCE ON ABDUCTION

Stathis Psillos

> Abduction, in the sense I give the word, is any reasoning of a large class of which the provisional adoption of an explanatory hypothesis is the type. But it includes processes of thought which lead only to the suggestion of questions to be considered, and includes much besides.
>
> Charles Peirce (2.544, note)

1 INTRODUCTION

Charles Sanders Peirce (1839-1914), the founder of American pragmatism, spent a good deal of his intellectual energy and time trying to categorise kinds of reasoning, examine their properties and their mutual relations. During this intellectual adventure, he was constantly breaking new ground. One of his major achievements was that he clearly delineated a space for non-deductive, that is ampliative, reasoning. In particular, he took it to be the case that there are *three* basic, irreducible and indispensable forms of reasoning. Deduction and Induction are the two of them. The third is what he came to call *abduction*, and whose study animated most of Peirce's intellectual life.

In his fifth lecture on Pragmatism, in 1903, Peirce claimed that "abduction consists in studying facts and devising a theory to explain them" (5.145).[1] And in the sixth lecture, he noted that "abduction is the process of forming an explanatory hypothesis" (5.171). He took abduction to be the only kind of reasoning by means of which new ideas can be introduced (cf. 5.171). In fact, he also thought that abduction is the mode of reasoning by means of which new ideas have actually been introduced: "All the ideas of science come to it by the way of Abduction" (5.145). "Abduction", he added, "consists in studying facts and devising a theory to explain them. Its only justification is that if we are ever to understand things at all, it must be in that way (5.145).

---

[1] All references to Peirce's works are to his *Collected Papers*, and are standardly cited by volume and paragraph number. The Collected Papers are not in chronological order. Every effort has been made to make clear the year in which the cited passages appeared.

Peirce never doubted the reality, importance, pervasiveness and reasonableness of explanatory reasoning. And yet he thought that explanatory reasoning had been understudied—its character as a distinct logical operation had not been understood. Nor had it been sufficiently distinguished from other basic forms of reasoning. In 1902, in the middle of his unfinished manuscript *Minute Logic*, he made it clear that he was fully aware of the unprecedented character of the task he had set to himself. In his study of 'Hypothetic inference', as he put it, he was "an explorer upon untrodden ground" (2.102).

In this chapter I will narrate the philosophical tale of this exploration. Section 2 will recount Peirce's debts to Kant and Aristotle. Section 3 will articulate and present Peirce's own two-dimensional framework for the study of reasoning and set out Peirce's key aim, viz., the study of the mode of reasoning that is both ampliative and generative of new content. Section 4 explains Peirce's early syllogistic approach to inference and discusses his division of ampliative reasoning into Hypothesis and Induction. Section 5 examines Peirce's mature approach to abduction. Section 6 focuses on the issue of the legitimacy of abduction *qua* mode of reasoning and relates it to Peirce's pragmatism. Section 7 relates Peirce's conception of inquiry as a three-stage project which brings together all three basic and ineliminable modes of reasoning, viz., abduction, deduction and deduction. The chapter concludes with some observations about Peirce's legacy.

2   IDEAS FROM KANT AND ARISTOTLE

In setting out for the exploration of the untrodden ground, Peirce had in his philosophical baggage two important ideas; one came from Kant and the other from Aristotle. From Kant he took the division of all reasoning into two broad types: explicative (or necessary) and ampliative. In his *Critique of Pure Reason*, Kant famously drew a distinction between analytic and synthetic judgements (A7/B11). He took it that analytic judgements are such that the predicate adds nothing to the concept of the subject, but merely breaks this concept up into "those constituent concepts that have all along been thought in it, although confusedly". For this reason, he added that analytic judgements can also be called "explicative". Synthetic judgements, on the other hand, "add to the concept of the subject a predicate which has not been in any wise thought in it, and which no analysis could possibly extract from it; and they may therefore be entitled ampliative". Peirce (cf. 5.176) thought that Kant's conception of explicative reasoning was flawed, if only because it was restricted to judgements of the subject-predicate form. Consequently, he thought that though Kant was surely right to draw the distinction between explicative and ampliative judgements, the distinction was not properly drawn.[2] He then took it upon himself to rectify this problem.

One way in which Kant's distinction was reformed was by a further division

---

[2] For Peirce's critique of Kant's conception of analytic judgements, see his *The Logic of Quantity* (4.85-4.93), which is chapter 17 of the *Grand Logic*, in 1893.

of ampliative reasoning into Induction and Hypothesis. In fact, Peirce found in Aristotle the idea that there is a mode of reasoning which is different from both deduction and Induction. In his *Prior Analytics*, chapter 25 (69a20ff), Aristotle introduced an inferential mode which he entitled *apagōgē* and was translated into English as 'reduction'. Aristotle characterised *apagōgē* as follows: "We have Reduction (1) when it is obvious that the first term applies to the middle, but that the middle applies to the last term is not obvious, yet nevertheless is more probable or not less probable than the conclusion". This is rather opaque,[3] but the example Aristotle used may help us see what he intended to say.[4] Let A stand for 'being taught' or 'teachable'; B for 'knowledge' and C for 'morality'. Is morality knowledge? That is, is it the case that C is B? This is not clear. What is evidently true, Aristotle says, is that knowledge can be taught, i.e., B is A. From this nothing much can be inferred. But if we hypothesise or assume that C is B (that morality is knowledge), we can reason as follows:

> C is B
> B is A
> Therefore, C is A.

That is: Morality is knowledge; Knowledge can be taught; therefore, morality can be taught. If the minor premise (C is B; morality is knowledge) is not less probable or is more probable than the conclusion (C is A; morality can be taught), Aristotle says, we have *apagōgē* : "for we are nearer to knowledge for having introduced an additional term, whereas before we had no knowledge that [C is A] is true". The additional term is, clearly, B and this, it can be argued, is introduced on the basis of explanatory considerations.[5]

In the uncompleted manuscript titled *Lessons from the History of Science* (c. 1896), Peirce noted that "There are in science three fundamentally different kinds of reasoning, Deduction (called by Aristotle {synagōgē} or {anagōgē}, Induction (Aristotle's and Plato's {epagōgē}) and Retroduction (Aristotle's {apagōgē}, but misunderstood because of corrupt text, and as misunderstood usually translated abduction)" (1.65).

Peirce formed the hypothesis that Aristotle's text was corrupt in some crucial respects and that Aristotle had in fact another kind of inference in mind than the one reconstructed above (and was also acknowledged by Peirce himself; cf. 7.250-7.252).[6] He took it that Aristotle was after an inference according to which the

---

[3]There is a second clause in Aristotle's definition, but it need not concern us here.

[4]There is actually some controversy over the exact rendering of Aristotle's text and the interpretation of 'reduction'. For two differing views, see W. D. Ross (1949, 480-91) and Smith (1989, 223-4).

[5]Ross [1949, p. 489] claims that *apagōgē* is a perfect syllogism and works on the assumption that if a proposition (which is not known to be true) is admitted, then a certain conclusion follows, which would not have followed otherwise. Smith [1989, p. 223] notes that *apagōgē* "is a matter of finding premises from which something may be proved".

[6]Here is how he explained things in his fifth lecture on Pragmatism in 1903: "[I]t is necessary to recognize three radically different kinds of arguments which I signalized in 1867 and which

minor premise (Case) of a syllogism is "inferred from its other two propositions as data" (7.249), viz., the major premise (Rule) and the conclusion (Result). Peirce took it that the proper form of hypothetical reasoning (the one that, according to Peirce, Aristotle was after in Prior Analytics, chapter 25) must be:[7]

Rule — M is P
Result — S is P
Case — S is M

which amounts to a re-organisation of the premises and the conclusion of the following deductive argument:

Rule — M is P
Case — S is M
Result — S is P.

But then it transpired to Peirce that there is yet another re-organisation of the propositions of this argument, viz.,

Case — S is M
Result — S is P
Rule — M is P

which, in his early period, he took it to characterise induction.

We shall discuss the details of Peirce's account of ampliative inference in the sequel. For the time being, let us just rest with the note that Peirce's studies of ampliative reasoning were shaped by his re-evaluation and critique of ideas present in Kant and Aristotle. As we are about to see in the next section, Peirce created his own framework for the philosophical study of reasoning, which was essentially two-dimensional.

3   PEIRCE'S TWO-DIMENSIONAL FRAMEWORK

In one of his last writings on the modes of reasoning, a letter he sent to Dr Woods in November 1913, Peirce summed up the framework within which he examined

---

had been recognized by the logicians of the eighteenth century, although [those] logicians quite pardonably failed to recognize the inferential character of one of them. Indeed, I suppose that the three were given by Aristotle in the Prior Analytics, although the unfortunate illegibility of a single word in his MS. and its replacement by a wrong word by his first editor, the stupid [Apellicon], has completely altered the sense of the chapter on Abduction. At any rate, even if my conjecture is wrong, and the text must stand as it is, still Aristotle, in that chapter on Abduction, was even in that case evidently groping for that mode of inference which I call by the otherwise quite useless name of Abduction—a word which is only employed in logic to translate the [{apagōgē}] of that chapter" (5.144).

[7]Peirce said of Aristotle: "Certainly, he would not be Aristotle, to have overlooked that question [whether the minor premise of a syllogism is not sometimes inferred from its major premise and the conclusion]; and it would no sooner be asked than he would perceive that such inferences are very common" (7.249). In 8.209 (c. 1905), Peirce expressed doubts concerning his earlier view that the text in chapter 25 of *Prior Analytics* was corrupt.

reasoning and its properties. As he explained (8.383-8.388), there are two kinds of desiderata or aims that logicians should strive for when they study types of reasoning: uberty and security. Uberty is the property of a mode of reasoning in virtue of which it is capable of producing extra content; its "value in productiveness", as he (8.384) put it. Security is the property of a mode of reasoning in virtue of which the conclusion of the reasoning is at least as certain as its premises. These two desiderata delineate a two-dimensional framework within which reasoning is studied. Peirce's complaint is that traditional studies of reasoning have focused almost exclusively on its "correctness", that is on "its leaving an absolute inability to doubt the truth of the conclusion as long as the premises are assumed to be true" (8.383). By doing so — by focusing only on security — traditional approaches have tended to neglect non-deductive modes of reasoning. They have confined their attention to deduction, which is the only mode of reasoning that guarantees security. This one-dimensional approach has, however, obscured the fact that there are different types of *reasoning*, which have different forms and require independent and special investigation. Induction and abduction are such types of reasoning. What is more, together with deduction they constitute the three ultimate, basic and independent modes of reasoning. This is a view that runs through the corpus of the Peircean work.

Peirce's two-dimensional framework suggests a clear way to classify the three modes of reasoning. Deduction scores best in security but worst in uberty. Abduction scores best in uberty and worst in security. Induction is between the two (cf. 8. 387). But why is uberty needed? One of Peirce's stable views was that reasoning should be able to generate new ideas; or new content. The conclusion of a piece of reasoning should be able to be such that it exceeds in information and content whatever is already stated in the premises. Deduction cannot possibly do that. So either all attempts to generate new content should be relegated to processes that do not constitute reasoning or there must be reasoning processes which are non-deductive. The latter option is the one consistently taken by Peirce. The further issue then is the logical form of non-deductive reasoning. Peirce was adamant that there are two basic modes of non-deductive reasoning. Throughout his intellectual life he strove to articulate these two distinct modes, to separate them from each other and to relate them to deductive reasoning. 'Abduction', 'Retroduction', 'Hypothetic Inference', 'Hypothesis' and 'Presumption' are appellations aiming to capture a distinct mode of reasoning — distinct from both induction and deduction. What makes this kind of reasoning distinctive is that it relies on explanation — and hence it carves a space in which explanatory considerations guide inference. In his letter to Dr Woods, he said: "I don't think the adoption of a hypothesis on probation can properly be called induction; and yet it is *reasoning* and though its *security* is low, its *uberty* is high" (8.388).

What exactly is reasoning? In his entry on ampliative reasoning in the *Dictionary of Philosophy and Psychology* (1901-2), Peirce wrote:

> Reasoning is a process in which the reasoner is conscious that a judgment, the conclusion, is determined by other judgment or judgments,

the premisses, according to a general habit of thought, which he may
not be able precisely to formulate, but which he approves as conducive
to true knowledge. By true knowledge he means, though he is not usu-
ally able to analyse his meaning, the ultimate knowledge in which he
hopes that belief may ultimately rest, undisturbed by doubt, in regard
to the particular subject to which his conclusion relates. Without this
logical approval, the process, although it may be closely analogous to
reasoning in other respects, lacks the essence of reasoning (2.773).

This is instructive in many respects. Reasoning is a directed movement of
thought (*from* the premises *to* the conclusion) that occurs according to a rule,
though the rule (what Peirce calls "a general habit of thought") might not be ex-
plicitly formulated. Actually, it is the task of the logician, broadly understood, to
specify, and hence make explicit, these rules. There is more however to reasoning
than being a directed movement of thought according to a rule. The rule itself
— the general pattern under which a piece of reasoning falls — must be truth-
conducive. Or at least, the reasoner should hold it to be truth-conducive. Here
again, it is the task of the logician to show how and in virtue of what a reasoning
pattern (a rule) is truth-conducive. Peirce puts the matter in terms of knowledge:
reasoning should lead to knowledge (and hence to truth). This is important be-
cause the reasoning process confers justification on the conclusion: it has been
arrived at by a knowledge-conducive process. But Peirce's account of knowledge
is thoroughly pragmatic. Knowledge is not just true belief — justification is also
needed. And yet, a belief is justified if it is immune to doubt; that is, if it is such
that further information relevant to it will not defeat whatever warrant there has
been for it.

If justification amounts to resistance to doubt, it is clear that there are two broad
ways in which a process of reasoning can confer justification on a belief. The first
is by making it the case that if the premises are true, the conclusion has to be
true. The second is by rendering a belief plausible and, in particular, by making
a belief available for further testing, which — ideally at least — will be able to
render this belief immune to revision. The security of deductive reasoning (and the
justification it offers to its conclusions) are related to the fact that no new ideas are
generated by deduction. But no new ideas are generated by induction either (cf.
5.145). Induction, understood as *enumerative* induction, generalises an observed
correlation from a (fair) sample to a population. It moves from 'All observed *A*s
have been *B*' to 'All *A*s are *B*'. This is clearly non-demonstrative reasoning; and it
is equally clearly ampliative, or content-increasing. But it is true that no new ideas
are generated by this kind of reasoning. The reason is simple. The extra content
generated by induction is simply a generalisation of the content of the premises; it
amounts to what one may call 'horizontal extrapolation'. Enumerative induction,
pretty much like deduction, operates with the principle 'garbage in, garbage out':
the descriptive vocabulary of the conclusion cannot be different from that of the
premises. Hence with enumerative induction, although we may arguably gain
knowledge of hitherto unobserved correlations between instances of the attributes

involved, we cannot gain 'novel' knowledge, i.e., knowledge of entities and causes that operate behind the phenomena. Peirce was quite clear on this: "[Induction] never can originate any idea whatever. No more can deduction. All the ideas of science come to it by the way of Abduction" (5.145).

For Peirce then there must be a mode of reasoning which is both ampliative *and* generates new ideas. How this could be possible preoccupied him throughout his intellectual life. The key idea was that new content is generated by explanation — better put, explanatory reasoning (viz., reasoning that is based on searching for and evaluating explanations) is both ampliative and has the resources to generate new content, or new ideas. But in line with his overall approach to reasoning, it has to have a rather definite logical form. What then is this form? Peirce changed his mind over this at least once in his life. The reason is that his first attempt to characterise explanatory reasoning was constrained by his overall syllogistic conception of inference. This conception did not leave a lot of room for manoeuvre when it came to the formal properties of reasoning. A rather adequate formal characterisation of explanatory reasoning required a broadening of Peirce's conception of logic, as the logic of inquiry. In fact, in 1882, Peirce came to see logic as "the method of methods", or "the art of devising methods of research" (7.59). In his later writings, he was more inclined to equate logic with the scientific method, broadly understood as "the general method of successful scientific research" (7.79).

We are about to see all this in detail in the next section. The general point, if you wish, is that Peirce's view of explanatory reasoning has gone through two distinct phases. In the first phase, he took explanatory reasoning to be a specific inferential pattern that stands on its own and is meant to capture the formation and acceptance of explanatory hypotheses. In the second phase, explanatory reasoning was taken to be part of a broader three-stage methodological pattern (the method of inquiry). Explanatory reasoning no longer stands on its own; and though Peirce's later view of explanatory reasoning shares a lot with his earlier view, (for instance the thought that explanatory reasoning is the sole generator of new content), the key difference is that explanatory reasoning yields conclusions that need further justification, which is achieved by means of deduction and induction. Abduction — Peirce's settled term for explanatory reasoning — leads to hypotheses with excess/fresh content. These hypotheses do not bear their justification on their sleeves. They need to be further justified — by deduction of predictions and their confirmation (which is Peirce's settled view of induction), because it is only this *further* process that can render them part of a body of beliefs that though fallible cannot be overturned by experience. Abductively generated beliefs should be subjected to further testing (encounters with possibly recalcitrant experience) and if they withstand it successfully (especially in the long run) they become true — in the sense that Peirce thinks of truths as doubt-resistant (or permanently settled) beliefs (truth: "a state of belief unassailable by doubt" 5.416).[8]

---

[8]Two of Peirce's commentators Arthur Burks [1946, p. 301] and K. T. Fann [1970, pp. 9-10] have rightly contrasted the two phases of Peirce's views of explanatory reasoning along the following lines: in the first phase, Hypothesis is an evidencing process, while in the second phase

4 HYPOTHESIS VS INDUCTION

"The chief business of the logician", Peirce said in 1878, "is to classify arguments; for all testing clearly depends on classification. The classes of the logicians are defined by certain typical forms called syllogisms" (2.619). As noted already in section 2, given this syllogistic conception of argument, typically exemplified by Barbara (S is M; M is P; therefore S is P), there is a division of ampliative types of argument into *Induction* and *Hypothesis*. Deduction is captured by a syllogism of the form:

D: {All *A*s are *B*; *a* is *A*; therefore, *a* is *B*}.

There are two re-organisations of the premises and the conclusion of this syllogism:

I: {*a* is *A*; *a* is *B*; therefore All *A*s are *B*}; and
H: {*a* is *B*; All *A*s are *B*; therefore *a* is *A*}.

Here is Peirce's own example (2.623)

DEDUCTION
Rule. — All the beans from this bag are white.
Case. — These beans are from this bag.
∴ Result. — These beans are white.

INDUCTION
Case. — These beans are from this bag.
Result. — These beans are white.
∴ Rule. — All the beans from this bag are white

HYPOTHESIS
Rule. — All the beans from this bag are white.
Result. — These beans are white.
∴ Case. — These beans are from this bag.

The crucial thing here is that Peirce took both *I* and *H* to be formal argument patterns, which characterise "synthetic" reasoning. So, making a hypothesis falls under an inferential pattern. Peirce said:

Suppose I enter a room and there find a number of bags, containing different kinds of beans. On the table there is a handful of white beans; and, after some searching, I find one of the bags contains white beans only. I at once infer as a probability, or as a fair guess, that this handful was taken out of that bag. This sort of inference is called *making an hypothesis.*

---

it is a methodological as well as an evidencing process.

The intended contrast here, I take it, is with the statement "These beans are from the white-beans bag" (the conclusion of H) being a *mere* guess, or a wild conjecture and the like. Though it is clear that the conclusion of H does not logically follow from the premises (this is simply to say that H is not D), it is a *conclusion*, that is, the result of an inferential process — a movement of thought according to a rule. Already in 1867, in *On the Natural Classification of Arguments*, (Proceedings of the American Academy of Arts and Sciences, vol. 7, April 9, 1867, pp. 261-87), he took it to be the case that the adoption of a hypothesis is an inference "because it is adopted for some reason, good or bad, and that reason, in being regarded as such, is regarded as lending the hypothesis some plausibility" (2.511, note). More generally, although both H and I are logically invalid, they are not meant to be explicative inferences but ampliative. Their conclusion is adopted on the basis that the premises offer *some* reason to accept it as plausible: were it not for the premises, the conclusion would not be considered, even prima facie, plausible; it would have been a *mere* guess.

The difference between I and H is, as Peirce put it, that induction "classifies", whereas hypothesis "explains" (2.636). Classification is not explanation and this implies that Induction and Hypothesis are not species of the same genus of ampliative reasoning, viz., explanatory reasoning. Induction is a more-of-the-same type of inference.[9] As noted already, the conclusion of an induction is a generalisation (or a rule) over the individuals mentioned in the premises. Hypothesis (or hypothetical inference) is different from induction in that the conclusion is a hypothesis which, if true, *explains* the evidence (or facts) mentioned in the premises. Here is how Peirce put it: "Hypothesis is where we find some very curious circumstance, which would be explained by the supposition that it was a case of a certain general rule, and thereupon adopt that supposition" (2.624). The supposition is adopted for a reason: it explains 'the curious circumstance'. The centrality of explanation in (at least a mode of) ampliative reasoning is a thought that Peirce kept throughout his intellectual engagement with the forms of reasoning.

A key idea that Peirce had already in the 1870s[10] was that Hypothesis is different from Induction in that the conclusion of H is typically a new kind of fact, or something "of a different kind from what we have directly observed, and frequently something which it would be impossible for us to observe directly" (2.640). The excess content (the new ideas) generated by Hypothesis concerns, in a host of

---

[9]In his seventh of his Lowell lectures in 1903, Peirce developed a rather elaborate theory of the several types of induction (cf. 7.110-7.130). He distinguished between three types of induction: Rudimentary or crude induction, which is a form of default reasoning, viz., if there is no evidence for A, we should assume that A is not the case; Predictive induction, where some predictions are being drawn from a hypothesis and they are further tested; and Statistical or Quantitative induction, where a definite value is assigned to a quantity, viz., it moves from a quantitative correlation in a sample to the entire class "by the aid of the doctrine of chances". Peirce goes on to distinguish between several subtypes of induction. The characteristic of all types is that their justification is that they will lead to convergence to the truth in the limit. Rudimentary induction is the weakest type, while statistical induction is the strongest one. See also his *The Variety and Validity of Induction*, (2.755-2.760), from manuscript 'G', c. 1905.

[10]In a series of six articles published in *Popular Science Monthly* between 1877 and 1878.

typical cases, unobservable entities that cause (and hence causally explain) some observable phenomena. Indeed, Peirce took this aspect of hypothetical reasoning to be one of the three reasons which suggest that Hypothesis and Induction are distinct modes of ampliative reasoning. As he put it in 1878: "Hypothetic reasoning infers very frequently a fact not capable of direct observation" (2.642). Induction lacks this capacity because it is constrained by considerations of similarity. Induction works by generalisation, and hence it presupposes that the facts mentioned in the conclusion of an inductive argument are similar to the facts mentioned in the premises. Hypothesis, on the other hand, is not constrained by similarity. It is perfectly possible, Peirce noted, that there are facts which support a hypothetically inferred conclusion, which are totally dissimilar to the facts that suggested it in the first place. The role played by similarity in Induction but not in Hypothesis is the second reason why they are distinct modes of ampliative reasoning. As he put it: (...) the essence of an induction is that it infers from one set of facts another set of similar facts, whereas hypothesis infers from facts of one kind to facts of another" (2.642).[11]

The example Peirce used to illustrate this feature of hypothetical reasoning is quite instructive. The existence of Napoleon Bonaparte is based on a hypothetical inference. It is accepted on the basis that it accounts for a host of historical records, where these records serve as the ground for the belief in the historical reality of Napoleon. There is no way, Peirce thought, this kind of inference be turned to an induction. If there were, we would have to be committed to the view that all further facts that may become available in the future and confirm the historical reality of Napoleon will be similar to those that have already been available. But it is certainly possible, Peirce suggested, that evidence that may become available in the future might well be of a radically different sort than the evidence already available. To illustrate this, he envisaged the possibility that "some ingenious creature on a neighboring planet was photographing the earth [when Napoleon was around], and that these pictures on a sufficiently large scale may some time come into our possession, or that some mirror upon a distant star will, when the light reaches it, reflect the whole story back to earth" (2.642). There is clearly no commitment to similarity in the case of Hypothesis. Further facts that may confirm the hypothesis that Napoleon was historically real may be of any sort whatever. Actually, this thought tallies well with his claim that hypotheses can gain extra strength by unifying hitherto unrelated domains of fact (see 2.639). A case like this is the kinetic theory of gases which Peirce took it to have gained in strength by relating (unifying) a "considerable number of observed facts of different kinds" (2.639).[12]

---

[11]Another salient point that distinguishes Induction from Hypothesis is that Hypothesis does not involve an enumeration of instances (see 2.632).

[12]The third reason for the distinction between Hypothesis and Induction is somewhat obscure. He takes it that Induction, yielding as it does a general rule, leads to the formation of a habit. Hypothesis, by contrast, yielding as it does an explanation, leads to an emotion. It's not quite clear what Peirce intends to say here. The analogy he uses is this: "Thus, the various sounds made by the instruments of an orchestra strike upon the ear, and the result is a peculiar musical

From the various examples of hypothetical reasoning Peirce offers,[13] he clearly thinks that Hypothesis (meaning: hypothetical inference) is pervasive. It is invariably employed in everyday life as well as in science. But as noted already, Peirce's approach to amplitative reasoning has been two-dimensional. Hypothesis might well score high in uberty, but it scores quite low in security. It is clear that Hypothesis does give us reasons to accept a conclusion, but the reasons might well be fairly weak. Here is Peirce again: "As a general rule, hypothesis is a weak kind of argument. It often inclines our judgment so slightly toward its conclusion that we cannot say that we believe the latter to be true; we only surmise that it may be so" (2.625). The point, clearly, is that the fact that a hypothesis might explain some facts is not, on its own, a conclusive reason to think that this hypothesis is true. Perhaps, Peirce's careful wording (compare: "it often inclines our judgement so slightly ...") implies that some conclusions of hypothetical inferences are stronger than others — that is, there might be some further reasons which enhance our degree of belief in the truth of the conclusion of a hypothetical inference. Indeed, Peirce went on to offer some rules as to how "the process of making an hypothesis should lead to a probable result" (2.633). This is very important because it makes clear that from quite early on, Peirce took it that hypothetical reasoning needs some, as it were, external support. It may stand on its own as a mode of reasoning (meaning: as offering grounds or reasons for a conclusion), but its strength (meaning: how likely the conclusion is) comes, at least partly, from the further testing that the adopted hypothesis should be subjected to. The three rules are (cf. 2.634):

1. Further predictions should be drawn from the adopted hypothesis.

2. The testing of the hypothesis should be severe. That is, the hypothesis should be tested not just against data for which it is known to do well but also against data that would prove it wrong, were it false.

3. The testing should be fair, viz., both the failures and the successes of the hypotheses should be noted.

emotion, quite distinct from the sounds themselves" (2.643). "This emotion", he carries on saying, "is essentially the same thing as an hypothetic inference, and every hypothetic inference involves the formation of such an emotion". My guess, motivated by the example above, is that Peirce means to highlight the fact that Hypothesis generates beliefs with extra content or new ideas, which exceed the content of whatever beliefs were meant to explain. This might be taken to generate in the mind a *feeling* of comprehension that was not there before. In a later piece, he seemed to equate the emotion involved in hypothetical inference with the fact that the adopted explanation removes the surprise of the explanandum (cf. 7.197).

[13](A) "I once landed at a seaport in a Turkish province; and, as I was walking up to the house which I was to visit, I met a man upon horseback, surrounded by four horsemen holding a canopy over his head. As the governor of the province was the only personage I could think of who would be so greatly honored, I inferred that this was he". (B) "This was an hypothesis. Fossils are found; say, remains like those of fishes, but far in the interior of the country. To explain the phenomenon, we suppose the sea once washed over this land. This is another hypothesis". (C) Numberless documents and monuments refer to a conqueror called Napoleon Bonaparte. Though we have not seen the man, yet we cannot explain what we have seen, namely, all these documents and monuments, without supposing that he really existed. Hypothesis again. (2.625).

Peirce did also suggest — somewhat in passing — that the proper ground for hypothetical inference involves comparison and elimination of alternative hypotheses: "When we adopt a certain hypothesis, it is not alone because it will explain the observed facts, but also because the contrary hypothesis would probably lead to results contrary to those observed" (2.628). But he did not say much about this. Later on, in 1901 (in *The Logic of Drawing History from Ancient Documents*), he did say a lot more on the proper ground for explanatory inference (by then called *abduction*). For instance, he insisted that the hypothesis that is adopted should be "likely in itself, and render the facts likely" (7.202). Part of the ground for stronger hypothetical inferences (abductions) comes from eliminating alternative and competing hypotheses, that if true, would account for the facts to be explained.

One interesting issue concerns the very nature of a hypothesis. The very term 'hypothesis' alludes to claims that are put forward as conjectures or as suppositions or are only weakly supported by the evidence. Peirce was alive to this problem, but he nonetheless stuck with the term 'hypothesis' in order to stress two things: first, hypotheses are explanatory; and second, they admit of various degrees of strength.[14] In his discussion of the case of the kinetic theory of gases (2.639 ff), Peirce makes clear that the outcome of hypothetical reasoning might vary from a "pure" hypothesis to a "theory". What it is does not vary with respect to how it is adopted — it is always adopted on the basis of explanatory considerations. But it varies with respect to its explanatory power, which may well change over time. For instance, the kinetic theory of gases was a "pure hypothesis" when it merely explained Boyle's law, but it became a "theory" when it unified a number of empirical laws of different kinds and received independent support by the mechanical theory of heat. The idea here is that a hypothesis gains in explanatory strength when it unifies various phenomena and when it gets itself unified with other background theories (like the principles of mechanics). "The successful theories" as Peirce put it, "are not pure guesses, but are guided by reasons" (2.638). In a move that became famous later on by Wilfrid Sellars, Peirce stressed that hypotheses gain in explanatory strength when they do not just explain an empirical law, but when they also explain "the deviations" from the law (2.638). Empirical laws are typically inexact and approximate and genuinely explanatory hypotheses replace them with stricter and more accurate theoretical models.

Already in the 1870s, Peirce took it to be case that hypothetical reasoning is indispensable. Immediately after he noted that Hypothesis is a weak type of argument, he added: "But there is no difference except one of degree between such an inference and that by which we are led to believe that we remember the occurrences of yesterday from our feeling as if we did so" (2.625). Bringing memory

---

[14]Concerning his own use of the term 'hypothesis', he said: "That the sense in which I have used 'hypothesis' is supported by good usage, I could prove by a hundred authorities. The following is from Kant: 'An hypothesis is the holding for true of the judgment of the truth of a reason on account of the sufficiency of its consequents.' Mill's definition (Logic, Book III, Ch. XIV §4) also nearly coincides with mine" (2.511, note).

in as an instance of hypothetical inference suggests that if hypothetical inference cannot be relied upon, there is no way to form any kind of beliefs that exceed what we now perceive. Hence, belief with any content that exceeds what is immediately given in experience requires and relies upon hypothetical reasoning. The very issue of the trustworthiness of memory is tricky, but this should be clear. There is no way to justify the reliability of memory without presupposing that it is reliable. Even if an experiment were to be made to determine the reliability of memory, it should still be the case that the results of the experiments should themselves be correctly remembered before they could play any role in the determination of the reliability of memory. All this might well imply that memory (and hence hypothetical inference) is too basic a mode of belief formation either to be fully doubted or to be justifiable on the basis of even more basic inferential modes. Indeed, to say that hypothetical reasoning is weak (or low in security, as Peirce would later on put it) is not to say that it is unjustified or unjustifiable. Rather, it implies that its justification is a more complex affair than the justification of deduction. It also implies that there can be better or worse hypothetical inferences, and the task of the logician is to specify the conditions under which a hypothetical inference is good.

Given Peirce's syllogistic conception of inference, H and I clearly have different logical forms. Besides, Peirce has insisted that it is only Hypothesis that explains — that is, that is based on explanatory considerations. But, it may be argued, presented as above, the difference between H and I is rather superficial.[15] The conclusions of both H and I are hypotheses, even though H and I have different logical forms. Besides, both types of conclusion seem to be explanatory. As Peirce himself put it on one occasion already in 1878: "(...) when we make an induction, it is drawn not only because it explains the distribution of characters in the sample, but also because a different rule would probably have led to the sample being other than it is" (2.628). This is a clear point to the effect that Induction is also based on explanatory considerations and is guided by them.

More generally, that laws (law-like generalisations) are explanatory of their instances has been part of traditional view of explanation that Peirce clearly shared. Peirce took it that explanation and prediction are the two sides of the same coin. Actually, his overall conception of explanation is that it amounts to "rationalisation", that is to rendering a phenomenon rational (rationally explicable), where this rationalisation consists in finding a reason why the phenomenon is the way it is, the reason being such that were it taken into account beforehand, the phenomenon would have been predicted with certainty or high probability. Here is how he put the matter: "(...) what an explanation of a phenomenon does is to supply a proposition which, if it had been known to be true before the phenomenon presented itself, would have rendered that phenomenon predictable, if not with certainty, at least as something very likely to occur. It thus renders that phenomenon rational, — that is, makes it a logical consequence, necessary

---

[15]If I read him correctly, Nagel [1938, p. 385] makes this point, but from a different angle.

or probable" (7.192).[16] It should be obvious then that law-like generalisations do explain their instances, and in particular, that they do explain the observed correlation between two properties or factors. Peirce went as far to a argue that regularisation is a type of rationalisation (cf. 7.199), where a regularisation makes some facts less isolated than before by subsuming them under a generalisation: *why* are these *A*s *B*? because *all A*s are *B*.

If laws are explanatory and if law-like statements (*qua* generalisations) are the products of Induction, it seems that H and I are closer to each other than Peirce thought. Hence, it might be argued, both H and I are modes of generation and acceptance of explanatory hypotheses, be they about singular facts (e.g., causes) or about generalisations (e.g., All *A*s are *B*). Besides, it seems that H involves (at least in many typical cases) a law-like generalisation in its premises, since in the syllogistic guise it has been presented thus far, the claim is that what explains a certain singular fact is another singular fact *and* a general fact in tandem, viz., H: {a is B; All *A*s are B; therefore a is A}. As Peirce acknowledged: "By hypothesis, we conclude the existence of a fact quite different from anything observed, from which, according to known laws, something observed would necessarily result" (2.636).

It seems reasonable to claim that the chief difference between H and I is that Induction involves what we have called 'horizontal extrapolation', whilst Hypothesis involves (or allows for) 'vertical extrapolation', viz., hypotheses whose content is about unobservable causes of the phenomena. Indeed, as has been stressed already, the very rationale for Hypothesis is that it makes possible the generation of new content or new ideas.

It turns out, however, that if Hypothesis is constrained by its syllogistic form, it cannot play its intended role as a creator of new content. Think of H presented as above:

   H: {All *A*s are B; a is B; therefore a is A}.

The conclusion of H might well be a hypothesis, but its content is not really new: it is already contained in the major premise. So the inference does not create new content; rather it unpacks content that is already present in the premises. The very syllogistic character of H leaves no choice here: premises and conclusion must share vocabulary; otherwise the conclusion cannot be inferred in the way H suggests. The inference process is such that the antecedent of the major premise is detached and is stated as the conclusion. This might be an illegitimate move in deductive inference but it captures the essence of H. In this process, no new content is created; instead some of the content of the premises is detached and is asserted.

This creates a certain tension in Peirce's account. Hypothesis is ampliative and the sole generator of new ideas or content. And yet, in the syllogistic conception

---

[16]The similarity with the standard Deductive-Nomological model of explanation developed by Hempel [1965] is quite striking. On a different occasion, Peirce noted that an explanation is "a syllogism exhibiting the surprising fact as necessarily consequent upon the circumstances of its occurrence together with the truth of the credible conjecture, as premises" (6.469).

of hypothetic inference, the new ideas or content must already be there before they are accepted as the conclusion of the inference. If Hypothesis is the sole generator of new content and ideas, and if this is the reason why it is, in the end, indispensable despite its insecurity, it must have been a great problem for Peirce that the syllogistic conception of reasoning, and of Hypothesis in particular, obscured this fact.

Perhaps this was part of the reason why Peirce abandoned the syllogistic conception of explanatory reasoning. Another part of the reason is that, as noted above, Peirce came to think that the difference between the logical forms of Induction and Hypothesis is not as fundamental as he initially thought. In 1902, he offered the following honest diagnosis of his earlier thinking about explanatory reasoning:

> (M)y capital error was a negative one, in not perceiving that, according to my own principles, the reasoning with which I was there dealing ['Hypothetic Inference'] could not be the reasoning by which we are led to adopt a hypothesis, although I all but stated as much. But I was too much taken up in considering syllogistic forms and the doctrine of logical extension and comprehension, both of which I made more fundamental than they really are. As long as I held that opinion, my conceptions of Abduction necessarily confused two different kinds of reasoning. When, after repeated attempts, I finally succeeded in clearing the matter up, the fact shone out that probability proper had nothing to do with the validity of Abduction, unless in a doubly indirect manner (2.102).

What then are the two different kinds of reasoning that Peirce's earlier syllogistic approach confused? It seems clear that the confusion was between the reasoning process by means of which hypotheses (with new and extra content) are being formulated and adopted on the basis of explanatory considerations and the reasoning process by means of which these hypotheses are rendered *likely.* In Peirce's later writings hypothetical inference is liberalised. It is no longer constrained by the syllogistic conception of inference. It becomes part of a broader methodological process of inquiry. Induction, on the other hand, is given the role of confirmation. In his eighth Lowell lecture in 1903, Peirce took abduction to be "any mode or degree of acceptance of a proposition as a truth, because a fact or facts have been ascertained whose occurrence would necessarily or probably result in case that proposition were true" (5.603).

5   THE ROAD TO ABDUCTION

In his unfinished manuscript of 1896, *Lessons from the History of Science*, Peirce employed the term 'retroduction' (or 'retroductive inference') to capture hypothetical reasoning — or perhaps, the species of it where the hypothesis concerns things

past. Here too, it is explanation that makes this mode of inference distinctive. As he put it:

> Now a retroductive conclusion is only justified by its explaining an observed fact. An explanation is a syllogism of which the major premiss, or rule, is a known law or rule of nature, or other general truth; the minor premise, or case, is the hypothesis or retroductive conclusion, and the conclusion, or result, is the observed (or otherwise established) fact (1.89).

As he explained, he took 'retroduction' to render into English the Aristotelian term *apag'=ogē*, which as Peirce noted (and as we have already seen in section 2), was "misunderstood because of corrupt text, and as misunderstood [it was] usually translated *abduction*" (1.65). But he did opt for the term 'abduction' in the end, though he also toyed with the term 'Presumption'. As he put it in 1903: "Presumption, or, more precisely, abduction (. . .), furnishes the reasoner with the problematic theory which induction verifies" (2.776). In the same context, he noted that "Logical or philosophical presumption is non-deductive probable inference which involves a hypothesis. It might very advantageously replace hypothesis in the sense of something supposed to be true because of certain facts which it would account for" (2.786).

It should be clear that abduction inherits some of the characteristics of Hypothesis, while it forfeits others. The two main points of contact are that a) it is explanatory considerations that guide abduction, *qua* an inference; and b) abduction is "the only kind of reasoning which supplies new ideas, the only kind which is, in this sense, synthetic" (2.777). But, unlike Hypothesis, abduction c) is not constrained by syllogistic formulations; d) *any* kind of hypothesis can be adopted on its basis, provided it plays an explanatory role. As Peirce notes: "Abduction must cover all the operations by which theories and hypotheses are engendered" (5.590). Besides, Peirce took it that his shift to abduction lay further emphasis on the fact that abduction is an insecure mode of reasoning; that the abductively adopted hypothesis is problematic; that it needs further testing. The rationale for abduction, then, is that if rational explanation is possible at all, it can only achieved by abduction. As he put it: "Its only justification is that its method is the only way in which there can be any hope of attaining a rational explanation" (2.777).

Peirce's classic characterisation of abduction *qua* inference, in his seventh lecture on Pragmatism titled *Pragmatism and Abduction* in 1903, is this (cf. 5.189):

> (CC)
>
> The surprising fact, C, is observed;
> But if A were true, C would be a matter of course,
> Hence, there is reason to suspect that A is true.

Immediately before (CC) he noted that abduction is "the operation of adopting an explanatory hypothesis" and that "the hypothesis cannot be admitted, even as

a hypothesis, unless it be supposed that it would account for the facts or some of them". The emphasis on C being a matter of course relates to Peirce's conception of explanation as rational expectability. What follows the classic characterisation is even more interesting. Peirce says: "Thus, A cannot be abductively inferred, or if you prefer the expression, cannot be abductively conjectured until its entire content is already present in the premise, 'If A were true, C would be a matter of course'". This claim captures the way in which abduction (in contradistinction to the earlier Hypothesis) can be genuinely ampliative *and* generative of new ideas and content. What Peirce implies, and what seems right anyway, is that the abductive inference generates both the major premise 'If A were true, C would be a matter of course' *and* licenses the conclusion that there is reason to accept A as true. Though A may be familiar in itself, it does not follow that it is the case that 'If A were true, C would be a matter of course'. *Qua* a hypothesis A has excess and new content vis-à-vis C precisely because it offers a reason (explains) why C holds.

The following then sounds plausible. Abduction is a dual process of reasoning. It involves the generation of some hypothesis A with excess content in virtue of which the explanandum C is accounted for, where the explanatory connection between A and C is captured by the counterfactual conditional 'If A were true, C would be a matter of course'. But it also allows the detachment of the antecedent A from the conditional and hence its acceptance in its own right. The detachment of the antecedent A requires reasons and these are offered by the explanatory connection there is between the antecedent and the consequent.

Peirce has had a rather broad conception of a surprising fact. He took it that the very presence of regularities in nature is quite surprising in that irregularity (the absence of regularity) is much more common than regularity in nature. Hence the presence of regular patterns under which sequences of events fall are, for Peirce, unexpected and requires explanation (cf. 7.189; 7.195). This suggests that Peirce would not take the regularities there are in nature as crude facts — which admit or require no further explanation. Regularities hold for a reason and their explanation amounts to finding the reason for which they hold, thereby rendering the phenomena rational (cf. 7.192). But aren't also deviations from a regularity surprising? Peirce insists that if an existing regularity is breached, this does require explanation (cf. 7.191). So both the regularity *and* the deviations from it require explanation, though the explanations offered are at different levels. A key feature of explanation according to Peirce is that it renders the explananda less "isolated" than they would have been in the absence of an explanation (cf. 7.199). This feature follows from the fact that explanation amounts to rational expectability. For, a fact is isolated if it does not fall under a pattern. And if it does not fall under a pattern, in its presence we "do not know what to expect" (7.201). Actually, abduction is justified, Peirce claimed, because it is the "only possible hope of regulating our future conduct rationally" (2.270), and clearly this rational regulation comes from devising explanations which render the facts less isolated.

A stable element of Peirce's thought on explanatory reasoning is that it is a *reasoning* process — it obeys a rule of a sort. In *The Logic of Drawing History from Ancient Documents* (1901), he insisted that abduction amounts to an adoption of a hypothesis "which is likely in itself, and renders the facts likely". He noted: "I reckon it as a form of inference, however problematical the hypothesis may be held" (7.202). And he queried about the "logical rules" that abduction should conform to. But why should he think that abduction is a reasoning process?

Recall that for him reasoning is a conscious activity by means of which a conclusion is drawn based on reasons. In his *Short Logic* (1893), he emphasised that reasoning (the making of inferences) amounts to the "conscious and controlled adoption of a belief as a consequence of other knowledge" (2.442). Reasoning is a voluntary activity, which among other things, involves considering and eliminating options (cf. 7.187). This is what abduction is and does. The point is brought home if we consider abduction as an eliminative inference. Not all possible explanatory hypotheses are considered. In answering the objection that abduction is not reasoning proper because one is free to examine whatever theories one likes, Peirce noted:

> The answer [to the question of what need of reasoning was there?] is that it is a question of economy. If he examines all the foolish theories he might imagine, he never will (short of a miracle) light upon the true one. Indeed, even with the most rational procedure, he never would do so, were there not an affinity between his ideas and nature's ways. However, if there be any attainable truth, as he hopes, it is plain that the only way in which it is to be attained is by trying the hypotheses which seem reasonable and which lead to such consequences as are observed (2.776).

What exactly is this criterion of reasonableness? Peirce took it that abduction is not a topic-neutral inferential pattern. It operates within a framework of background beliefs, depends on them and capitalises on them. It is these beliefs that determine reasonableness or plausibility. Here is Peirce, in 1901:

> Of course, if we know any positive facts which render a given hypothesis objectively probable, they recommend it for inductive testing. When this is not the case, but the hypothesis seems to us likely, or unlikely, this likelihood is an indication that the hypothesis accords or discords with our preconceived ideas; and since those ideas are presumably based upon some experience, it follows that, other things being equal, there will be, in the long run, some economy in giving the hypothesis a place in the order of precedence in accordance with this indication (7.220).

Background beliefs play a dual role in abduction. Their first role is to eliminate a number of candidates as "foolish". What hypotheses will count as foolish will surely depend on how strong and well supported the background beliefs are. But

Peirce also insisted that some hypotheses must be discarded from further consideration *ab initio*. They are the hypotheses that, by their very nature, are untestable (cf. 6.524). This call for testability is a hallmark of Peirce's pragmatism. In his *Lectures on Pragmatism*, he famously noted that "the question of pragmatism" is "nothing else than the question of the logic of abduction" (5.196). As he went on to explain, the link between the two is testability. The Maxim of Pragmatism is that admissible hypotheses must be such that their truth makes a difference in experience. In the present context, Peirce put the Maxim thus:

> (...) [T]he maxim of pragmatism is that a conception can have no logical effect or import differing from that of a second conception except so far as, taken in connection with other conceptions and intentions, it might conceivably modify our practical conduct differently from that second conception (5.196).

In other words, there can be no logical difference between two hypotheses that results in no difference in experience. Abduction, according to Peirce, honours this maxim because it "puts a limit upon admissible hypotheses" (5.196). And the limit is set by the logical form of abduction CC, which has already been noted. The major premise of an abductive inference is: 'If A were true, C would be a matter of course'. For A to be admissible at all it must be the case that it renders C explicable and *expectable*. We have already seen that Peirce equated explanation with rational expectability and this clearly implies that he took it that explanation yields predictions (or even that explanation and prediction are the two sides of the same coin). Predictions are, ultimately, what differentiates between admissible (*qua* testable) hypotheses and inadmissible (*qua* untestable) ones. Here is Peirce's own way to put the point:

> Admitting, then, that the question of Pragmatism is the question of Abduction, let us consider it under that form. What is good abduction? What should an explanatory hypothesis be to be worthy to rank as a hypothesis? Of course, it must explain the facts. But what other conditions ought it to fulfil to be good? The question of the goodness of anything is whether that thing fulfils its end. What, then, is the end of an explanatory hypothesis? Its end is, through subjection to the test of experiment, to lead to the avoidance of all surprise and to the establishment of a habit of positive expectation that shall not be disappointed. Any hypothesis, therefore, may be admissible, in the absence of any special reasons to the contrary, provided it be capable of experimental verification, and only insofar as it is capable of such verification. This is approximately the doctrine of pragmatism (5.197).

The second role that background beliefs play in abduction concerns the ranking of the admissible candidates in an "order of preference"; or the selection of hypotheses. Accordingly, the search for explanatory hypotheses is not blind but guided by reasons. The search aims to create, as Peirce (5.592) nicely put it,

"good" hypotheses. So the search will be accompanied by an evaluation of hypotheses, and by their placement in an order of preference according to how good an explanation they offer. In *Hume on Miracles* (1901) Peirce put this point as follows:

> The first starting of a hypothesis and the entertaining of it, whether as a simple interrogation or with any degree of confidence, is an inferential step which I propose to call abduction. This will include a preference for any one hypothesis over others which would equally explain the facts, so long as this preference is not based upon any previous knowledge bearing upon the truth of the hypotheses, nor on any testing of any of the hypotheses, after having admitted them on probation (cf. 6.525).

It is clear from this passage that the preferential ranking of competing hypotheses that would explain the facts, were they true, cannot be based on judgements concerning their truth, since if we already knew which hypothesis is the true one, it would be an almost trivial matter to infer this as against its rivals. So the ranking should be based on different criteria.

What are they? The closest Peirce comes to offering a systematic treatment of this subject is in his *Abduction*, which is part of *The Logic of Drawing History from Ancient Documents* (1901). The principles which should "guide us in abduction or the process of choosing a hypothesis" include:

A. Hypotheses should explain *all* relevant facts.[17]

B. Hypotheses should be licensed by the existing background beliefs;

C. Hypotheses should be, as far as possible, simple ("incomplex" (7.220-1));

D. Hypotheses should have unifying power ("breadth" (7.220-1));[18]

E. Hypotheses should be further testable, and preferably entail novel predictions (7.220).[19]

6   FROM THE INSTINCTIVE TO THE REASONED MARKS OF TRUTH

The picture of abduction that Peirce has painted is quite complex. On the face of it, there may be a question of its coherence. Abduction is an inference by

---

[17] "Still, before admitting the hypothesis to probation, we must ask whether it would explain all the principal facts" (7.235).

[18] Peirce characterised unifying power thus: "The purpose of a theory may be said to be to embrace the manifold of observed facts in one statement, and other things being equal that theory best fulfils its function which brings the most facts under a single formula" (7.410).

[19] Peirce stressed that "the strength of any argument of the Second Order depends upon how much the confirmation of the prediction runs counter to what our expectation would have been without the hypothesis" (7.115).

means of which explanatory hypotheses are admitted, but it is not clear what this admission amounts to. Nor is it clear whether there are rules that this mode of inference is subject to. In a rather astonishing passage that preceded Peirce's classic characterisation of abduction noted above, Peirce stressed: "It must be remembered that abduction, although it is very little hampered by logical rules, nevertheless is logical inference, asserting its conclusion only problematically or conjecturally, it is true, but nevertheless having a definite logical form" (5.188). How can it be that abduction has a definite logical form (the one suggested by CC above) and yet not be hampered by logical rules? Besides, Peirce made the seemingly strange point that "(...) abduction commits us to nothing. It merely causes a hypothesis to be set down upon our docket of cases to be tried" (5.602).

To resolve the possible tensions here we need to take into account Peirce's overall approach to ampliative reasoning. Peirce was adamant that the conclusion of an abductive inference can be accepted only on "probation" or "proviosionally". Here is one of the very many ways in which Peirce expressed this key thought of his: "Abduction, in the sense I give the word, is any reasoning of a large class of which the provisional adoption of an explanatory hypothesis is the type" (2.544, note). One important reason why explanatory hypotheses can only be accepted on probation comes from the history of science itself, which is a history of actual abductions. Though it is reasonable to accept a hypothesis as true on the basis that "it seems to render the world reasonable", a closer look at the fate of explanatory hypotheses suggests that they were subsequently controverted because of wrong predictions. "Ultimately" Peirce said,

> the circumstance that a hypothesis, although it may lead us to expect some facts to be as they are, may in the future lead us to erroneous expectations about other facts, — this circumstance, which anybody must have admitted as soon as it was brought home to him, was brought home to scientific men so forcibly, first in astronomy, and then in other sciences, that it became axiomatical that a hypothesis adopted by abduction could only be adopted on probation, and must be tested (7.202).

Peirce did consider the claim that abduction might admit of a strict logical form based on Bayes's theorem. Well, he did not put it quite that way, but this is what he clearly meant when he said that according to a common theory, reasoning should be "guided by balancing probabilities, according to the doctrine of inverse probability" (2.777). The idea here is that one updates one's degree of belief in a proposition by using Bayes's theorem: $\text{Prob}_{\text{new}}(H)=\text{Prob}_{\text{old}}(H/e)$, where $\text{Prob}(H/e)=\text{Prob}(e/H) \times \text{Prob}(H)/\text{Prob}(e)$. "Inverse probabilities" are what later on became known as likelihoods, viz., $\text{Prob}(e/H)$. As Peirce immediately added, this approach to reasoning relies "upon knowing antecedent probabilities", that is prior probabilities. But he was entirely clear that this Bayesian approach could not capture the logical form of abduction because he thought that prior probabilities in the case of hypotheses were not available. Peirce was totally

unwilling to admit subjective prior probabilities — if there were well-defined prior probabilities applied to hypotheses, they would have to be "solid" statistical probabilities "like those upon which the insurance business rests". But when it comes to hypotheses, the hope for solid statistical facts is futile:

> But they are not and cannot, in the nature of things, be statistical facts. What is the antecedent probability that matter should be composed of atoms? Can we take statistics of a multitude of different universes? An objective probability is the ratio of frequency of a specific to a generic event in the ordinary course of experience. Of a fact per se it is absurd to speak of objective probability. All that is attainable are subjective probabilities, or likelihoods, which express nothing but the conformity of a new suggestion to our prepossessions; and these are the source of most of the errors into which man falls, and of all the worst of them (2.777).

It might be objected here that Peirce's last point is unfair to probabilists, since he himself has tied the adoption of explanatory hypotheses to background beliefs, which could well be the sources of most of the errors into which man falls. Fair enough. Peirce would not think this is an objection to *his* views, precisely because the dependence of abduction on background beliefs is the reason he thought that it could not, in the first instance and on its own, yield probable results and that its conclusion should be accepted on probation, subject to further testing. So his point is that those who think that by relying on subjective prior probabilities can have a conception of inference which yields *likely* hypotheses delude themselves.

What all this means is that abductive inference *per se* is not the kind of inference that can or does lead to likely conclusions. It's not as if we feed a topic-neutral and algorithmic rule with suitable premises and it returns likely conclusions. Abduction is not like that at all. Peirce insisted that when it comes to abduction "yielding to judgments of likelihood is a fertile source of waste of time and energy" (6.534). So abduction is not in the business of conferring probabilities on its conclusions. But this is not to imply abduction is neither an inference nor a means, *qua* inference, to yield reasonably held conclusions.

In his fifth lecture on Pragmatism (1903), Peirce drew a distinction between validity and strength (which is different from the one between uberty and security noted in section 3). An argument is valid, Peirce suggested "if it possesses the sort of strength that it professes and tends toward the establishment of the conclusion in the way in which it pretends to do this" (5.192). This might sound opaque, but the underlying idea is that different sort of inferences aim at different things and hence cannot be lamped together. Deduction aims at truth-preservation or truth-maintenance: if the premises are true, the conclusion has to be true. In deduction, validity and strength coincide because the conclusion of a deductive argument is at least as secure as its premises. But this is peculiar to *deductive* inference. Other inferential patterns may be such that validity and strength do not coincide. An inference may be weak and yet valid. It may be weak in that the conclusion of the inference might not be strongly supported by the premises, and yet it may be valid in Peirce's sense above: the inference does not pretend to license stronger conclusions than it actually does. As he put it: "An argument is none the less logical for being weak, provided it does not pretend to a strength that it does not possess" (5.192). Abduction is a weak inference, but it can be reasonable

nonetheless (or "valid", as Peirce would put it). Unlike deduction, abduction does not advertise itself as truth-preserving. Its aim is the generation of extra content and the provision of reasons for its adoption (based on explanatory considerations). Here is Peirce's own way to out the point: "The conclusion of an abduction is problematic or conjectural, but is not necessarily at the weakest grade of surmise, and what we call assertoric judgments are, accurately, problematic judgments of a high grade of hopefulness" (5.192). Peirce had actually examined this issue in his earlier *Notes on Ampliative Reasoning* (1901-2). There, after noting that "an argument may be perfectly valid and yet excessively weak" (2.780), he went on to suggest that the strength of abduction is a function of its eliminative power.[20] The strength of an abductively inferred hypothesis depends on "the absence of any other hypothesis". But this would suggest that abduction is *very* weak, since how can it possibly be asserted that all other potentially explanatory hypotheses have been eliminated?

To avoid rendering abduction excessively weak, Peirce suggested that strength might be measured in terms of "the amount of wealth, in time, thought, money, etc., that we ought to have at our disposal before it would be worth while to take up that hypothesis for examination". This introduces a new factor into reasoning — over and above the requirements of explanation and testability noted above. We can call this factor 'economy' echoing Peirce's own characterisation of it.

Peirce tied economy to a number of features of abductive reasoning. In his eighth Lowell lecture in 1903, he stressed that "the leading consideration in Abduction" is "the question of Economy–Economy of money, time, thought, and energy" (5.600). Economy is related to the range of potential explanations that may be entertained and be subjected to further testing (cf. 6.528). It is related to the eliminative nature of abduction. Economy dictates that when there is need for choice between competing hypotheses which explain a set of phenomena, some crucial experiment should be devised which eliminates many or most of the competitors.[21] Economy is also related to the use of Ockham's razor (cf. 6.535). The demand for, and the preference of, simple explanation is "a sound economic principle" because simpler explanations are more easily tested.

The general point here is that abduction — *qua* reasoning — is subjected to criteria that do not admit a precise logical formulation. But these criteria are necessary for the characterisation of abduction nonetheless if abduction is to be humanly possible. In contradistinction to Descartes, Peirce was surely not inter-

---

[20]In the case of Induction, Peirce noted that the larger the number of instances that form the inductive basis, the stronger the induction. But, he added, weak inductions (based on small numbers of instances) are perfectly valid (cf. 2.780).

[21]Here is how Peirce put it: "Let us suppose that there are thirty-two different possible ways of explaining a set of phenomena. Then, thirty-one hypotheses must be rejected. The most economical procedure, when it is practicable, will be to find some observable fact which, under conditions easily brought about, would result from sixteen of the hypotheses and not from any of the other sixteen. Such an experiment, if it can be devised, at once halves the number of hypotheses" (6.529).

ested in the project of pure enquiry. Inference, in particular ampliative inference, does not operate in a vacuum; nor is it subjected to no constraints other than the search for the truth. Nor does it occur in an environment of unlimited resources of time and energy. Principles of economy govern abductive inference precisely because abduction has to work its way through a space of hypotheses that is virtually inexhaustible. So either no abductive inference would be possible or there should be criteria that cut down the space of hypotheses to a reasonable size (cf. 2.776).

It might be thought that these considerations of economy render abduction totally whimsical. For, one may wonder, what possibly could be the relation between abduction and truth? Note, however, that this worry would be overstated. Principles of economy are principles which facilitate the further testing of the selected hypotheses. Hence, they can facilitate finding out whether a hypothesis is true in the only sense in which Peirce can accept this, viz., in the sense of making a hypothesis doubt-resistant. But there is a residual worry that is more serious. If abduction does not operate within a network of background of true beliefs, there is no way in which it can return hypotheses which have a good chance of being true. How can these true background beliefs emerge? In at least two different places, Peirce argues that the human mind has had the power to imagine correct theories, where this power is a "natural adaptation" (5.591). On one of these two occasions, he clearly associated this power of the human mind ("the guessing instinct") with the principles of economy in abductive reasoning. These principles work because the mind has the power to hit upon the truth in a relatively small number of trials. Here is how he put it:

> In very many questions, the situation before us is this: We shall do better to abandon the whole attempt to learn the truth, however urgent may be our need of ascertaining it, unless we can trust to the human mind's having such a power of guessing right that before very many hypotheses shall have been tried, intelligent guessing may be expected to lead us to the one which will support all tests, leaving the vast majority of possible hypotheses unexamined (6.530).

Peirce does not prove this claim, how could he? He does say in its support that truth has survival value (cf. 5.591). But it is not clear that this is anything other than speculation. A more likely ground for Peirce's claim is quasi-transcendental, viz., that unless we accept that the human mind has had this power to guess right, there can be no rational explanation of why it has come up with some true theories in the first place. Peirce tries to substantiate this claim by means of a further argument. True theories cannot be a matter of chance because given all possible theories that could have been entertained, stumbling over a true one is extremely unlikely. The possible theories, Peirce said, "if not strictly innumerable, at any rate exceed a trillion – or the third power of a million; and therefore the chances are too overwhelmingly against the single true theory in the twenty or thirty thousand years during which man has been a thinking animal, ever having

come into any man's head" (5.591). Note that this kind of argument is based on a statistical claim, which might be contentious: how can we come up with such statistics in the first place? Be that as it may, Peirce's key point here is that though abduction does not wear its justification on its sleeve, it is reasonable to think that abduction does tend to operate within networks of true background beliefs.

It is fair to say that though abduction cannot have a foundational role, its products cannot be doubted *en masse*, either. Its justification, *qua* mode of inference, comes from the need for rational explanation and in particular from the commitment to the view that rational explanation is possible; that the facts "admit of rationalization, and of rationalization by us" (7.219). Interestingly, Peirce claims that this commitment embodies another hypothesis, and as such it is the product of "a fundamental and primary abduction". As Peirce put it: "it is a primary hypothesis underlying all abduction that the human mind is akin to the truth in the sense that in a finite number of guesses it will light upon the correct hypothesis". This creates an air of circularity, of course. In essence, a grand abduction is employed to justify the possibility of abductive inference. Peirce does not address this problem directly. For him it seems that this circularity is the inevitable price that needs to be paid if human understanding is at all possible. Explanation aims at (and offers) understanding, but unless it is assumed that the human mind has a capacity or power to reach the truth in a finite number of trials, hitting the right explanations would be a random walk. It is no surprise, then, that Peirce brings in instinct once more. He draws a distinction between two kinds of considerations "which tend toward an expectation that a given hypothesis may be true": the purely instinctive and the reasoned ones (7.220). The instinctive considerations kick in when it comes to the primary hypothesis that the human mind has a power to hit upon the truth. This is not reasoned, though it is supported by an induction on the past record of abductive inferences. As Peirce put it, "it has seldom been necessary to try more than two or three hypotheses made by clear genius before the right one was found". The reasoned considerations kick in when a body of background beliefs has emerged which has some measure of truth in it. Then, the choice among competing hypotheses is guided by criteria noted above, e.g., breadth and incomplexity. For Peirce, however, it would be folly to try to hide the claim that "the existence of a natural instinct for truth is, after all, the sheet-anchor of science. From the instinctive, we pass to reasoned, marks of truth in the hypothesis" (7.220).[22]

In one of the first systematic treatments of Peirce's views of abduction, Harry Frankfurt (1958, 594) raised what might be called Peirce's paradox. This is that Peirce appears to want to have it both ways, viz., "that hypotheses are the products of a wonderful imaginative faculty in man and that they are products of a certain sort of logical inference". It should be clear by now that this paradox is only

---

[22]The role of instinct in abduction is raised and discussed, in more or less the same way, in Peirce's fifth lecture on Pragmatism, in 1903 (5.171-5.74). There, Peirce expresses his view that the instinct of guessing right is accounted for by evolution.

apparent. Abduction involves a guessing instinct *and* is a reasoned process, but for Peirce these two elements operate at different levels. The guessing instinct is required for the very possibility of a trustworthy abductive inference. The reasoned process operates within an environment of background beliefs and aims to select among competing hypotheses on the basis of explanatory considerations.

7   THE THREE STAGES OF INQUIRY

In Peirce's mature thought, we have already seen, abduction covers a cluster of operations that generate and evaluate explanatory hypotheses. Peirce was adamant, it was noted, that abduction is *not* the kind of inference that returns likely hypotheses. It's not in the business of producing judgements of likelihood. This is not to say, we have stressed, that abduction is not trustworthy. Rather, its trustworthiness is a function of the background beliefs within which it operates. But is it not the case that, in the end of the day, we want theories or hypotheses that are likely to be true? Peirce never doubted this. In 1901 he summed this up by saying: "A hypothesis then has to be adopted which is likely in itself and renders the facts likely. This process of adopting a hypothesis as being suggested by the facts is what I call abduction" (7.202). But how can abduction lead to likely hypotheses if it is not meant to do so?

In his mature writings Peirce treated abduction as the first part of a three-stage methodological process, the other two stages being deduction and induction. The burden of likelihood is carried not by abduction in and of itself but by the other two methods which complement abduction in the overall method of inquiry. Abduction might confer plausibility or reasonableness on its conclusion, but their probability is determined by their further testing. Here is a long but nice summary by Peirce himself, offered in 1908:

> The whole series of mental performances between the notice of the wonderful phenomenon and the acceptance of the hypothesis, during which the usually docile understanding seems to hold the bit between its teeth and to have us at its mercy, the search for pertinent circumstances and the laying hold of them, sometimes without our cognizance, the scrutiny of them, the dark laboring, the bursting out of the startling conjecture, the remarking of its smooth fitting to the anomaly, as it is turned back and forth like a key in a lock, and the final estimation of its Plausibility, I reckon as composing the **First Stage of Inquiry**. Its characteristic formula of reasoning I term Retroduction, i.e. reasoning from consequent to antecedent (6.469).
>
> Retroduction does not afford security. The hypothesis must be tested. This testing, to be logically valid, must honestly start, not as Retroduction starts, with scrutiny of the phenomena, but with examination of the hypothesis, and a muster of all sorts of conditional experiential consequences which would follow from its truth. This constitutes the

> **Second Stage of Inquiry**. For its characteristic form of reasoning
> our language has, for two centuries, been happily provided with the
> name Deduction (6.470).

> The purpose of Deduction, that of collecting consequents of the hypoth-
> esis, having been sufficiently carried out, the inquiry enters upon its
> **Third Stage**, that of ascertaining how far those consequents accord
> with Experience, and of judging accordingly whether the hypothesis
> is sensibly correct, or requires some inessential modification, or must
> be entirely rejected. Its characteristic way of reasoning is Induction
> (6.472).

Abduction is the sole method by means of which new ideas are introduced. It
is the only method by means of which the phenomena are 'rationalised' by being
explained. But to get from an abductively inferred hypothesis to a judgement of
probability, this hypothesis should be subjected to further testing. According to
Peirce:

> The validity of a presumptive adoption of a hypothesis for examination
> consists in this, that the hypothesis being such that its consequences
> are capable of being tested by experimentation, and being such that
> the observed facts would follow from it as necessary conclusions, that
> hypothesis is selected according to a method which must ultimately
> lead to the discovery of the truth, so far as the truth is capable of being
> discovered, with an indefinite approximation to accuracy (2.781).

Taken on its own, abduction is the method of generation and ranking of hypotheses
which potentially explain a certain explanandum. Peirce says: "The first starting
and the entertaining of [a hypothesis], whether as a simple interrogation or with
any degree of confidence, is an inferential step which I propose to call abduction"
(6.525). But these hypotheses should be subjected to further testing which will
determine, ultimately, their degree of confirmation.

Accordingly, Peirce suggests that abduction should be embedded in a broader
framework of inquiry so that the hypotheses generated and evaluated by abduction
can be further tested. The result of this testing is the confirmation or disconfir-
mation of the hypotheses. So, Peirce sees abduction as the first stage of the
reasoners' attempt to add reasonable beliefs into their belief-corpus in the light of
new phenomena or observations. The process of generation and first evaluation of
hypotheses (abduction) is followed by deduction — i.e., by deriving further pre-
dictions from the abduced hypotheses — and then by induction which now Peirce
understands as the process of testing these predictions and hence the process of
confirming the abduced hypothesis (cf. 7.202ff).

"As soon as a hypothesis has been adopted", Peirce (7.203) says, the next step
"will be to trace out its necessary and probable experiential consequences. This
step is deduction". And he adds:

> Having, then, by means of deduction, drawn from a hypothesis predic-
> tions as to what the results of experiment will be, we proceed to test
> the hypothesis by making the experiments and comparing those pre-
> dictions with the actual results of the experiment. (. . .) When, (. . .),
> we find that prediction after prediction, notwithstanding a preference
> for putting the most unlikely ones to the test, is verified by experiment,
> whether without modification or with a merely quantitative modifica-
> tion, we begin to accord to the hypothesis a standing among scientific
> results (7.206).

Induction, then, is given an overall different role than the one it had in his ear-
lier thinking. It now captures the methods by means of which hypotheses are
confirmed. Hence, in the transition from his earlier views to his later ones, what
really changed is not Peirce's conception of explanatory reasoning, but rather his
views on induction. Induction changed status: from a distinct mode of amplia-
tive reasoning with a definite syllogistic form which leads to the acceptance of a
generalisation as opposed to a fact (early phase) to the general process of testing
a hypothesis. As Peirce out it: "This sort of inference it is, from experiments
testing predictions based on a hypothesis, that is alone properly entitled to be
called induction" (7.206). Induction is a process "for testing hypotheses already
in hand. The induction adds nothing" (7.217).

Induction is no less indispensable than abduction in the overall process of in-
quiry — but its role is clearly different from the role of abduction. Peirce put this
point in a picturesque way when he said that our knowledge of nature consists in
building a "cantilever bridge of inductions" over the "chasm that yawns between
the ultimate goal of science and such ideas of Man's environment", but that "every
plank of [this bridge] is first laid by Retroduction alone" (6.475).

Peirce kept his view that abduction and induction are distinct modes of reasoning. In *The Logic of Drawing History from Ancient Documents* (1901), he noted
that abduction and induction are "the opposite poles of reason, the one the most
ineffective, the other the most effective of arguments" (7.218). Abduction is "the
first step of scientific reasoning, as induction is the concluding step". Abduction
is "merely preparatory".

> Abduction makes its start from the facts, without, at the outset, having
> any particular theory in view, though it is motived by the feeling that
> a theory is needed to explain the surprising facts. Induction makes its
> start from a hypothesis which seems to recommend itself, without at
> the outset having any particular facts in view, though it feels the need
> of facts to support the theory. Abduction seeks a theory. Induction
> seeks for facts. In abduction the consideration of the facts suggests
> the hypothesis. In induction the study of the hypothesis suggests the
> experiments which bring to light the very facts to which the hypothesis
> had pointed.

Nonetheless, abduction and induction have a common feature: "that both lead to

the acceptance of a hypothesis because observed facts are such as would necessarily or probably result as consequences of that hypothesis".

Hence, Peirce has moved a long way from his earlier view on induction. Abduction covers all kinds of explanatory reasoning (including explanation by subsumption under a generalisation), while induction is confirmation. What is important to note is that Peirce took it that induction is justified in a way radically distinct from the way abduction is justified. He thought that induction is, essentially, a self-corrective method[23], viz., that "although the conclusion [of induction] at any stage of the investigation may be more or less erroneous, yet the further application of the same method must correct the error" (5.145). Being a frequentist about probabilities, Peirce clearly thought that a consistent application of the straight rule of induction will converge in the limit to the true relative frequency of a certain factor A in a class of events B.

In one of the most interesting studies of Peirce's abduction, Douglas R. Anderson (1986, 162) noted that Peircean abduction "is a possibilistic inference whose test is *in futuro*". This claim goes a long way in capturing the essence of Peircean abduction. Peirce employed the Aristotelian idea of "esse in futuro" to capture a mode of being which is potential, and not actual. For him, potentialities as well as laws of nature have their esse *in futuro*. Abduction, it might be claimed, has its justification *in futuro* — or, better put, it has its *full* justification *in futuro*. This means that although a hypothesis might be reasonably accepted as plausible based on explanatory considerations (abduction), the *degree of confidence* in this hypothesis is not thereby settled. Rather it is tied to the degree of confirmation of this hypothesis, where the latter depends, ultimately, on the future performance of the hypothesis, viz., on how well-confirmed it becomes by further evidence. This conception of justification *in futuro* tallies well with Peirce'a account of knowledge and truth. The aim of inquiry is to get doubt-resistant beliefs. As noted already, truth itself boils down to doubt-resistant belief. In *What Pragmatism Is*, in 1905, Peirce said:

> You only puzzle yourself by talking of this metaphysical 'truth' and metaphysical 'falsity,' that you know nothing about. All you have any dealings with are your doubts and beliefs, with the course of life that forces new beliefs upon you and gives you power to doubt old beliefs. If your terms 'truth' and 'falsity' are taken in such senses as to be definable in terms of doubt and belief and the course of experience (as for example they would be, if you were to define the 'truth' as that to a belief in which belief would tend if it were to tend indefinitely toward absolute fixity), well and good: in that case, you are only talking about doubt and belief. But if by truth and falsity you mean something not definable in terms of doubt and belief in any way, then you are talking of entities of whose existence you can know nothing, and which Ockham's razor would clean shave off. Your problems would

---

[23] "That Induction tends to correct itself, is obvious enough" (5.776).

be greatly simplified, if, instead of saying that you want to know the 'Truth,' you were simply to say that you want to attain a state of belief unassailable by doubt (5.416).

All beliefs, then, which are not certain should be subjected to further testing — it is only this further testing (or, at least, the openness to further testing) that can render beliefs permanently settled and hence doubt-resistant. The justification of all fallible beliefs is *in futuro*. Abduction generates and recommends beliefs; but the process of their becoming doubt-resistant is external to abduction — there is where induction rules.

8   LOOKING AHEAD

Peirce had the intellectual courage to explore uncharted territories, but this exploration did not leave behind a full and comprehensive map. Despite his expressed wish to write short book on "the real nature" of explanatory reasoning, he left behind papers, notes and unfinished manuscripts and, with them, a big challenge to his followers to reconstruct his thinking and put together a coherent and comprehensive theory of ampliative reasoning.

A few decades passed after Peirce's death in 1913 before philosophers started to appreciate the depth, richness and complexity of Peirce's views of abduction. It was not until the publication of the first two volumes of his collected papers in 1931-2, that philosophers started to pay a more systematic attention to Peirce's philosophy, and to his writings on abduction, in particular. In a paper published a few years after Peirce's death, Professor Josiah Royce (who bequeathed Peirce's manuscripts to the Harvard philosophy department and the Harvard library) brought to attention Peirce's *Lectures on Pragmatism* as well as his *Lowell Lectures on Logic* in 1903-4 and made the following characteristic comment [Royce, 1916, p. 708]:

> It was these latter [the Lowell Lectures] which James described as 'flashes of brilliant light relieved against Cimmerian darkness — 'darkness' indeed to James as to many others must have seemed those portions on 'Existential Graphs' or 'Abduction'.

William James's reported view of Peirce's writings on abduction was far from atypical. There is virtually no attempt for a reconstruction or exegesis of Peirce's views of abduction before Arthur Burks's (1946). In his long an instructive review of the first two volumes of Peirce's Collected Papers, Ernest Nagel (1933, 382) devoted only a few lines on abduction noting that "*Presumptive reasoning*, (...) (also called abduction, retroduction, hypothesis), consists in inferring an explanation, cause, or hypothesis from some fact which can be taken as a consequence of the hypothesis". And Hans Reichenbach made the following passing note in his [1938, p. 36]:

> I admire Charles Peirce as one of the few men who saw the relations
> between induction and probability at an early time; but just his re-
> marks concerning what he calls 'abduction' suffer from an unfortunate
> obscurity which I must ascribe to his confounding the psychology of
> scientific discovery with the logical situation of theories in relation to
> observed facts.

When Peirce's views were studied more carefully, there were two broad ways in
which they were developed. The first focused on the issue of justification and
reliability of ampliative reasoning; the second focused on the process of discovery
of explanatory theories. Gilbert Harman's [1965] paper on Inference to the Best
Explanation (IBE) argued that the best way to conceive of abduction *qua* an
inferential method was to see it as the method of inferring to the truth of the
best among a number of competing rival explanations of a set of phenomena. On
Harman's view, abduction is the mode of inference in which a hypothesis H is
accepted on the basis that a) it explains the evidence and b) no other hypothesis
explains the evidence as well as H does.

In a sense, IBE ends up being a liberalised version of Peircean abduction; it is
defended as the mode of ampliative reasoning that can encompass hypothetico-
deductivism and enumerative induction as special cases.[24] One important issue
in this way of thinking about abduction concerns its justification: why should
it be taken to be the case that IBE is truth-conducive? Here the issue of the
justification of IBE has been tied to the prospects of the defence of scientific realism
in the philosophy of science.[25] Another important issue concerns the virtues of
hypotheses that make up goodness of explanation, or measure explanatory power.
The identification of these virtues has not gone much further than what Peirce
suggested (see, for instance, [Thagard, 1978]). But the justification of the truth-
conductive character of these virtues has become a subject of intense debate (see
[McMullin, 1992]). A third issue concerns the relationship between abduction, *qua*
IBE, and the Bayesian theory of confirmation and belief updating (see [Lipton,
2004]).

It was Norman Russell Hanson in the 1950s who suggested that Peirce's ab-
duction should be best seen as a logic of discovery. The then dominant tradition
was shaped by Reichenbach's distinction between the context of discovery and the
context of justification and the key thought (shared by Karl Popper and others as
well) was that discovery was not subject to rules — it obeyed no logic; it was sub-
ject only to a psychological study (see Reichenbach's comment on Peirce above).
Hanson suggested that discovery falls under rational patterns and argued that
this was Peirce's key idea behind abduction. He took it that a logic of discovery
is shaped by the following type of structure: it proceeds retroductively, from an
anomaly to the delineation of a kind of explanation H which fits into an organised
pattern of concepts [1965, p. 50].

---

[24]For more on this, see Psillos [2002].
[25]For more on this, see Psillos [1999, chapter 4].

In the 1980s, the study of abduction found a new home in Artificial Intelligence. The study of reasoning, among other things, by computer scientists unveiled a variety of modes of reasoning which tend to capture the defeasible, non-monotonic and uncertain character of human reasoning. The study of abduction became of prominent aspect of this new focus on reasoning. In this respect, pioneering among the researchers in AI has been Bob Kowalski. Together with his collaborators, Kowalski attempted to offer a systematic treatment of both the syntax and the semantic of abduction within the framework of Logic Programming. The aim of an abductive problem is to assimilate a new datum O into a knowledge-base (KB). So, KB is suitably extended by a certain hypothesis H into KB' such that KB' incorporates the datum O. Abduction is the process through which a hypothesis H is chosen (see [Kakas *et al.*, 1992; 1997]). Others, notably Bylander and his collaborators (1991), have aimed to offer computational models of abduction which capture its evaluative element.[26] Abduction has been used in a host of areas such as fault diagnosis (where abduction is used for the derivation of a set of faults that are likely to cause a certain problem); belief revision (where abduction is used in the incorporation of new information in a belief corpus); as well as scientific discovery, legal reasoning; natural language understanding, and model-based reasoning. In these areas, there have been attempts to advance formal models of abductive reasoning so that its computational properties are clearly understood and its relations to other kinds of reasoning becomes more precise.

A rich map of the conceptual and computational models of abduction is offered in Gabbay and Woods [2005]. In this work, Gabbay and Woods advance their own formal model of abduction that aims to capture some of the nuances of Peirce's later account. They treat abduction as a method to solve an ignorance-problem, where the latter is a problem not solvable by presently available cognitive resources. Given a choice between surrender (leaving the problem unsolved) and subduance (looking for novel cognitive resources), Gabbay and Woods promote abduction as a middle way: ignorance is not (fully) removed, but becomes the basis for looking for resources upon which reasoned action can be based. The abduced hypothesis does not become known, but it is still the basis for further exploration and action.

Circa 1897, Peirce wrote this:

> The development of my ideas has been the industry of thirty years. I did not know as I ever should get to publish them, their ripening seemed so slow. But the harvest time has come, at last, and to me that harvest seems a wild one, but of course it is not I who have to pass judgment. It is not quite you, either, individual reader; it is experience and history (1.12).

Both experience and history have now spoken. Peirce's theory of abduction still yields fruits and promises good harvests for many years to come.

---

[26]See also [Josephson and Josephson, 1994]. For a good survey of the role of abduction in AI, see [Konolige, 1996].

FURTHER READING

Perhaps the most important early writings on Peirce's theory of abduction are by Burks [1946], Frankfurt [1958], and Fann [1970]. A very significant more recent article is Anderson [1986]. Even more recent work that discusses aspects of Peirce's views of abduction are Hofmann [1999] and Paavola [2007]. An excellent, brief but comprehensive account of Peirce's philosophy of pragmatism is given in Misak [1999]. Thagard's [1981] is a brief but suggestive account of the relation between abduction and hypothesis, while his [1977] explains the relation between Induction and Hypothesis. On Peirce's account of Induction, see Goudge [1940], Jessup [1970] and Sharpe [1970]. On issues related to the abduction as a logic of discovery, see Hanson [1965]. The classic book-length treatment of Inference to the Best Explanation is by Lipton [1991]. A recent thorough discussion of the rival interpretations of Peirce (IBE vs logic of discovery) is given in McKaughan, D. J. [2008]. For an emphasis on computational aspects of abduction, see Aliseda [2006]. The role of abduction in science is discussed in Magnani [2001]. On the relation between abduction and Bayesian confirmation, see the symposium on Peter Lipton's *Inference to the Best Explanation* in *Philosophy and Phenomenological Research*, 74: 421-462, (2007) (Symposiasts: Alexander Bird, Christopher Hitchcock and Stathis Psillos). For a development of the Peircean two-dimensional framework see Psillos [2002] and [2009].

BIBLIOGRAPHY

[Aliseda, 2006] A. Aliseda. *Abductive Reasoning. Logical Investigations into Discovery and Explanation*, Synthese Library vol.330, Springer, 2006.

[Anderson, 1986] D. R. Anderson. The Evolution of Peirce's Concept of Abduction, *Transactions of the Charles S. Peirce Society* **22**: 145-64, 1986.

[Burks, 1946] A. Burks. Peirce's Theory of Abduction, *Philosophy of Science* **13**: 301-306, 1946.

[Bylander *et al.*, 1991] T. Bylander, D. Allemang, M. C. Tanner, and J. R. Josephson. The Computational Complexity of Abduction, *Artificial Intelligence* **49**: 25-60, 1991.

[Fann, 1970] K. T. Fann. *Peirce's Theory of Abduction*, The Hague: Martinus Nijhoff, 1970.

[Frankfurt, 1958] H. Frankfurt. Peirce's Notion of Abduction, *The Journal of Philosophy* **55**: 593-7, 1958.

[Gabbay and Woods, 2005] D. M. Gabbay and J. Woods. *The Reach of Abduction: Insight and Trial,* volume 2 of *A Practical Logic of Cognitive Systems,* Amsterdam: North-Holland, 2005.

[Goodge, 1940] T. A. Goodge. Peirce's Treatment of Induction, *Philosophy of Science* **7**: 56-68, 1940.

[Hanson, 1965] N. R. Hanson. Notes Towards a Logic of Discovery, in R.J. Bernstein (ed.) *Critical Essays on C.S. Peirce*, Yale University Press, 1965.

[Harman, 1965] G. Harman. Inference to the Best Explanation, *The Philosophical Review* **74**: 88-95, 1965.

[Hempel, 1965] C. G. Hempel. *Aspects of Scientific Explanation*, New York: The Free Press, 1965.

[Hofmann, 1999] M. Hofmann. Problems with Peirce's Concept of Abduction, *Foundations of Science* **4**: 271-305, 1999.

[Josephson and Josephson, 1994] R. Josephson and S. G. Josephson, eds. *Adducive Inference*, Cambridge University Press, Cambridge, 1994.

[Jessup, 1974] J. A. Jessup. Peirce's Early Account of Induction, *Transactions of the Charles S. Peirce Society* **10**: 224-34, 1974.

[Kakas *et al.*, 1992] A. C. Kakas, R. A. Kowalski, and F. Toni. Abductive Logic Programming, *Journal of Logic and Computation* **2**: 719-70, 1992.

[Kakas *et al.*, 1997] A. C. Kakas, R. A. Kowalski, and F. Toni. The Role of Abduction in Logic Programming, in D. Gabbay et al. (eds) *Handbook in Artificial Intelligence and Logic Programming*, Oxford: Oxford University Press, 1997.

[Konolige, 1996] K. Konolige. Abductive Theories in Artificial Intelligence, in G. Brewka (ed.) *Principles of Knowledge Representation*, CSLI Publications, 1996.

[Lipton, 1991] P. Lipton. *Inference to the best Explanation* (2nd enlarged edition, 2004), London: Routledge, 1991.

[Magnani, 2001] L. Magnani. Abduction, Reason and Science. Processes of Discovery and Explanation, New York: Kluwer Academic, 2001.

[McKaughan, 2008] D. J. McKaughan. From Ugly Duckling to Swan: C. S. Peirce Abduction and the Pursuit of Scientific Theories, *Transactions of the Charles S. Peirce Society* **44**: 446-68, 2001.

[McMullin, 1992] E. McMullin. *The Inference that Makes Science*, Milwaukee: Marquette University Press, 1992.

[Misak, 1999] C. Misak. American Pragmatism — Peirce, in C. L. Ten (ed.) *Routledge History of Philosophy, Volume 7, The Nineteenth Century*, London: Routledge, 1999.

[Nagel, 1933] E. Nagel. Charles Peirce's Guesses at the Riddle, *The Journal of Philosophy* **30**: 365-86, 1933.

[Paavola, 2007] S. Paavola. On the Origin of Ideas: An Abductivist Approach to Discovery, *Philosophical Studies from the University of Helsinki* 15, 2007.

[Peirce, 1931–1958] C. S. Peirce. *Collected Papers of Charles Sanders Peirce*, C. Hartshorne & P. Weiss (eds) (volumes 1-6) and A. Burks (volumes 7 and 8), Cambridge MA: Belknap Press, 1931-1958.

[Psillos, 1999] S. Psillos. *Scientific Realism: How Science Tracks Truth*, London: Routledge, 1999.

[Psillos, 2002] S. Psillos. Simply the Best: A Case for Abduction, in A. C. Kakas and F. Sadri (eds) *Computational Logic: From Logic Programming into the Future*, LNAI 2408, Berlin-Heidelberg: Springer-Verlag, pp.605-25, 2002.

[Psillos, 2009] S. Psillos. *Knowing the Structure of Nature*, Palgrave-MacMillan, 2009.

[Reichenbach, 1938] H. Reichenbach. On Probability and Induction, *Philosophy of Science* **5**: 21-45, 1938.

[Ross, 1949] W. D. Ross. *Aristotle's Prior and Posterior Analytics*, (with intr. and commentary), Oxford: Clarendon Press, 1949.

[Royce, 1916] J. Royce. Charles Sanders Peirce, *The Journal of Philosophy* **13**: 701-9, 1916.

[Sharpe, 1970] R. Sharpe. Induction, Abduction and the Evolution of Science, *Transactions of the Charles S. Peirce Society* **6**: 17-33, 1970.

[Smith, 1989] R. Smith. *Aristotle — Prior Analytics* (translation, with intr., notes and commentary), Indianapolis: Hackett Publishing Company, 1989.

[Thagard, 1977] P. Thagard. On the Unity of Peirce's Theory of Hypothesis, *Transactions of the Charles S. Peirce Society* **113**: 112-21, 1977.

[Thagard, 1978] P. Thagard. The Best Explanation: Criteria for Theory Choice, *The Journal of Philosophy* **75**: 76-92, 1978.

[Thagard, 1981] P. Thagard. Peirce on Hypothesis and Abduction, in K. Ketner *et al.*, eds., *Proceedings of the C.S. Peirce Bicentennial International Congress*, Texas: Texas Tech University Press, 271-4, 1981.