Simply the Best: A Case for Abduction*

Stathis Psillos

Department of Philosophy and History of Science, University of Athens,
37 John Kennedy Str. 16121 Athens, Greece.
psillos@phs.uoa.gr,
http://www.uoa.gr/dhps/postgrad1/profs/psillos.htm

Abstract. This paper formulates what I think is *the basic problem* of any attempt to characterise the abstract structure of scientific method, viz., that it has to satisfy two conflicting desiderata: it should be ampliative (content-increasing) and it should confer epistemic warrant on its outcomes. Then, after two extreme solutions to the problem of the method, viz., Enumerative Induction and the Method of Hypothesis, are examined, the paper argues that abduction, suitably understood as Inference to the Best Explanation, offers the best description of scientific method and solves the foregoing problem in the best way: it strikes the best balance between ampliation and epistemic warrant.

1 Introduction

In the last decade there has been a lot of work on abduction, both among philosophers and researchers in Artificial Intelligence (AI). Philosophers have mostly tried to unravel the conceptual problems that this mode of reasoning faces[1], whereas workers in AI have looked into its computational modelling.[2] Pioneering among the researchers in AI has been Bob Kowalski. Together with his collaborators, Kowalski has attempted to offer a systematic treatment of both the syntax and the semantic of abduction, with an eye to how Logic Programming can offer the appropriate framework to deal with these issues. It is this primarily theoretical work that will be, I think, the lasting influence of Kowalski's work on our thinking about abduction. In

* This essay is dedicated to Bob Kowalski for his very generous help and the long time we spent in London discussing about philosophy of science and Artificial Intelligence. His inquisitive mind and sharp criticism made me think harder about the philosophical problems of abduction. Many thanks are due to two anonymous readers for this volume and John Norton for useful comments. An earlier version of this paper was presented at the NORDPLUS Intensive Programme on Inference to the Best Explanation in Iceland. Comments made by Jan Faye, Olav Gjelsvik, Mikael Karlsson and Bengt Hansson were particularly useful.

[1] Some recent philosophical work includes [2], [6], and [36]. A fresh approach to abduction has been presented in [11] where Fodor uses the very fact that reasoners employ abduction to raise some important worries against computational theories of mind.

[2] For appraisals of the recent work on abduction in AI, see [1], [25], [28], [37], [49] and the papers in [10].

A.C. Kakas, F. Sadri (Eds.): Computat. Logic (Kowalski Festschrift), LNAI 2408, pp. 605-625, 2002.
© Springer-Verlag Berlin Heidelberg 2002

particular, Kowalski saw very clearly that a number of tangles in the foundations of AI could be dealt with successfully by taking abduction seriously and by incorporating it within AI.[3] In this article, however, my aim is not to deal with the philosophical implications and the possible problems of the analysis of abduction within Logic Programming. I have tried to do this in [42], which can usefully be seen as a companion to the present article. Instead, in this paper I will do two things. First, I shall formulate what I think is *the basic problem* of any attempt to characterise the abstract structure of scientific method, viz., that it has to satisfy two conflicting desiderata: it should be ampliative (content-increasing) and it should confer epistemic warrant on its outcomes (cf. [13], [41]). Second, and after I have examined two extreme solutions to the problem of the method, viz., Enumerative Induction and the Method of Hypothesis, I will try to show that abduction, suitably understood as Inference to the Best Explanation (henceforth, IBE), offers the best description of scientific method and solves the foregoing problem in the best way: it strikes the best balance between ampliation and epistemic warrant. So, the paper to follow will aim to offer a philosophical vindication of the recent interest in abduction among researchers in AI.

The general framework I will follow is John Pollock's [38] analysis of defeasible reasoning in terms of the presence or absence of *defeaters*. This framework makes possible to investigate the conditions under which defeasible reasoning can issue in warranted beliefs. I shall also raise and try to answer some general philosophical questions concerning the epistemic status of abduction.

In what follows, I shall deliberately leave aside all the substantive issues about the nature of explanation.[4] This is partly because they are just too many to be dealt with in this article and partly because I think that--barring some general platitudes about the nature of explanation--my claims about IBE should be neutral vis-à-vis the main theories of explanation.[5] At any rate, I think that the very possibility of Inference to the Best Explanation as a warranted ampliative method must be examined independently of specific models of the explanatory relationship between hypotheses and evidence. Ideally, IBE should be able to accommodate different conceptions of what explanation is. This last thought implies that abduction (that is, IBE) is not usefully seen as a species of ampliative reasoning, but rather as a *genus* whose several species are distinguished by plugging assorted conceptions of explanation in the reasoning schema that constitutes the genus. So, for instance, if the relevant notion of explanation is revealing of causes, then IBE becomes an inference to the best causal explanation. Or, if the relevant notion of explanation is subsumption under laws, then IBE becomes as a kind of inference to the best Deductive-Nomological explanation, and so forth. Given that there is too much disagreement on the notion of explanation, and given that no account offered in the literature so far seems to cover fully all aspects of explanation, it seems to me methodologically useful to treat the reference to explanation in IBE as a 'placeholder' which can be spelled out in different ways in

[3] The paper [24] in this volume contains a very useful analysis of Abductive Logic Programming.

[4] These are dealt in detail in my [43].

[5] The relevant literature is really massive. Some important recent items include [27], [30] and [45].

different contexts. Some philosophers may think that this approach to IBE renders it an unnatural agglomeration of many different types of reasoning where explanatory considerations are involved. But I think it is at least premature to call this agglomeration 'unnatural'. After all, as I hope to show in this piece, the general ways in which explanatory considerations can enter into defeasible reasoning can be specified without a prior commitment to the nature of the explanatory relation.

2 Ampliation and Epistemic Warrant

Any attempt to characterise the abstract structure of scientific method should make the method satisfy two general and intuitively compelling desiderata: it should be ampliative and epistemically probative. Ampliation is necessary if the method is to deliver informative hypotheses and theories, viz., hypotheses and theories which exceed in content the observations, data, experimental results and, in general, the experiences which prompt them. This 'content-increasing' aspect of scientific method is indispensable, if science is seen, at least prima facie, as an activity which purports to extend our knowledge (and our understanding) beyond what is observed by means of the senses. But this ampliation would be merely illusory, qua increase of *content*, if the method was not epistemically probative: if, that is, it did not convey epistemic warrant to the excess content produced thus (viz., hypotheses and theories). To say that the method produces--as its output--more information than what there is in its input is one thing. To say that this extra information can reasonably be held to be warranted is quite another. Now, the real problem of the scientific method is that these two plausible desiderata are not jointly satisfiable. Or, to weaken the claim a bit, the problem is that there seems to be good reason to think that they are not jointly satisfiable. The tension between them arises from the fact that ampliation does not carry its epistemically probative character on its sleeves. When ampliation takes place, the output of the method can be false while its input is true. The following question then arises: what makes it the case that the method conveys epistemic warrant to the intended output rather than to any other output which is consistent with the input? Notice that ampliation has precisely the features that deduction lacks. Suppose one thought that a purely deductive method is epistemically probative in the following (conditional) sense: if the input (premises) is warranted, then the method guarantees that the output cannot be less warranted than the input. No ampliative method can be epistemically probative in the above sense. But can there be any other way in which a method can be epistemically probative? If the method is not such that the input excludes all but one output, in what sense does it confer any warrant on a certain output?

'In no sense', is the strong sceptical (Humean) answer. The sceptic points out that any attempt to strike a balance between ampliation and epistemic warrant is futile for the following reason. Given that ampliative methods will fail to satisfy the aforementioned conditional, they will have to base any differential epistemic treatment of outputs which are consistent with the input on some *substantive and contingent assumptions*, (e.g., that the world has a natural-kind structure, or that the world is governed by universal regularities, or that observable phenomena have unobservable causes, etc.). It is these substantive assumptions that will do all the work in conferring epistemic warrant on some output rather than another. But, the sceptic goes on, what else, other than ampliative reasoning itself, can possibly establish that these

substantive and contingent assumptions are true of the world? Arguing in a circle, the sceptic notes, is inevitable and this simply means, he concludes, that the alleged balance between ampliation and epistemic warrant carries no rational compulsion with it. In other words, the sceptic capitalises on the fact that in a purely deductive (non-ampliative) method, the transference of the epistemic warrant from the premises to the conclusion is parasitic on their formal (deductive) relationship, whereas in an ampliative method the alleged transference of the epistemic warrant from the premises to the conclusion depends on substantive (and hence challengeable) background beliefs and considerations.[6]

A standard answer to the problem of method is to grant that the sceptic has won. But I think this is too quick. Note that the sceptical challenge is far from intuitively compelling. It itself relies on a substantive *epistemic* assumption: that any defence of an ampliative but epistemically probative method should simply mirror some formal relations between the input and the output of the method and should depend on no substantive and contingent assumptions whose truth cannot be established by independent means. This very assumption is itself subject to criticism.[7] First, if it is accepted, it becomes *a priori* true that there can be no epistemically probative ampliative method. Yet, it may be reasonably argued that the issue of whether or not there can be an ampliative yet epistemically probative method should hinge on information about the actual world and its structure (or, also on information about those possible worlds which have the same nomological structure as the actual). A proof that a method could be both ampliative and epistemically probative in *all* possible worlds (that is, a proof which we have *no* reasons to believe is forthcoming) would certainly show that it can have these features in the actual world. But the very request of such a proof (one that could persuade the sceptic) relies on the substantive assumption that an epistemically probative method should be totally insensitive to the actual features (or structure) of the world. This request is far from compelling. After all, we primarily need our methods to be the right ones for the world we live in. If the range of their effectiveness is larger, then that's a pleasant bonus. But we can live without it. Second, if the sceptical assumption is accepted, even the possibility of epistemically probative demonstrative reasoning becomes dubious. For truth-transmission, even though it is guaranteed by deductive reasoning, requires some truths to start with. Yet, the truth of any substantive claims that feature in the premises of a deductive argument can only be established by ampliative reasoning, and hence it is equally open to the sceptical challenge.[8] Naturally, the point here is not that relations of deductive entailment between some premise P and a conclusion Q fail to offer an epistemic warrant for accepting Q, *if one already warrantedly accepts P*. Rather, the point is that coming to accept as true a premise P with any serious content will typically involve some ampliative reasoning. The sceptical challenge is not

[6] Philosophical attempts to offer circular justifications of ampliative modes of reasoning have been analysed in [40, chapter 4] and in [32].

[7] For a rather compelling criticism of the sceptical challenge to induction and of its philosophical presuppositions, see [35].

[8] It might be claimed that some self-evident beliefs are ampliative and yet certain enough to be the deductive foundations of all knowledge. But a) it is contentious whether there are such beliefs; and b) even if there were, they would have to be implausibly rich in content, since deduction cannot create any new content.

incoherent. But if its central assumption is taken seriously, then what is endangered is not just the very possibility of any kind of learning from experience, but also any kind of substantive reasoning.

There is, however, something important in a mild reading of the sceptical answer to the problem of method: if we see it as a challenge to offer a satisfactory account of method which is both ampliative and epistemically probative, then we can at least make some progress in our attempt to understand under what conditions (and under what substantive assumptions) the two desiderata can co-exist.

3 Between Two Extremes

In order to start making this progress, we need to see how the two standard accounts of scientific method fare vis-à-vis the two desiderata. So, we'll look at Enumerative Induction (EI) and crude hypothetico-deductivism (HD) (or, the 'method of hypothesis') and compare them in terms of the strength of ampliation and the strength of the epistemic warrant. But let me first make an important note.

3.1 Defeasibility and Defeaters

The very idea of ampliation implies that the outcome of the application of an ampliative method (or of a mode of ampliative reasoning) can be defeated by new information or evidence. So, unlike deductive methods, ampliative methods are *defeasible*. The issue here is not just that further information can make the output not to logically follow from the input. It is rather that further information can remove the *warrant* for holding the output of the method. So, further information can make the previous input not be strong enough to warrant the output. Following Pollock ([38] chapter 2, section 3; [39]), we can call "prima facie" or "defeasible" any type of reason which is not conclusive (in the sense that it is not deductively linked with the output it is a reason for). Given that ampliative reasoning is defeasible, we can say that such reasoning provides *prima facie warrant* for an output (belief). What Pollock has rightly stressed is that to call a warrant (or a reason) prima facie is not to degrade it, *qua* warrant or reason. Rather, it is to stress that a) it can be defeated by further reasons (or information); and b) its strength, *qua* reason, is a function of the presence or absence of "defeaters". "Defeaters" are the factors (generally, reasons or information) that, when they are taken into account, can remove the prima facie warrant for an outcome (belief). On Pollock's insightful analysis of reasoning and warrant, the presence or absence of defeaters is directly linked with the degree to which one is warranted to hold a certain belief. Suppose that a subject S has a prima facie (nonconclusive) reason R to believe Q. Then S is warranted to believe that Q on the basis of R, *unless* either there are further reasons R' such that, were they to be taken into account, they would lead S to doubt the integrity of R as a reason for Q, or there are strong (independent) reasons to hold not-Q. Generalising this idea to the problem of method, we may say that the presence or absence of defeaters is directly linked with the degree to which an ampliative method can confer epistemic warrant on an outcome, that is, the degree to which it can be epistemically probative. So, to say that S is prima facie warranted to accept the outcome Q of an ampliative method is to say that although it is possible that there are defeaters of the outcome Q, such

defeaters are not actual. In particular, it is to say that S has considered several possible defeaters of the reasons offered for this outcome Q and has shown that they are not present. If this is done, we can say that there are no *specific* doubts about the outcome of the method and, that belief in this outcome is prima facie warranted.

This talk of defeaters is not abstract. There are general *types* of defeater that one can consider. Hence, when it comes to considering whether an outcome is warranted, there are certain things to look at such that, if present, they would remove the warrant for the outcome. Even if it is logically possible that there could be considerations that would undercut the warrant for the outcome (a possibility that follows from the very idea of defeasibility), the concrete issue is whether or not there actually are such considerations (actual defeaters).[9] Besides, if the reasoner has done whatever she can to ensure that such defeaters are not present in a particular case, there is a strong sense in which she has done what it can plausibly be demanded of her in order to be epistemically justified. Pollock ([38], 38-39) has identified two general types of defeater: "rebutting" and "undercutting". Suppose, for simplicity, that the ampliative method offers some prima facie reason P for the outcome Q. A factor R is called a rebutting defeater for P as a reason for Q if and only if R is a reason for believing not-Q. And a factor R is called an undercutting defeater for P as a reason for Q if and only if R is a reason for denying that P offers warrant for Q.[10] So, considering whether or not Q is warranted on the basis of P one has to consider whether or not there are rebutting and undercutting defeaters. Taking all this into account, let us look at the two extreme cases of ampliative method.

3.2 Enumerative Induction

Enumerative Induction (EI) is based on the following: if one has observed *n* As being B and *no* As being not-B, and if the evidence is enough and variable, then one should infer that (with high probability) 'All As are B'. The crux of EI is that ampliation is effected by *generalisation*. We observe a pattern among the data (or, among the instances of two attributes), and then generalise it so that it covers all the values of the relevant variables (or all instances of the two attributes). For obvious reasons, we can call EI, the "more-of-the-same" method (cf. [31], 16). The prime advantage of EI is that it is content-increasing in a, so to speak, 'horizontal way': it allows the acceptance of generalisations based on observed evidence in a way that stays close to what has been actually observed. In particular, no new entities (other then those referred to in (descriptions of) the data) are introduced by the ampliation. Let me call this *minimal ampliation*. The basic substantive assumptions involved in this ampliation are that a) there are projectable regularities among the data; and b) the pattern detected among the data (or the observations) in the sample is representative of the pattern (regularity) in the whole relevant population. The prima facie warrant that EI confers on its

[9] As Pollock ([38], 39) notes the mere presence of a defeater R' is not enough to remove the prima facie warrant for a belief Q. For, being itself a reason, R' might also be subject to defeaters. Hence, faced with a possible defeater R', we should examine whether R' can itself be (or actually is) defeated by other reasons (what Pollock calls "defeater defeaters").

[10] Pollock frames this in terms of the subjunctive conditional: R is a reason to deny that *P would be* true unless Q *were* true.

outcomes is based on these substantive assumptions. But this warrant--and the assumptions themselves--are subject to evaluation. EI admits of both undercutting and rebutting defeaters. If there are specific reasons to doubt that the pattern among the data can be projected to a lawful regularity in the population, then the projection is not warranted.[11] If there are specific reasons to doubt the fairness of the sample, then the projection is also no longer warranted. Note that although the sample may be unfair (e.g., the sample might involve only ravens in a certain region), the conclusion (viz., that all ravens are black) may well be true. Yet, knowing that the sample was unfair does remove the warrant for the conclusion. These are cases of undercutting defeaters. Besides, EI admits of rebutting defeaters. If we find a negative instance (e.g. a black swan) the warrant (e.g. for the conclusion that all swans are white) is completely removed.[12] So, in EI we know precisely what kind of defeaters can remove the prima facie warrant for making the ampliation (generalisation). And, on very many occasions, we a) can certify the presence or absence of defeaters; and b) we can withhold the conclusion until we have reasons to believe that the potential defeaters are not present (e.g. by making meticulous search for cases which would rebut the conclusion). Given the very specific character of defeaters in EI, and the general feasibility of the search for defeaters, we can say that EI can be *maximally epistemically probative* (among ampliative methods). Here again, the point is not that the sceptic loses. Nor is it that EI *is* maximally epistemically probative. Rather, the point is that if ampliative--and hence defeasible--methods can be warranted at all based on the presence or absence of defeaters, and given that in the case of EI we know exactly what defeaters we should look for and how to do it, EI fares best in terms of how warranted an outcome of a successful (undefeated) application of EI can be.

So, EI is minimally ampliative and maximally epistemically probative. But this is precisely the problem with EI: that what we gain in (epistemic) austerity we lose in strength (of ampliation). EI is too restrictive. It cannot possibly yield any hypothesis about the causes of the phenomena. Nor can it introduce new entities. The basic problem is that the input and the output of EI are couched in the same vocabulary: conclusions that state generalisations are necessarily couched in the vocabulary of the premises. Hence, EI cannot legitimately introduce new vocabulary. Hence, it cannot possibly be used to form ampliative hypotheses that refer to entities whose descriptions go beyond the expressive power of the premises.[13]

[11] This is essentially what Goodman [12] observed in his notorious "new riddle of induction".

[12] This may be a bit too strong, since we know that we can always fault the observation. We may, for instance, insist that the observed swan was not really black. Or we may make it part of the meaning of the term 'swan' that all swans are white. On this last move, a black swan cannot really be a swan. But such manoeuvres, though logically impeccable, do not always have the required epistemic force to save the generalisation from refutation. In any case, in EI we know exactly what sort of manoeuvres we have to block in order to render a generalisation rebutted.

[13] Goodman-type stories of the form 'All observed emeralds are green. Therefore, all emeralds are grue' involve a different vocabulary between premises and conclusion only in a trivial way. For predicates such as 'grue' are fully definable in terms of the vocabulary of the premises (plus other antecedently understood vocabulary). So, for instance, 'grue' is defined as: 'green if observed before 2001 and blue thereafter'.

3.3 The Method of Hypothesis

Let us turn to the crude version of the 'method of hypothesis' (HD). This is based on the following: Form a hypothesis H and derive some observational consequences from it. If the consequences are borne out, then the hypothesis is confirmed (accepted). If they are not borne out, then the hypothesis is disconfirmed (rejected). So, the crux of the method is that a hypothesis is warrantedly accepted on the basis of the fact that it entails all available relevant evidence. In HD, ampliation is effected by *confirmation*. An ampliative hypothesis H is accepted because it gets confirmed by the relevant evidence. To be sure, the operation of HD is more complicated. The observational consequences follow from the conjunction of H with some statements of initial conditions, other auxiliary assumptions and some bridge-principles which connect the vocabulary in which H is couched and the vocabulary in which the observational consequences are couched. It is this bridge-principles that make HD quite powerful, since they allow for what I shall call 'vertical extrapolation'--to be contrasted with the 'horizontal extrapolation' characteristic of EI. The content of H may well be much richer than the content of the relevant observational consequences and the deductive link between the two contents is guaranteed by the presence of bridge-principles. The prime attraction of HD is precisely that is can be content-increasing in a, so to speak, 'vertical way': it allows the acceptance of hypotheses about the, typically unobservable, causes of the phenomena. In particular, new entities (other then those referred to in the data) are introduced by the ampliation. So, in contrast to EI, let me call this *maximal ampliation*. The basic substantive assumptions involved in this type of ampliation are that a) there are causally and explanatory relevant entities and regularities behind the observed data or phenomena; and b) the pattern detected among the data (or the observations) is the causal-nomological outcome of entities and processes behind the phenomena. What about the warrant that HD confers on its outcomes? As in the case of EI, we should look at the possible defeaters of the reasons offered by HD for the acceptance of a hypothesis H. The rebutting defeaters seem to be clear-cut: if the predicted observation is not borne out, then--by modus tollens--the hypothesis is refuted. This seems quite compelling, yet there are well-known problems. As we have just seen, it is typically the case that, in applications of HD, the predictions follow from the conjunction of the hypothesis with other auxiliary assumptions and initial and boundary conditions. Hence, when the prediction is *not* borne out, it is the whole cluster of premises that gets refuted. But HD alone cannot tell us how to apportion praise and blame among them. At least one of them is false but the culprit is not specified by HD. It might be that the hypothesis is wrong, or some of the auxiliaries were inappropriate. So, a possible rebutting defeater (a negative prediction) does not carry with it the epistemic force to defeat the hypothesis and hence to remove the warrant for it. (This is a version of the well-known Duhem-Quine problem.) In order for the rebutting defeater to do its job, we need further information, viz., whether the hypothesis is warranted enough to be held on, or whether the auxiliaries are vulnerable to substantive criticism etc. But all these considerations go a lot beyond the deductive link between hypotheses and data that forms the backbone of HD and are not incorporated by the logical structure of HD. What about the undercutting defeaters? Here, it's not clear what these are. It seems a good idea to say that an undercutting defeater for a hypothesis H which does conform to the observations is another hypothesis H* which also conforms to the observations.

For if we know that there is another H*, then it seems that our confidence about H is negatively affected. The prima facie warrant for H (based as it is on the fact that H entails the evidence) may not be totally removed, but our confidence that H is correct will surely be undermined. To put the same point in a different way, if our warrant for H is solely based on the fact that it entails the evidence, then insofar as there is another hypothesis H* which also entails the evidence, H and H* will be equally warranted. It may be that H* entails H, which means that, on probabilistic considerations, H will be at least as probable as H*. But this is a special case. The general case is that H and the alternative hypothesis H* will be mutually inconsistent. Hence, HD will offer no way to discriminate between them in terms of warrant. The existence of each alternative hypothesis will act as an undercutting defeater for the rest of them. Given that, typically, for any H there will be alternative hypotheses which also entail the evidence, HD suffers from the existence of just too many undercutting defeaters. All this can naturally lead us to the conclusion that HD is *minimally epistemically probative*, since it does not have the resources to show how the undercutting defeaters can be removed.[14]

So, HD is maximally ampliative and minimally epistemically probative. But this is precisely the problem with it: that what we gain in strength (of ampliation) we lose in (epistemic) austerity. Unlike EI, it can lead to hypotheses about the causes of the phenomena. And it can introduce new entities. That is, it can also be 'vertically ampliative'. But, also unlike EI, HD is epistemically too permissive. Since there are, typically, more than one (mutually incompatible) hypothesis which entail the very same evidence, if a crude 'method of hypothesis' were to license any of them as probably true, it would also have to license all of them as probably true. But this permissiveness leads to absurdities. The crude 'method of hypothesis' simply lacks the discriminatory power that scientific method ought to have.[15]

4 A Case for Abduction

Faced with these two extreme solutions to the problem of the scientific method, the question is whether there can be a characterisation of the method that somehow moves in-between them. So far, we have noted that ampliation is inversely proportional to epistemic warrant. This is clearly not accidental, since ampliation amounts to risk and the more the risk taken, the less the epistemic security it enjoys. But it is an open issue whether or not there can be a way to strike a balance between ampliation and epistemic warrant, or (equivalently) between strength and austerity. In particular, it is an open issue whether there can be a characterisation of the method which strikes a balance between EI's restrictive ampliation and HD's epistemic permissiveness. I want to explore the suggestion that abduction, if suitably understood as Inference to the Best Explanation (IBE), can offer the required trade-off. But first, what is abduction?

[14] For a telling critique of hypothetico-deductivism see [29]. However, Laudan wrongly assimilates Inference to the Best Explanation to hypothetico-deductivism.

[15] It may be objected that EI is equally epistemically permissive since, on any evidence, there will be more than one generalisation which entails it. Yet in order to substantiate this claim for the case of EI, one is bound to produce alternative generalisations which either are non-projectible or restate merely sceptical doubts (e.g., that all ravens are black when someone observes them).

4.1 What Is Abduction?

I am going to leave aside any attempt to connect what follows with Peirce's views on abduction.[16] Rather, I shall take Harman's [15] as the *locus classicus* of the characterisation of IBE. "In making this inference", Harman notes, "one infers, from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis. In general, there will be several hypotheses that might explain the evidence, so one must be able to reject all such alternative hypotheses before one is warranted in making the inference. Thus one infers, from the premise that a given hypothesis would provide a 'better' explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true" (1965, 89). Following Josephson ([22], 5), IBE can be put schematically thus (A):

> D is a collection of data (facts, observations, givens).
> H explains D (would, if true, explain D)
> No other hypothesis can explain D as well as H does.
>
> ---
>
> Therefore, H is probably true.[17]

It is important to keep in mind that, on IBE, it is not just the semantic relation between the hypothesis and the evidence which constitutes the prima facie warrant for the acceptance of the hypothesis. Rather, it is the *explanatory quality* of this hypothesis, on its own but also taken in comparison to others, which contributes essentially to the warrant for its acceptability. So, what we should be after here is a kind of measure of the explanatory power of a hypothesis. Explanatory power is connected with the basic function of an explanation, viz., providing understanding. Whatever the formal details of an explanation, it should be such that it enhances our understanding of why the explanandum-event happened. This can be effected by incorporating the explanandum into the rest of our background knowledge by providing some link between the explanandum and other hypotheses that are part of our background knowledge. Intuitively, there can be better and worse ways to achieve this incorporation--and hence the concomitant understanding of the explanandum. For instance, an explanation which does not introduce gratuitous hypotheses in the explanatory story it tells, or one that tallies better with the relevant background knowledge, or one that by incorporating the explanandum in the background knowledge it enhances its unity, offers a better understanding and, hence has more explanatory power.

I think the evaluation of explanatory power takes place in two directions. The *first* is to look at the specific background information (beliefs) which operate in a certain application of IBE. The *second* is to look at a number of structural features (standards) which competing explanations might possess. The prime characteristic of IBE is that it cannot operate in a "conceptual vacuum", as Ben-Menahem ([2], 330) put it. Whatever else one thinks of an explanation, it must be such that it establishes

[16] For Peirce's views the interested reader should look at [4], [8], [14], [47] and [9].

[17] Here I am using the word 'probably' with no specific interpretation of the probability calculus in mind. Its use implies only that the conclusion does not follow from the premises in the way that a deductive argument would have it.

some causal-nomological connection between the explanandum and the explanans. The details of this connection--and hence the explanatory story that they tell--will be specified relative to the available background knowledge. So, to say that a certain hypothesis H is the best explanation of the evidence is to say, at least in part, that the causal-nomological story that H tells tallies best with background knowledge. This knowledge must contain all relevant information about, say, the types of causes that, typically, bring about certain effects, or the laws that govern certain phenomena etc. At least in non-revolutionary applications of IBE, the relevant background knowledge can have the resources to discriminate between better and worse potential explanations of the evidence. So, the explanatory power of a potential explanation depends on what other substantive information there is available in the background knowledge.[18] Let me call 'consilience' this feature of IBE which connects the background knowledge with the potential explanation of the evidence.

Consilience: Suppose that there are two potentially explanatory hypotheses H_1 and H_2 but the relevant background knowledge favours H_1 over H_2. Unless there are specific reasons to challenge the background knowledge, H_1 should be accepted as the best explanation.

Yet, to a certain extent, there is room for a structural specification of the best explanation of a certain event (or piece of evidence). That is, there are structural standards of explanatory merit which mark the explanatory power of a hypothesis and which, when applied to a certain situation, rank competing explanations in terms of their explanatory power. These standards operate crucially when the substantive information contained in the relevant background knowledge cannot forcefully discriminate between competing potential explanations of the evidence. The following list, far from being complete, is an indication of the relevant standards.[19]

Completeness: Suppose that only one explanatory hypothesis H explains all data to be explained. That is, all other competing explanatory hypotheses fail to explain some of the data, although they are not refuted by them. H should be accepted as the best explanation.
Importance: Suppose that two hypotheses H_1 and H_2 do not explain all relevant phenomena, but that H_1, unlike H_2, explains the most salient phenomena. Then H_1 is to be preferred as a better explanation.
Parsimony: Suppose that two composite explanatory hypotheses H_1 and H_2 explain all data. Suppose also that H_1 uses fewer assumptions than H_2. In particular, suppose that the set of hypotheses that H_1 employs to explain the data is a proper subset of the hypotheses that H_2 employs. Then H_1 is to be preferred as a better explanation.

[18] A reader has pressed me to explain how the background knowledge can discriminate among competing hypotheses that, if true, would explain a certain explanandum. I don't think there is a deep mystery here. In a lot of typical cases where reasoners employ IBE, there is just one 'best explanation' that the relevant background knowledge makes possible. Finding it consists in simply searching within the relevant background knowledge. For more on this issue, and for an interesting scientific example, see [40], 217-219.
[19] For a fuller discussion see [48].

Unification: Suppose that we have two composite explanatory hypotheses H^k and H^j a body of data $e_1,...,e_n$. Suppose that for every piece of data e_i (i=1,...,n) to be explained H^j introduces an explanatory assumption H^j_i such that H^j_i explains e_i. H^k, on the other hand, subsumes the explanation of all data under a few hypotheses, and hence it unifies the explananda. Then H^k is a better explanation than H^j.
Precision: Suppose that H_1 offers a more precise explanation of the phenomena than H_2, in particular an explanation that articulates some causal-nomological mechanism by means of which the phenomena are explained. Then H_1 is to be preferred as a better explanation.

Such standards have a lot of intuitive pull. Besides, they can characterise sufficiently well several instances of application of IBE in scientific practice (cf. [46], [48]). But even if one granted that these standards have some genuine connection with explanatory quality or merit, one could question their epistemic status: why are they anything more than pragmatic virtues? (cf. [51]) If to call a certain virtue 'pragmatic' is to make it non-cognitive, to relegate it to a merely self-gratifying 'reason' for believing things, then it should be clear that the foregoing explanatory virtues (standards) are not pragmatic. For they possess a straight cognitive function. As Thagard [49] has persuasively argued, such standards safeguard the explanatory coherence of our total belief corpus as well as the coherence between our belief corpus and a new potential explanation of the evidence. To say that a hypothesis that meets these standards has the most explanatory power among its competitors is to say that it has performed best in an explanatory coherence test among its competitors. Explanatory coherence is a cognitive virtue because, on some theories of justification at least, it is a prime way to confer justification on a belief or a corpus of beliefs (cf. [3], [17]). Naturally, the warrant conferred on the chosen hypothesis, viz., that it fares better than others in an explanatory-quality test and that, as a result of this, it enhances the explanatory coherence of the belief corpus, is a defeasible warrant. But this is as it should be. The problem might be thought to be that there is no algorithmic way to connect all these criteria (with appropriate weights) so that they always engender a clear-cut ranking. And the obvious rivalries among some of the criteria suggest that a lot of judgement should be exercised in this ranking. Such problems would be fatal only for those who thought that a suitable description of the method would have to be algorithmic, and in particular that it would have to employ a simple and universal algorithm. This aspiration should not have been taken seriously in the first place. Note also that although a simple and universal algorithm for IBE is not possible, there have been implementations of IBE, e.g., by Thagard [49] which employ a variety of algorithms. Besides, although IBE may be characterised at a very general and abstract level in the way presented above, there is good reason to think that many specific applications (e.g., in medical diagnosis) may employ important domain-specific criteria which require more careful empirical study.

4.2 Some Philosophical Issues

Some philosophers have expressed doubts about IBE which are based on the following worry: why should the information that a hypothesis is the best explanation

of the evidence be a prima facie *reason* to believe that this hypothesis is true (or likely to be true)? Cartwright ([5], 4) for instance, has argued that the foregoing question cannot be successfully answered.[20] Meeting this challenge will have to engage us in a proper understanding of the interplay between substantive background knowledge and considerations of explanatory coherence in rendering IBE a legitimate mode of inference. Those readers who feel that these doubts are ill-motivated or just philosophical can skip the rest of this section.

So, what sort of *inference* is IBE? There are two broad answers to this. (1) We infer to the probable truth of the likeliest explanation insofar as and because it is the *likeliest* explanation. On this answer, what matters is how likely the explanatory hypothesis is. If it is likely we infer it; if it isn't we don't. (2) The best explanation, *qua* explanation, is likely to be true (or, at least more likely to be true than worse explanations). That is, the fact that a hypothesis H is the *best* explanation of the evidence issues a warrant that H is likely. In his ([31], 61-65), Lipton has noted that the first answer views IBE as an inference to the Likeliest Potential Explanation, while the second views it as an inference to the Loveliest Potential Explanation. The loveliest potential explanation is "the one which would, if correct, be the most explanatory or provide the most understanding" (op.cit., p.61). If we go for the Likeliest Potential Explanation, then Cartwright's challenge evaporates. For, best explanation and epistemic warrant are linked *externally* via some considerations of likelihood.[21] If there are reasons to believe that a certain hypothesis is likely (or the likeliest available), then there is no further issue of epistemically warranted acceptance. But if we go for the Likeliest Potential Explanation (i.e., the first answer above) then, IBE loses all of its excitement. For what is particularly challenging with IBE is the suggestion--encapsulated in answer (2) above--that the fact that a hypothesis is the *best* explanation (i.e. the loveliest one) *ipso facto* warrants the judgement that it is likely. If the loveliness of a potential explanation is shown to be a symptom of its truth, then Cartwright's challenge is met in a significant and *internal* way.[22] Lipton's own strategy has been to impose two sorts of filters on the choice of hypotheses. One selects a relatively small number of potential explanations as plausible, while the other selects the best among them as the actual explanation. Both filters should operate with explanatory considerations. That is, both filters should act as explanatory-quality tests. Still, although plausibility might have to do with explanatory considerations, why should plausibility have anything to do with likelihood? Here, Lipton's answer is to highlight the *substantive assumptions* that need

[20] She does believe however in a special case of IBE, viz., inference to the most likely cause (cf. [5], 6).

[21] Note that here I am using the term "likelihood" informally and not in the statistical sense of it. An attentive reader has pressed me to elaborate on the possible relation between IBE and Bayesianism. I have attempted to offer a few thoughts on this matter in [42]. Suffice it to say here that I take IBE to be a way to assign a kind of objective prior probabilities to hypotheses whose posterior degree of confirmation--in light of further evidence for them--can be calculated by Bayesian techniques.

[22] Failure to discriminate between the Likeliest and the Loveliest Explanation seems to be the reason why Ben-Menahem ([2], 324) claims that "[t]here is nothing particularly deep about the inference to the best explanation. At least there is nothing particularly deep about it qua type of inference".

to be in place for IBE (as Inference to the Loveliest Potential Explanation) to be possible. Explanatory considerations enter into the first filter (that of selecting a small number of hypotheses) by means of our substantive background knowledge that favours hypotheses that cohere well with (or are licensed by) our background beliefs (cf. [31], 122). Insofar as these background beliefs are themselves likely, then IBE operates within an environment of likely hypotheses. Given that the background beliefs themselves have been the product of past applications of IBE, they have been themselves imputed by explanatory considerations. So, the latter enter implicitly in the first filter and explicitly in the second (that of choosing the best among the competing hypotheses that are licensed by the background beliefs). We can see the crux of all this by looking at Josephson's aforementioned schema (A) for IBE. The crucial judgement for the inference to take place is that no other hypothesis explains D as well as H. This judgement is the product of a) filtering the competing hypotheses according to substantive background knowledge and b) choosing among *them* by explanatory considerations. The upshot of all this is that the application of IBE relies on substantive background knowledge. Without it, IBE as an *inference* is simply impotent.[23] But notice that the structural features that make an explanation better than another are part and parcel of the background knowledge. They are just this more abstract part of it which tells us how to evaluate potential explanations. Notice also that these general structural features are complemented by particular ones when it comes to specific applications of IBE. As Josephson ([22], 14) has noted, in specific cases the likelihood of the chosen 'best explanation' H will depend on considerations such as "how decisively H surpasses the alternatives" and "how much confidence there is that all plausible explanations have been considered (how thorough was the search for alternative explanations)".

But suppose that all this is not convincing. Suppose, that is, that we haven't made a case for the claim that the best (loveliest) explanation and the likeliest explanation may reasonably be taken to coincide in light of the relevant background knowledge. There is still an indirect answer available to Cartwright's challenge. Note that we are concerned with the prima facie warrant for accepting a hypothesis H. The question then is: is the fact that H is rendered the best explanation of the evidence a prima facie reason for its acceptance? If, following Pollock ([38], 124), we view justification as "epistemic permissibility", it is obvious that the answer to the foregoing question can only be positive. For to say that the fact that H is the best explanation of the evidence is a *reason* for the acceptance of H is to say that a) it is all right (i.e., it is permissible) to believe in H on this basis; and b) that this permissibility is grounded on the explanatory connection between H and the evidence. It is this explanatory connection which makes the acceptance of H prima facie reasonable since it enhances the coherence of our total belief corpus. By incorporating H in our belief corpus BC as the best explanation of the evidence we enhance the capacity of BC to deal with new information and we improve our understanding not just of why the evidence is the way it is but also of how this evidence gets embedded in our belief corpus. To see how all this works out, note the following. It is explanatory (causal-nomological) connections which hold our belief corpus together. It is such connections which organise the individual beliefs that form it and make the corpus useful in understanding, planning, anticipating etc. (cf. [16]). Faced with a choice among competing explanatory

[23] I have defended the reliability of IBE in some detail in my ([40], 81-90 & 212-2).

hypotheses of some event, we should appeal to reasons to eliminate some of them.[24] Subjecting these hypotheses to an explanatory-quality test is the prime way to afford these reasons. Those hypotheses which fare badly in this test get eliminated. For, by having done badly in the test, they have failed at least some of the intuitively compelling criteria of explanatory power. So, they have either failed to cohere well with the relevant background information, or have left some of the data unaccounted for, or have introduced gratuitous assumptions into the explanatory story, or what have you. If this test has a clear winner (the best explanation), then this is the only live option for acceptance. In the end, what IBE does is to enhance the explanatory coherence of a background corpus of belief by choosing a hypothesis which brings certain pieces of evidence into line with this corpus. And it is obviously reasonable to do this enhancement by means of the best available hypotheses. This coherence-enhancing role of IBE, which has been repeatedly stressed by Harman ([16], [17], [18]), Lycan [33] and Thagard ([46], [49]), is ultimately the warrant-conferring element of IBE.

Some philosophers think that there may be a tension between the two prime aspects of IBE that I have described above, viz., its reliance on considerations of explanatory coherence and its dependence on substantive background beliefs. Day and Kincaid ([6], 275) for instance, argue that if IBE is primarily seen as relying on considerations of explanatory coherence, it becomes "redundant and uninformative". For it reduces to "nothing more than a general admonition to increase coherence ([6], 279). And if IBE is primarily seen as being dependent on substantive background knowledge, it "does not name a fundamental pattern of inference" ([6], 282). Rather, they argue, it is an instance of a strategy "that infers to warranted beliefs from background information and the data", without necessarily favouring an explanatory connection between hypotheses and the data (cf. ibid.). Day and Kincaid favour a *contextual* understanding of IBE, since, they say, it has "no automatic warrant" and its importance "might well differ from one epistemic situation to the next" ([6], 282). I think, however, that a) the two aspects of IBE are not in any tension; and b) they engender a rather general and exciting mode of ampliative reasoning. Certainly, more work needs to be done on the notion of coherence and its link with explanation. But if we adopt what Lycan [33] has called "explanationism", it should be clear that explanatory coherence is a vehicle through which an inference is performed and justified. IBE is the mode of inference which effects ampliation via explanation and which licenses conclusions on the basis of considerations which increase explanatory coherence. Yet, as I have noted above, it is wrong to think that the achievement (or enhancement) of explanatory coherence is just a formal-structural matter. Whatever else it is, the best explanation of the evidence (viz., the one that is the best candidate for an enhancement of the explanatory coherence of a belief corpus) has some substantive content which is constrained (if not directly licensed) by the relevant substantive background knowledge. So, substantive background information is not just the material on which some abstract considerations of explanatory coherence should be imposed. It is also the means by which this coherence is achieved. To infer to the best explanation H of the evidence is to search within the relevant background knowledge for explanatory hypotheses and to select the one (if there is one) which

[24] Normally, we need to eliminate all but one of them (insofar as they are mutually incompatible, of course), but we should surely allow for ties.

makes the incorporation of the evidence into this background corpus the most explanatorily coherent one. The selection, as we have seen, will be guided by both the substantive background knowledge and some relatively abstract structural standards. That this process is not an inference can be upheld only if one entertains the implausible views that to infer is to deduce and that to infer is to have "an automatic warrant" for the inference. Not all changes in the background knowledge will be based on explanatory considerations. But given that some (perhaps most) are, IBE will have a distinctive (and exciting) role to play.

To sum up, the prima facie reasonableness of IBE cannot be seriously contested. Even if one can question the link between best explanation and truth, one cannot seriously question that the fact that a hypothesis stands out as the best explanation of the evidence offers defeasible reasons to warrantedly accept this hypothesis.[25]

4.3 Abduction and the Two Desiderata

This preliminary defence of the reasonableness of IBE was necessary in order to dispel some natural doubts towards it.[26] Now, we need to see how IBE fares vis-à-vis EI and HD. I will suggest that both EI and HD are extreme cases of IBE, but while EI is an interesting limiting case, HD is a degenerate one whose very possibility shows why IBE is immensely more efficient. Besides, I will argue that IBE has all the strengths and none of the weaknesses of either EI or HD.

That proper inductive arguments are instances of IBE has been argued by Harman [16] and been defended by Josephson ([22], [23]) and Psillos [42]. The basic idea is that good inductive reasoning involves comparison of alternative potentially explanatory hypotheses. In a typical case, where the reasoning starts from the premise that 'All As in the sample are B', there are (at least) two possible ways in which the reasoning can go. The first is to withhold drawing the conclusion that 'All As are B', even if the relevant predicates are projectable, based on the claim that the observed correlation in the sample is due to the fact that the sample is biased. The second is to draw the conclusion that 'All As are B' based on the claim that that the observed correlation is due to the fact that there is a nomological connection between being A and being B such that All As are B. This second way to reason implies (and is supported by) the claim that the observed sample is not biased. What is important in any case is that which way the reasoning should go depends on explanatory considerations. Insofar as the conclusion 'All As are B' is accepted, it is accepted on the basis it offers a better explanation of the observed frequencies of As which are B in the sample, in contrast to the (alternative potential) explanation that someone (or something) has biased the sample. And insofar as the generalisation to the whole population is not accepted, this judgement will be based on providing reasons that the biased-sample hypothesis offers a better explanation of the observed correlations in the sample. Differently put, EI is an extreme case of IBE in that a) the best

[25] Here I am leaving aside van Fraassen's [52] claim that the reasons for acceptance are merely pragmatic rather than epistemic. For a critical discussion of his views see ([40] 171-76) and ([20] chapter 4).

[26] Van Fraassen ([50], 160-70) suggested that IBE--conceived as a rule--is incoherent. Harman [19] and Douven [7] have rebutted this claim.

explanation has the form of a nomological generalisation of the data in the sample to the whole relevant population and b) the nomological generalisation is accepted, if at all, on the basis that it offers the best explanation of the observed correlations on the sample. HD, on the other hand, is a limiting but degenerate case of IBE in the following sense: if the only constraint on an explanatory hypothesis is that it deductively entails the data, then any hypothesis which does that is a potential explanation of the data. If there is only one such hypothesis, then it is automatically the 'best' explanation. But it is trivially so. The very need for IBE is suggested by the fact that HD is impotent, as it stands, to discriminate between competing hypotheses which entail (and hence explain in this minimal sense) the evidence.

How, then, does IBE fare vis-à-vis the two desiderata for the method, viz. ampliation and epistemic warrant? Remember that EI is minimally ampliative and maximally epistemically probative, whereas HD is maximally ampliative and minimally epistemically probative. Like HD, IBE is maximally ampliative: it allows for the acceptance of hypotheses which go far beyond the data not just in a horizontal way but also in a vertical one. And given that EI is a special case of IBE, IBE can--under certain circumstances--be as epistemically probative as EI. But unlike HD, IBE can be epistemically probative in circumstances that HD becomes epistemically too permissive. For IBE has the resources to deal with the so-called 'multiple explanations' problem (cf. [42], 65). That is, IBE can rank competing hypotheses which all, prima facie, explain the evidence in terms of their explanatory power and therefore evaluate them.[27] In order to see how this evaluative dimension of IBE can issue in epistemic warrant, let us examine the types of defeaters to the reasons offered by IBE.

Recall from section 3 that to say that one is prima facie warranted to accept the outcome of an ampliative method is to say that one has considered several possible defeaters of the reasons offered for this outcome and has shown that they are not present. If this is done, we noted there, there are no *specific* doubts about the warrant for the outcome of the method. Recall also that there are two general types of defeater, rebutting and undercutting ones. Naturally, if there is an observation which refutes the best explanation of the evidence so far, then this is a rebutting defeater of the best explanation. But IBE fares better than HD vis-à-vis the Duhem-Quine problem. For, although any hypothesis can be saved from refutation by suitable adjustments to some auxiliary assumptions (and hence although any rebutting defeater can be neutralised), IBE can offer means to evaluate the impact of a recalcitrant piece of evidence on the conclusion that the chosen hypotheses is the best explanation of the evidence. HD does not have the resources to perform this evaluation. If the sole constraint on the acceptance of the hypothesis is whether or not it entails the evidence, it is clear that a

[27] As one of the anonymous readers observed, abduction, as this is typically used in Logic Programming, does not require ranking of competing hypotheses in terms of their explanatory power. In particular, it does not require that no other hypothesis be a better explanation than the one actually chosen. This is indeed so. But, as I have argued [42], this is precisely the problem that suggests that the computational modelling of abduction in Logic Programming should be more complicated than it actually is. In many cases of abductive Logic Programming it is already a difficult (and valuable) task to generate an explanation of a certain event. But, as many advocates of abductive Logic Programming are aware, there will typically be competing explanations of the event to be explained (cf. [25]). So there is bound to be need to discriminate between them in terms of their explanatory power. This point of view is also entertained by [24] in this volume.

negative observation can only refute the hypothesis. If the hypothesis is to be saved, then the blame should be put on some auxiliaries, but--staying within HD--there is no independent reason to do so. In IBE, the required independent reasons are provided by the relevant explanatory considerations: if there are strong reasons to believe that a hypothesis is the best explanation of the evidence, there is also reason to stick to this hypothesis and make the negative observation issue in some changes to the auxiliary assumptions. After all, if a hypothesis has been chosen as the best explanation, then it has fared best in an explanatory-quality test with its competing rivals. So unless there is reason to think that it is superseded by an even better explanation, or unless there is reason to believe that the recalcitrant evidence points to one of the rivals as a better explanation, to stick with the best explanatory hypothesis is entirely reasonable. This last thought brings us to the role of undercutting defeaters in IBE. Recall that in the case of HD, any other hypothesis which entails the same evidence as H is an undercutting defeater for (the warrant for) H. And given that there are going to be a lot of such alternative hypotheses, the warrant for H gets minimised. But in IBE it is simply not the case that any other hypothesis which entails the evidence offers an explanation of it. For it is not required that the explanatory relation between the evidence and the hypothesis be deductive (cf. [31], 96).[28] Even if we focus on the special case in which this relation is deductive, IBE dictates that we should look beyond the content of each potential explanatory hypothesis and beyond the relations of deductive entailment between it and the evidence in order to appraise its explanatory power. Two or more hypotheses may entail the same evidence, but one of them may be a better explanation of it. So, the presence of a worse explanation cannot act as a possible undercutting defeater for the acceptance of the best explanatory hypothesis. The choice of the best explanation has already involved the consideration of possible undercutting defeaters (viz., other potential explanations of the evidence) and has found them wanting. The judgement that a certain hypothesis is the best explanation of the evidence is warranted precisely because it has rested on the examination and neutralisation of possible undercutting defeaters. To be sure, IBE is defeasible. And the discovery of an even better explanation of the evidence will act as an undercutting (sometimes even as a rebutting defeater) of the chosen hypothesis. But this is harmless for two reasons. First, given the information available at a time *t*, it is reasonable to infer to the best available explanation H of the present evidence even if there may be even better possible explanations of it. The existence of hitherto unthought of explanations is a contingent matter. H has fared in the explanatory-quality test better than its extant competitors. Hence it has neutralised a number of possible undercutting defeaters. That there may be more possible undercutting defeaters neither can be predicted, nor can it retract from the fact that it is prima facie reasonable to accept H. In any case, if the search for other potential explanations has been thorough, and if the present information does not justify a further exploration of the logical space of potentially explanatory hypotheses, there is no *specific* reason to

[28] A hypothesis might explain an event without entailing it. It might make it occurrence probable; or it might be such that it makes the occurrence of the event more probable than it was before the explanatory hypothesis was taken into account. More generally, IBE should be able to take the form of statistical explanation either in the form of the Hempelian Inductive-Statistical model (cf. [21]) or in the form of Salmon's Statistical-Relevance model (cf. [44]).

doubt that the current best explanation is simply the best explanation. If such doubts arise later on they are welcome, but do not invalidate our present judgement.[29]

The natural conclusion of all this is that IBE admits of clear-cut undercutting defeaters, but unlike HD it has the resources to show when a potential undercutting defeater can be neutralised. And it also admits of clear-cut rebutting defeaters, but unlike HD it can explain how and why such a possible defeater can be neutralised. So, when its comes to its epistemically probative character, IBE can reach the maximal epistemic warrant of EI (since EI is an extreme case of IBE), but it goes far beyond the minimal epistemic warrant of HD (since it offers reasons to evaluate competing hypotheses in an explanatory-quality test). And when it comes to ampliation, like HD and unlike EI, it reaches up to maximal ampliation (cf. the following chart).

	EI	**HD**	**IBE**
Ampliation	Minimal	Maximal	Maximal
Epistemic Warrant	Maximal	Minimal	Far more than minimal and up to maximal

5 Conclusion

I have argued that abduction, understood as Inference to the Best Explanation, satisfies in the best way the two desiderata of ampliation and epistemic warrant and also strikes the best balance between the role that background knowledge plays in ampliative reasoning and the role that explanatory considerations (as linked with the demand of explanatory coherence) plays in justifying an inference. I will then conclude with a couple of issues that need more attention in future work.

One such issue is the connection between Kowalski's work on argumentation and the approach to IBE suggested in this paper. Kowalski and Toni [26] have suggested that practical reasoning can be understood as a "dialectic process" in which two reasoners present defeasible arguments in favour of their respective positions. Part of the reasoning process is, then, for each side to present defeaters for the other side's arguments. The possibility is then open that we can think of cases where the best explanation of an event is sought as cases in which reasoners argue for their favoured hypotheses being the 'best explanation' and defend it against the defeaters offered by the other side. It may indeed be useful to see how the abstract framework for argumentation that Kowalski and Toni have put forward, and which makes heavy use of defeaters, can be enlarged (or customised) to incorporate cases of conclusions reached by IBE. Obviously, more work needs to be done on the notion of explanatory coherence and also on the role of coherence in justification. But the good news so far seems to be that IBE can emerge as the general specification of scientific method which promises to solve in the best way its central philosophical problem.

[29] In his [37], Pereira makes some interesting observations as to how defeasibility considerations can be captured within Logic Programming, especially in connection with the role that negation plays within this framework.

References

1. Aliseda, A.: Seeking Explanations: Abduction in Logic. Philosophy of Science and Artificial Intelligence. ILLC Dissertation Series (1997) Amsterdam: University of Amsterdam
2. Ben-Menahem, Y.: The Inference to the Best Explanation. Erkenntnis 33 (1990) 319-344
3. BonJour, L.: The Structure of Empirical Knowledge. (1985) Cambridge MA: Harvard University Press
4. Burks A.: 'Peirce's Theory of Abduction. Philosophy of Science 13 301-306
5. Cartwright, N.: How the Laws of Physics Lie. (1983) Oxford: Clarendon Press
6. Day, T. & Kincaid, H.: Putting Inference to the Best Explanation in its Place. Synthese 98 (1994) 271-295
7. Douven, I.: Inference to the Best Explanation Made Coherent. Philosophy of Science 66 (Proceedings) (1999) S424-435
8. Fann, K.T.: Peirce's Theory of Abduction. (1970) Martinus Nijhoff
9. Flach, P. & Kakas, A.: Abductive and Inductive Reasoning: Background and Issues. In Flach, P. & Kakas, A. (eds.): Abduction and Induction: Essays on their Relation and Integration. (2000) Dordrecht: Kluwer Academic Publishers
10. Flach, P. & Kakas, A. (eds.): Abduction and Induction: Essays on their Relation and Integration. Dordrecht: Kluwer Academic Publishers
11. Fodor, G.: The Mind Doesn't Work That Way. (2000) MIT Press
12. Goodman, N.: Fact, Fiction and Forecast. (1954) Cambridge MA: Harvard University Press
13. Gower, B.: Scientific Method: An Historical and Philosophical Introduction. (1998) London: Routledge.
14. Hanson, N.R.: Notes Towards a Logic of Discovery. In Bernstein, R. J. (ed.): Critical Essays on C. S. Peirce. (1965) Yale University Press.
15. Harman, G.: Inference to the Best Explanation. The Philosophical Review 74 (1965) 88-95
16. Harman, G.: Reasoning and Explanatory Coherence. American Philosophical Quarterly 17 (1979) 151-157
17. Harman, G.: Change in View: Principles of Reasoning. (1986) Cambridge MA: MIT Press
18. Harman, G.: Rationality. In Smith, E. E. & Osherson, D. N. (eds.) An Invitation to Cognitive Science Vol. 3 (Thinking) (1995) Cambridge MA: MIT Press
19. Harman, G.: Pragmatism and the Reasons for Belief. In Kulp, C. B. (ed.) Realism/Anti-realism and Epistemology. (1996) New Jersey: Rowan & Littlefield
20. Harman, G.: Reasoning, Meaning and Mind. (1999) Oxford: Oxford University Press
21. Hempel, C.: Aspects of Scientific Explanation. (1965) New York: Basic Books
22. Josephson, J. et al.: Abductive Inference. (1994) Cambridge: Cambridge University Press
23. Josephson, J.: Smart Inductive Generalisations are Abductions. In Flach, P. & Kakas, A. (eds.) Abduction and Induction: Essays on their Relation and Integration. (2000) Dordrecht: Kluwer Academic Publishers
24. Denecker, M & A.C. Kakas.: Abduction in Logic Programming. This volume
25. Kakas, A.C., Kowalski, R.A., & Toni, F.: Abductive Logic Programming. Journal of Logic and Computation 2 (1992) 719-770
26. Kowalski, R. A. & Toni, F.: Abstract Argumentation. Artificial Intelligence and Law 4 (1996) 275-296
27. Kitcher, P.: Explanatory Unification. Philosophy of Science 48 (1981) 251-81
28. Konolige, K.: Abductive Theories in Artificial Intelligence. In Brewka, G. (ed.) Principles of Knowledge Representation. (1996) CSLI Publications

29. Laudan, L.: Damn the Consequences. The Proceedings and Addresses of the American Philosophical Association 6 (1995) 27-34
30. Lewis, D.: Causal Explanation. In his Philosophical Papers, Vol.2, (1986) Oxford University Press
31. Lipton, P.: Inference to the Best Explanation. (1991) London: Routledge
32. Lipton, P.: Tracking Track Records. Proceedings of the Aristotelian Society Suppl. Volume 74 (2000) 179-205
33. Lycan, W.: Judgement and Justification. (1988) Cambridge: Cambridge University Press
34. Lycan, W.: Explanationism, ECHO, and the Connectionist Paradigm. Behavioural and Brain Sciences 12 (1989) 480
35. Mellor, D. H.: The Warrant of Induction. (1988) Cambridge: Cambridge University Press
36. Niiniluoto, I.: Defending Abduction. Philosophy of Science 66 (Proceedings) (1999) S436-S451
37. Pereira, L. M.: Philosophical Impingement of Logic Programming. In Gabbay, D. & Woods, J. (eds) Handbook of History and Philosophy of Logic. (2001) Kluwer Academic Press
38. Pollock, J.: Contemporary Theories of Knowledge. (1986) New Jersey: Rowan & Littlefield
39. Pollock, J.: Defeasible Reasoning. Cognitive Science 11 (1987) 481-518
40. Psillos, S.: Scientific Realism: How Science Tracks Truth. (1999) London: Routledge
41. Psillos, S.: Review of Gower, B: Theories of Scientific Method. Ratio XII (1999) 310-316
42. Psillos, S.: Abduction: Between Conceptual Richness and Computational Complexity. In Flach, P. & Kakas, A. (eds.) Abduction and Induction: Essays on their Relation and Integration. (2000) Dordrecht: Kluwer Academic Publishers
43. Psillos, S.: Causation and Explanation. (forthcoming) Acumen
44. Salmon, W.: Scientific Explanation and the Causal Structure of the World. (1984) Princeton: Princeton University Press
45. Salmon, W.: Four Decades of Scientific Explanation. (1989) Minnesota University Press
46. Thagard, P.: Best Explanation: Criteria for Theory Choice. Journal of Philosophy 75 (1978) 76-92
47. Thagard, P.: Peirce on Hypothesis and Abduction. In C. S. Peirce Bicentennial International Congress. (1981) Texas University Press
48. Thagard, P.: Computational Philosophy of Science. (1988) Cambridge MA: MIT Press
49. Thagard, P.: Explanatory Coherence. Behavioural and Brain Sciences 12 (1989) 435-502
50. Thagard, P. & Shelley, C.: Abductive Reasoning: Logic, Visual Thinking and Coherence. In Dalla Chiara, M. L. (ed.) Logic and Scientific Methods. (1997) Kluwer Academic Publishers
51. van Fraassen, B.C.: The Scientific Image. (1980) Oxford: Clarendon Press
52. van Fraassen, B.C.: Laws and Symmetry. (1989) Oxford: Clarendon Press