Inference to the Best Explanation: a comparison of approaches

David H. Glass[1]

Abstract. In the form of inference known as inference to the best explanation (IBE) there are various ways to characterize what is meant by the best explanation. This paper considers a number of such characterizations including several based on confirmation measures and several based on coherence measures. The goal is to find a measure which adequately captures what is meant by 'best' and that also yields the true explanation with a high degree of probability. Computer simulations are used to show that the overlap coherence measure achieves this goal, enabling the true explanation to be identified almost as often as an approach which simply selects the most probable explanation.

1 INTRODUCTION

In many scenarios human reasoning seems to involve producing adequate explanations of the phenomena under consideration. In many artificial intelligence applications, however, reasoning and inference can be carried out without any explicit account of explanation. This naturally raises the question as to how explanations can be extracted from such applications. This is crucial if users are to trust the reliability of the inferences made. In probabilistic systems, for example, users often find it difficult to make sense of the reasoning process unless suitable explanations are available. Unfortunately the automatic generation of explanations requires an adequate account of explanation to be given and this is a notoriously difficult problem.

In addition to extracting explanations for the benefit of the user, explanations can play a more fundamental role in reasoning systems. The system could be designed to generate a range of explanations as part of the reasoning process and then go on to select the best one. This form of reasoning is known as abduction or inference to the best explanation (IBE) and has attracted a great deal of interest in the both artificial intelligence (see [13] and [7] for example) and philosophy (see [14] for example). As well as giving an account of explanation, and thus enabling the system to determine whether a proposed explanation should be accepted as a possibility, IBE requires some mechanism for comparing competing explanations. Clearly it would be useful to have a measure for the quality of an explanation and thus to provide an ordering of competing explanations.

A major difficulty for IBE is that there is no generally agreed account of explanation. Considerable effort has been expended by philosophers trying to overcome difficulties with the deductive-nomological and inductive-statistical accounts of explanation as proposed by Hempel [12]. Salmon, for example, gave an account of explanation in terms of statistical relevance (see his account in [19]). However, Salmon along with other philosophers came to realize that statistical relationships alone could not adequately account for explanation: an adequate account would require causality to be taken into account [20]. Statistical relationships may well be important, but this is because they provide evidence for underlying causal relationships.

One of the difficulties with a causal approach is that the concept of causality is just as problematic as explanation. Nevertheless, there has been considerable attention given to causality recently both in philosophy of science and in artificial intelligence. In particular, work on causality within the context of causal models has given rise to accounts of causality which are both practical and philosophically defensible as discussed in [18, 10]. This has opened up the possibility that explanation can be described in terms of such accounts [11].

In light of these points it is helpful to distinguish two components that are required for a full account of explanation:

a) an account of what constitutes an explanation, and
b) a suitable methodology for comparing competing explanations.

This paper builds on earlier work which drew on recent work on probabilistic accounts of coherence in order to meet requirement (b) by providing a measure to rank explanations [9]. In recent years a number of probabilistic accounts of coherence have been proposed and implications for the coherence theory of justification investigated ([3, 16, 17, 2, 1]). Since there has been general agreement that coherence on its own does not result in a high likelihood of truth, the focus has been on the question of whether coherence is truth conducive so that more coherence gives rise to a higher probability of truth. Olsson [17] presents an impossibility theorem to the effect that there is no truth conducive coherence measure. Bovens and Hartmann [1] also present an impossibility result, but argue that its impact on coherentism can be circumvented by adopting a partial ordering of information sets on the basis of coherence, i.e. in some cases one set can be identified as more coherent than another while in other cases no such comparison is possible.

In [9] coherence was considered as a relation between an hypothesis and the evidence for it. The motivation was to find a measure of coherence which matches our intuitive understanding of the concept and to investigate how such a conception might relate to explanation. A connection between the notions of explanation and coherence was established by noting that a condition for a satisfactory account of the relation "... better explanation than ..." turned out to be essentially the same as a plausible condition for the relation "...more coherent than...". After identifying a suitable measure of coherence, several scenarios were presented to illustrate some advantages of this approach over other accounts of 'best explanation'.

This paper expands on this earlier research in two ways. First, in addition to the measures previously used to quantify 'best explana-

[1] School of Computing and Mathematics, University of Ulster, Newtownabbey, Co. Antrim, BT37 0QB, United Kingdom, email: dh.glass@ulster.ac.uk

tion' several other coherence and confirmation measures are considered as possible alternatives. It should be noted that not all of these measures were proposed as measures for ranking explanations, but they do appear to be plausible candidates nevertheless. Second, a computational approach is adopted so that instead of comparing measures on particular scenarios with specifically selected probabilities, they can be compared over numerous scenarios where the probabilities are selected randomly. This gives a picture as to how well the different measures function on average in terms of identifying the actual explanation that is responsible for the evidence.

The structure of the paper is as follows. Section 2 presents a number of possible ways to compare competing explanations. These approaches are then tested using computer simulations in section 3. Section 4 discusses the relevance of this work for the feasibility of IBE as a mode of inference and section 5 presents conclusions.

2 WHAT IS THE BEST EXPLANATION?

In attempting to provide a methodology for comparing competing explanations it is worth noting Hempel's distinction between potential and actual explanations [12]. An actual explanation is one which, as a matter of fact, explains the explanandum in question. A potential explanation is one which, if true, would be an actual explanation. This section considers different approaches for comparing potential explanations. In this context the goal of IBE can be understood as selecting the actual explanation from the potential explanations. Different forms of IBE arise from which of the approaches is used to select the best explanation from the potential explanations.

It is also worth noting another distinction that has been emphasized by Lipton [14] who distinguished between the loveliest explanation and likeliest explanation. To quote Lipton, "We want a model of inductive inference to describe what principles we use to judge one inference more likely than another, so to say that we infer the likeliest explanation is not helpful ([14], p. 60). It seems that the goal for defenders of IBE is to give an account of 'best explanation' in terms of loveliness and show that a feature of such an explanation will be its likeliness, i.e. high posterior probability.

2.1 Approaches based directly on Bayes' theorem

If the goal of IBE is to provide an account of 'best explanation' that will typically have a high posterior probability, then Bayes' theorem provides an obvious starting point. Suppose that there are n hypotheses H_i where $i = 1, \ldots, n$, then the posterior probability of each hypothesis given evidence E is given by,

$$Pr(H_i|E) = \frac{Pr(E|H_i)}{Pr(E)} \times Pr(H_i), \tag{1}$$

where all the probabilities are assumed to be conditioned on appropriate background evidence k which has been suppressed in the notation.

As pointed out in [9], the most probable explanation (MPE) approach simply takes the best explanation to be the one with the highest posterior probability. This means that hypothesis H_1 is better than H_2 if and only if

$$Pr(H_1|E) > Pr(H_2|E). \tag{2}$$

Of course, adopting this approach guarantees that the best explanation will be the one that is most probable given the evidence, but it makes IBE trivial since this success has been achieved simply by defining 'best' as 'most probable given the evidence'. In Lipton's terminology the 'loveliest' explanation has simply been defined as the 'likeliest' explanation.

A second approach discussed in [9] is the maximum likelihood (ML) approach. Taking the first term on the RHS of equation (1) hypothesis H_1 is defined to be better than H_2 if and only if

$$Pr(E|H_1) > Pr(E|H_2). \tag{3}$$

This has certainly some merit to it since good explanations often do make the occurrence of the relevant evidence highly probable. In fact, ideally an hypothesis will deductively entail that the evidence will occur. The problem, however, is that there is no good reason for thinking that an hypothesis with a high likelihood will also have a high posterior probability unless it also has a high prior probability. Thus, despite its merits, we might expect that in many cases IBE, understood as inference to the hypothesis with the maximum likelihood, will not be a good approach for finding true (or highly probable) hypotheses.

A middle way is provided in [4] who define hypothesis H_1 as better than H_2 if and only if

$$Pr(E|H_1) > Pr(E|H_2) \quad \text{and} \quad Pr(H_1) > Pr(H_2). \tag{4}$$

This approach is referred to in [9] as a conservative Bayesian (CB) approach. A problem with CB is also pointed out in [9] since there are many cases in which the ML and MPE approaches agree as to which of two hypotheses H_1 and H_2 is best and yet CB fails to order them. For example, the ML and MPE approaches will agree in all cases where the priors of the competing explanations are equal and in many cases where the explanation with the greater likelihood has a lower prior, yet in such cases CB does not provide an ordering.

Before going on to look at other approaches, is is worth pausing to ask whether there is a preferred Bayesian account of 'best explanation'. According to Bayesianism, the rational agent updates her degrees of belief according to conditionalization. For example, if she has a prior probability for an hypothesis H_i of $Pr(H_i)$, then conditionalization requires that after taking evidence E into account her probability should be updated via Bayes' theorem as defined in equation (1) so that her posterior probability for H_i is $Pr(H_i|E)$. If she is required to infer one hypothesis, then it seems that this should be the one that is most probable. If so, it might seem that the MPE approach as expressed in (2) is the preferred Bayesian account of 'best explanation', but this is not necessarily the case. The reason for this is that the Bayesian might wish to maintain that she is interested in the most probable hypothesis but that this need not be the one which is the best explanation; it is probability that is important, not explanation. Furthermore, there is no requirement for the Bayesian to infer one of the hypotheses and so even if the MPE approach is adopted this still does not mean that Bayesianism is a form of IBE.

Indeed, van Fraassen [22] has gone further and argued that Bayesianism and IBE are conflicting approaches. Others [15, 14] have responded by arguing that IBE need not involve any departure from Bayesian probabilities and that explanatory considerations may come into play in implementing Bayesian reasoning. A difficulty with this approach is that it seems to require that the 'best explanation' be *defined* as the most probable explanation. But as noted above this makes IBE trivial since in this case the 'best explanation' is guaranteed to be the most probable explanation by definition. The goal of IBE is to give an account of 'best explanation' that is conceptually distinct from 'most probable explanation' and yet show that the best explanation will often be the one that is most probable. The aim in this paper is to see whether there is an account of 'best explanation' that will achieve this.

2.2 Approaches based on confirmation theory

A confirmation measure of the degree to which a piece of evidence E confirms an hypothesis H, denoted $c(E, H)$, is a measure which satisfies

(i) $c(H, E) > 0 \ \ iff \ \ Pr(H|E) > Pr(H)$
(ii) $c(H, E) = 0 \ \ iff \ \ Pr(H|E) = Pr(H)$
(iii) $c(H, E) < 0 \ \ iff \ \ Pr(H|E) < Pr(H)$

where Pr is a probability function. Another way of putting this is to say that E confirms (disconfirms) H if and only if there is a positive (negative) probabilistic dependence between E and H. It is important to emphasize that confirmation in the sense used here relates to the impact of the evidence on the probability of the hypothesis rather than simply being the posterior probability of the hypothesis given the evidence. This means that the degree to which E confirms H is a measure of how much *more* probable the evidence E makes the hypothesis H.

A large number of confirmation measures have been proposed in the literature (see for example [5]). Here, only three are considered. First, the ratio measure is given by $r(H, E) = log\left[Pr(H|E)/Pr(H)\right]$, which can be used to rank explanations such that H_1 is defined to be better than H_2 if and only if

$$log\left[\frac{Pr(H_1|E)}{Pr(H_1)}\right] > log\left[\frac{Pr(H_2|E)}{Pr(H_2)}\right]. \tag{5}$$

However, since $Pr(H_i|E)/Pr(H_i) = Pr(E|H_i)/Pr(E)$, it turns out that this results in an identical ordering of explanations as the ML approach considered in the last section. For this reason, the ratio measure will not be considered further here.

An alternative confirmation measure is the difference measure which is given by $d(H, E) = Pr(H|E) - Pr(H)$ and so enables H_1 to be defined as better than H_2 if and only if,

$$Pr(H_1|E) - Pr(H_1) > Pr(H_2|E) - Pr(H_2). \tag{6}$$

The final confirmation measure considered here is the likelihood ratio given by $l(H, E) = log\left[Pr(E|H)/Pr(E|\neg H)\right]$ which enables H_1 to be defined as better than H_2 if and only if,

$$log\left[\frac{Pr(E|H_1)}{Pr(E|\neg H_1)}\right] > log\left[\frac{Pr(E|H_2)}{Pr(E|\neg H_2)}\right]. \tag{7}$$

It turns out that when there are only two mutually exclusive and exhaustive hypotheses H_1 and $\neg H_1$ all confirmation measures will agree as to which is the best explanation. This is because if E confirms H_1 then it disconfirms H_2 and so the degree of confirmation E provides for H ($\neg H$) will be positive (negative) for all confirmation measures. This does not apply when more than two hypotheses are being considered.

2.3 Approaches based on coherence

In [9] a case was made for using a coherence measure known as the overlap measure proposed in [16, 8] to rank explanations. For an hypothesis H and evidence E the measure in question is given by

$$C_O(H, E) = \frac{Pr(H \wedge E)}{Pr(H \vee E)} \tag{8}$$

whenever $Pr(H \vee E) \neq 0$ and so the suggestion was to define H_1 as better than H_2 if and only if

$$\frac{Pr(H_1 \wedge E)}{Pr(H_1 \vee E)} > \frac{Pr(H_2 \wedge E)}{Pr(H_2 \vee E)}. \tag{9}$$

Another coherence measure proposed in [6] is given by first defining a confirmation measure as

$$F(H, E) = \frac{Pr(E|H) - P(E|\neg H)}{P(E|H) + P(E|\neg H)} \tag{10}$$

and then using this to define a coherence measure, which we shall refer to as the Fitelson measure, as $C_F(H, E) = \{F(H, E) + F(E, H)\}/2$. Using this measure H_1 can be defined to be better than H_2 if and only if

$$C_F(H_1, E) > C_F(H_2, E). \tag{11}$$

It is worth noting that this measure is a confirmation measure as well as a coherence measure.

It turns out that a further coherence measure, the Shogenji measure, proposed in [21], which is defined for H and E as $\frac{Pr(H,E)}{Pr(H) \cdot Pr(E)}$, provides an equivalent ordering to the ML approach discussed earlier. For this reason it will not be considered further.

3 A COMPARISON OF APPROACHES

The goal in this section is to compare the following approaches to ranking explanations:

(MPE) the most probable explanation approach as expressed in (2),
(ML) the maximum likelihood approach as expressed in (3),
(CB) the conservative Bayesian approach as expressed as in (4),
(DIFF) the approach based on the difference confirmation measure as expressed as in (6),
(LR) the approach based on the likelihood ratio confirmation measure as expressed as in (7),
(OCM) the approach based on the overlap coherence measure as expressed as in (9),
(FCM) the approach based on the Fitelson coherence measure as expressed as in (11).

It is not immediately obvious, however, how to compare the different approaches. Since the goal in IBE is to infer the actual explanation then it would seem that MPE would be the most appropriate approach since it will yield the explanation with the highest probability given the evidence. As has already been pointed out, however, this would make IBE trivial and furthermore MPE does not really seem to capture the notion of 'best explanation' adequately. This can be seen from the fact that an explanation can be the most probable one simply because it has a high prior probability and even though it gives a low probability for the evidence, i.e. has a low likelihood.

Another way to address the issue is to look at what kinds of features make an explanation a good one and then see which approach best takes these into account. This is essentially to consider which measure best captures various explanatory virtues. In [9] a case was made that OCM was to be preferred in this respect to MPE, ML and CB and some scenarios were used to motivate this preference. This way of proceeding is somewhat subjective, however, and undoubtedly a case could be made for some of the other approaches on the list since all of them have their merits.

3.1 Methodology

Here the method for comparing the approaches is rather different. Recall that IBE involves first of all consider the explanatory merits of the potential explanations and then inferring the best one as being true or probably true. Hence, the suitability of IBE as an inductive methodology will depend on how often it enables us to identify the actual (or true) explanation. It will, of course, be impossible to do better in this latter respect than MPE, but as we have seen MPE is inadequate as an account of 'best explanation'. The goal is then to see which of the other approaches best approximates MPE, i.e. which of the other approaches yields the actual explanation most often.

In order to do this, computer experiments have been carried out to see how the various approaches perform. The idea is to take a given number of mutually exclusive and exhaustive hypotheses and randomly assign prior probabilities (adding to one) to them. Random values of the likelihoods $Pr(E|H_i)$ are then attributed to the hypotheses. One of the hypotheses is then selected randomly according to the prior probability distribution and designated as the actual hypothesis. Whether E occurs is then decided randomly based on the likelihood ratio for the actual hypothesis. If E occurs the hypothesis which is the best explanation according to each of the approaches above is identified and if it corresponds to the actual hypothesis this is considered a success, otherwise it is a fail. The entire exercise is then repeated to get an average picture of the performance of each approach. It is also repeated for different numbers of hypotheses.

The procedure is summarised in the algorithm below.

initialize number of hypotheses N and number of repetitions R
for $i = 1$ to R **do**
  set count$_E$ and count$_j$ for each approach to zero
  for $k = 1$ to N **do**
    set the prior probability of hypothesis H_k randomly (ensuring they sum to one)
    set the likelihood of hypothesis H_k randomly
  end for
  select one hypothesis based on the prior probability distribution and designate it the actual explanation H_A
  select whether E or $\neg E$ occurs based on the likelihood of H_A
  if E occurs **then**
    increment count$_E$
    for each approach (MPE) to (FCM) denoted j **do**
      select the hypothesis that is the best explanation H_B
      if H_B = H_A **then**
        increment count$_j$
      end if
    end for
  end if
end for
for each approach (MPE) to (FCM) **do**
  print count$_j$ / count$_E$
end for

3.2 Results

Results were obtained for values of N, the number of competing hypotheses, ranging from 2 to 10. In each case 100,000 repetitions were carried out to ensure that the results were accurate. The results are displayed in Figure 1. The accuracy is the number of cases in which a given measure identifies the actual explanation expressed as a percentage of cases in which the evidence E occurs.

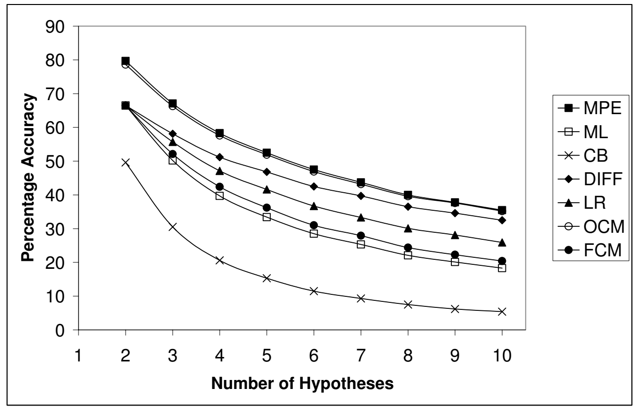

Figure 1. Accuracy plotted as a function of the number of competing hypotheses for each of the different approaches.

It is clear from the results that the percentage accuracy decreases as a function of the number of hypotheses for all of the approaches considered. This, of course, is exactly as expected since as the number of hypotheses increases there are more ways to identify an hypothesis that is not the actual explanation. Also, the MPE approach performs best for all values of the number of hypotheses. This again is as expected since it effectively sets the standard against which the other approaches are to be compared.

Clearly, the CB approach performs worst. In fact, it performs worse than simply selecting an hypothesis at random. For example, when $N = 10$ the accuracy is just 5.4%, whereas 10% could be achieved by random selection. On the other hand, it must be noted that the CB approach is the only one considered here that does not provide a complete ordering of hypotheses and so it also has a much lower false positive rate than other approaches. Nevertheless, it seems clear that the conservative Bayesian approach is too conservative.

The results also illustrate the point, noted in section 2, that all the confirmation measures ML (which is equivalent to the ratio measure), DIFF, LR, and FCM (which is also a coherence measure) yield identical results at $N = 2$, but diverge for $N > 2$. For these higher values of N, DIFF has the best performance, followed by LR, FCM and ML respectively. As N becomes large the results for DIFF get much closer to the MPE results. This seems reasonable since the confirmation measure used in the DIFF approach is the difference measure given by $d(H,E) = Pr(H|E) - Pr(H)$ and so the prior probability $Pr(H)$ becomes less important as the number of hypotheses increases.

Recall that three of the approaches correspond to coherence measures, OCM, FCM and ML (which corresponds to the Shogenji measure, see section 2). Of these, FCM performs slightly better than ML, but OCM outperforms all the other measures. In fact, OCM tracks MPE so closely that it is not possible to distinguish them in the figure. For all values of N the OCM result is within a couple of percent of the MPE result. This means that at $N = 2$, for example, MPE and OCM yield the hypothesis which is the actual explanation almost 80% of the time while the confirmation measures only get the correct result two-thirds of the time and CB half of the time.

Table 1 summarises the performances of the different approaches averaged over values of N in terms of how well they compare with MPE. The OCM approach does remarkably well, identifying the actual explanation 99% as often as MPE, with DIFF coming in second place with a score of 89%. CB is a long way behind the other approaches with a score of just 30%.

Table 1. The ratio of the accuracy of each approach to the MPE approach averaged over values of N from 2 to 10.

Approach	Average percentage of MPE result
OCM	99
DIFF	89
LR	78
FCM	68
ML	63
CB	30

4 DISCUSSION

The results indicate that the OCM approach using the overlap coherence measure proposed in [16, 8] performs much better than the other measures when compared against the benchmark of the MPE approach. In effect, the OCM approach is almost as good at identifying the actual explanation as the MPE approach. This seems to suggest that OCM does provide a good way of comparing explanations or alternatively a good way of quantifying what is meant by the 'best explanation', but how does this relate to the viability of IBE as an approach to inductive reasoning?

In [9] it was argued that the OCM approach provided a good way of making IBE precise and so IBE can be understood as *inference to the most coherent explanation*, where coherence refers to the coherence between the explanation and the evidence and the coherence measure used is the overlap measure. There it was claimed that it had a number of advantages over MPE, ML and CB and that it provides a good way of linking the goodness of an explanation with its probability of being true without simply defining 'best' as 'most probable'. It was pointed out that any approach which does not define 'best' as 'most probable' will inevitably conflict with MPE in some cases, nevertheless if IBE is to be a viable form of inductive reasoning it should tend to yield explanations that are highly probable. It was claimed that OCM was such an approach, but no experimental evidence was presented.

This paper presents evidence from computer simulations to back up this claim and actually supports a much stronger claim: OCM not only tends to yield highly probable explanations, but it yields the actual explanation almost as frequently as the MPE approach which simply selects the most probable explanation. It is difficult to see how any alternative account of IBE could do better.

This still leaves a question concerning the importance of IBE as an inductive reasoning method distinct to Bayesianism. Two approaches were discussed in [9]. First, perhaps IBE is intended as a descriptive account of how humans actually reason, whereas Bayesianism is the normatively correct way that humans should reason. If so, the rationality of IBE depends on how well it tracks Bayesianism, i.e. how frequently it yields the most probable explanation. As we have seen if IBE is understood as inference to the most coherent explanation, it does remarkably well. Alternatively, perhaps IBE is intended to be a rival to Bayesianism. After all, in some cases the goal of inference is not to find the hypothesis that is most probable given the evidence. As Lipton points out, "... high probability is not the only aim of inference. Scientists also have a preference for theories with great content, even though that is in tension with high probability, since the more one says the more likely it is that what one says is false" ([14], p. 116). Scientists are typically interested in theories which are as precise as possible and testable rather than being vague and compatible with both a piece of evidence and its negation. The account of IBE presented here seems appropriate in this context since the cases where it will diverge from Bayesianism are those when the posterior probability is high because of a high prior and despite a low likelihood.

5 CONCLUSIONS

Various approaches to quantifying the goodness of explanations so that they can be compared and ranked have been considered. These include several simple approaches arising directly from Bayes' theorem, several approaches based on confirmation measures and several approaches based on coherence measures. Results have been presented to show how well each of these approaches performs in terms of identifying the actual explanation. In one sense the MPE approach, which simply identifies 'best explanation' with the explanation that is 'most probable given the evidence', is ideal from an inductive point of view, but it makes IBE trivial and does not provide an adequate account of 'best explanation'. Instead MPE can be seen as the benchmark against which the performance of other approaches should be assessed.

The results show that of the other measures the OCM approach, which uses the overlap coherence measure to identify the 'best explanation', identifies the actual explanation almost as often as the MPE approach. Yet this account seems much more plausible as an account of 'best explanation'. Thus, the research presented here goes some way to vindicating IBE as a form of reasoning provided it is understood as inference to the most coherent explanation where coherence is measured using the overlap measure.

REFERENCES

[1] L. Bovens and S. Hartmann, *Bayesian Epistemology*, Oxford University Press, Oxford, 2003.

[2] L. Bovens and S. Hartmann, 'Solving the riddle of coherence', *Mind*, **112**, 601–633, (2003).

[3] L. Bovens and E. Olsson, 'Coherence, reliability and Bayesian networks', *Mind*, **109**, 685–719, (2000).

[4] U. Chajewska and J. Y. Halpern, 'Defining explanation in probabilistic systems', in *Proceedings of the 13th Conference on Uncertainty in AI*, pp. 62–71, (1997).

[5] B. Fitelson, 'The plurality of Bayesian measures of confirmation and the problem of measure sensitivity', *Philosophy of Science*, **66**, S362–S378, (1999).

[6] B. Fitelson, 'A probabilistic theory of coherence', *Analysis*, **63**, 194–199, (2003).

[7] P. A. Flach and A. C. Kakas, *Abduction and Induction: Essays on their relation and integration*, Kluwer Academic Publishers, 2000.

[8] D. H. Glass, 'Coherence, explanation and Bayesian networks. In Proceedings of the 13th Irish Conference on AI and cognitive science', *LNAI*, **2464**, 177–182, (2002).

[9] D. H. Glass, 'Coherence measures and inference to the best explanation', *Synthese*, **157**, 275–296, (2007).

[10] J. Y. Halpern and J. Pearl, 'Causes and explanations: a structural-model approach-part I: Causes', in *Proceedings of the 17th Conference on Uncertainty in AI*, pp. 194–202, (2001).

[11] J. Y. Halpern and J. Pearl, 'Causes and explanations: a structural-model approach-part II: Explanations', in *Proceedings of the 17th International Joint Conference on AI*, pp. 194–202, (2001).

[12] C. G. Hempel, *Aspects of Scientific Explanation*, Free Press, 1965.

[13] J. R. Josephson and S. G. Josephson, *Abductive Inference: Computation, Philosophy and Technology*, Cambridge University Press, Cambridge, 1994.

[14] P. Lipton, *Inference to the Best Explanation*, Routledge, London, 2nd edn., 2004.

[15] S. Okasha, ‘Van Fraassen’s critique of inference to the best explanation’, *Studies in the History and Philosophy of Modern Science*, **31**, 691–710, (2000).

[16] E. J. Olsson, ‘What is the problem of coherence and truth?’, *Journal of Philosophy*, **99**, 246–272, (2002).

[17] E. J. Olsson, *Against Coherence*, Oxford University Press, Oxford, 2005.

[18] J. Pearl, *Causality: Models, Reasoning and Inference*, Cambridge University Press, Cambridge, 2000.

[19] W. C. Salmon, *Four Decades of Scientific Explanation*, University of Minnesota Press, 1989.

[20] W. C. Salmon, *Causality and Explanation*, Oxford University Press, Oxford, 1998.

[21] T. Shogenji, ‘Is coherence truth-conducive?’, *Analysis*, **59**, 338–345, (1999).

[22] B. C. van Fraassen, *Laws and Symmetry*, Clarendon Press, Oxford, 1989.