JID:IJA AID:8254 /FLA [m3G; v1.243; Prn:17/09/2018; 11:54] P.1 (1-10)
International Journal of Approximate Reasoning ••• (••••) •••–•••

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

An evaluation of probabilistic approaches to inference to the best explanation ☆

David H. Glass

School of Computing, Ulster University, Shore Road, Newtownabbey, Co. Antrim, BT37 0QB, UK

A R T I C L E I N F O

Article history:
Received 1 March 2018
Received in revised form 13 September 2018
Accepted 13 September 2018
Available online xxxx

Keywords:
Explanation
Abduction
Bayesian
Computer simulation

A B S T R A C T

This paper presents results of computer simulations for a number of different probabilistic versions of inference to the best explanation (IBE), which are distinguished by the probabilistic measures used to identify the best explanation. Simulation results are presented which include cases involving ignorance of a catch-all hypothesis, uncertainty regarding the prior probability distribution over the remaining hypotheses, initial elimination of implausible hypotheses, and variations in the number of pieces of evidence taken into consideration. The results show that at least some versions of IBE perform very well in a wide range of cases. In particular, the results for all approaches remain very similar (or improve in some cases) when just the two hypotheses with the highest prior probabilities are retained and the rest are eliminated from consideration.

1. Introduction

Inference to the best explanation (IBE) or abduction is an ampliative mode of reasoning which is often defended as central to scientific reasoning, but also seems to capture aspects of evidential reasoning as it occurs more generally, including in everyday life. IBE proceeds by considering a number of plausible candidate hypotheses in a given evidential context and then comparing these hypotheses in order to make an inference to the one that best explains the relevant evidence. It has been discussed widely in both the philosophy of science and computer science literature [1–4].

There has been a lot of interest in the relationship between IBE and probability, particularly in debates about the compatibility or otherwise between IBE and Bayesian inference [5–8] and on IBE or abduction in the context of Bayesian networks [9–14]. Here the focus is on using probabilistic approaches to determine which one of a collection of competing hypotheses provides the best explanation of the evidence. These approaches enable two key questions about IBE to be addressed. First, they provide ways of making IBE precise by specifying what is meant by 'best'. Various probabilistic measures have been proposed to this end in the literature [15–19]. Second, they make it possible to investigate whether selecting the best explanation is a good strategy for inferring truth. In this context, some studies have been carried out using computer simulations to evaluate how well various measures perform in hypothesis selection tasks and provide more general defences of versions of IBE based on probabilistic measures [18,20,21]. Hence, although the paper does not attempt to address all the philosophical issues surrounding IBE (see for example [2,3]), the results obtained are very relevant to those debates.

☆ This paper is part of the Virtual special issue on Defeasible and Ampliative Reasoning, Edited by Ivan Varzinczak, Richard Booth, Giovanni Casini, Szymon Klarman and Gilles Richard.

E-mail address: dh.glass@ulster.ac.uk.

https://doi.org/10.1016/j.ijar.2018.09.004
0888-613X/© 2018 Published by Elsevier Inc.

An interesting aspect of previous work is that in cases where there is uncertainty in the prior probability distribution over the hypotheses, inference based on a coherence measure for ranking hypotheses outperformed the approach that simply selects the most probable hypothesis in light of the evidence [18]. The current paper builds on this work by providing a more adequate and realistic account of how probability can be used in IBE and a more systematic evaluation of how IBE so construed performs as a mode of reasoning when ignorance of a catch-all hypothesis, uncertainty regarding the prior distribution, initial elimination of implausible hypotheses, and variations in the number of pieces of evidence available are taken into account.

These are important considerations for IBE in general. For example, whether in the context of scientific reasoning or in everyday, commonsense reasoning, it would be unrealistic to assume complete knowledge of all possible hypotheses. As such, inclusion of a catch-all hypothesis allows for IBE to be modelled and evaluated in a more realistic way. Similarly, even among the known hypotheses, it would be unreasonable to expect that they should all be evaluated in detail since the number of hypotheses could be large and some hypotheses might be considered very implausible based on background knowledge. Hence, modelling this practical aspect of IBE enables us to see how it affects its performance.

Rather than exploring IBE in the context of particular applications, whether in science or everyday life, the goal of the paper is to evaluate IBE as a general mode of inference. More specifically, the goal is to investigate the performance of various probabilistic models of IBE. Thus, the focus is of a conceptual nature, but by incorporating more realistic aspects of IBE into the simulations the results may also have implications for how IBE could be justified and used in practical applications.

2. Measures for comparing hypotheses

A number of measures have been proposed in the literature to quantify how well a hypothesis h explains evidence e. These include the following measure of explanatory power proposed by Schupbach and Sprenger [16]:

$$\mathcal{E}_{SS}(e,h) = \frac{P(h|e) - P(h|{\sim}e)}{P(h|e) + P(h|{\sim}e)}, \tag{1}$$

an alternative measure of explanatory power proposed by Crupi and Tentori [19]:

$$\mathcal{E}_{CT}(e,h) = \begin{cases} \dfrac{P(e|h) - P(e)}{1 - P(e)} & \text{if } P(e|h) \geq P(e) \\ \dfrac{P(e|h) - P(e)}{P(e)} & \text{if } P(e|h) < P(e), \end{cases} \tag{2}$$

another measure that has been discussed by Good [22] and McGrew [23]:

$$\mathcal{E}_{GM}(e,h) = \ln\left[\frac{P(e|h)}{P(e)}\right], \tag{3}$$

and the overlap coherence measure (OCM) used to rank explanations by Glass [17,18]:

$$\mathcal{E}_{OCM}(e,h) = \frac{P(h \wedge e)}{P(h \vee e)}. \tag{4}$$

Criticisms of some of these measures were presented by Glymour [24], while a response has been given by Glass [21]. In order to respond to a criticism based on the fact that advantages of the $\mathcal{E}_{OCM}$ diminished as the sample size (i.e. the number of samples of evidence) increased, the following alternative product coherence measure (PCM) was proposed and shown to retain the advantages with increasing sample size:

$$\mathcal{E}_{PCM}(e,h) = P(e|h) \times P(h|e). \tag{5}$$

The strategy used in this paper and discussed in detail in section 3 is to consider a set of hypotheses $\{h_i\}$ for evidence e and select the hypothesis which gives the maximum value of a particular measure, $\mathcal{E}$. It would be possible to use all the measures defined above, but it turns out that $\mathcal{E}_{SS}$, $\mathcal{E}_{CT}$ and $\mathcal{E}_{GM}$ all give the same ranking of hypotheses, giving the same result as selecting the hypothesis with the maximum likelihood, $P(e|h_i)$.[1]

Another measure that will be considered is the posterior probability of the hypotheses in light of the evidence, $P(h_i|e)$. The hypothesis that maximizes posterior probability is often referred to in the artificial intelligence literature as the most probable explanation (MPE). Arguably, this is a poor definition of 'best explanation' [15,17], it nevertheless provides a standard against which to compare the various explanatory measures.

[1] This is straightforward to show for $\mathcal{E}_{CT}$ and $\mathcal{E}_{GM}$. From the definition of $\mathcal{E}_{SS}$ it is easy to show that $\mathcal{E}_{SS}(e,h_1) > \mathcal{E}_{SS}(e,h_2)$ if and only if $P(h_1|e)P(h_2|{\sim}e) > P(h_1|{\sim}e)P(h_2|e)$. Using Bayes' theorem to replace each term and then rearranging, we can see that this expression is true if and only if $P(e|h_1) > P(e|h_2)$, provided that P is a regular probability function. See also theorem 1 in [19].

Please cite this article in press as: D.H. Glass, An evaluation of probabilistic approaches to inference to the best explanation, Int. J. Approx. Reason. (2018), https://doi.org/10.1016/j.ijar.2018.09.004