Competing hypotheses and abductive inference

David H. Glass[1] ✆

Abstract

This paper explores the nature of competition between hypotheses and the effect of failing to model this relationship correctly when performing abductive inference. In terms of the nature of competition, the importance of the interplay between direct and indirect pathways, where the latter depends on the evidence under consideration, is investigated. Experimental results show that models which treat hypotheses as mutually exclusive or independent perform well in an abduction problem that requires identifying the most probable hypothesis, provided there is at least some positive degree of competition between the hypotheses. However, even in such cases a significant limitation of these models is their inability to identify a second hypothesis that may well also be true.

Keywords Abduction · Inference to the best explanation · Hypothesis competition · Computer simulation

1 Introduction

Abduction or abductive inference is a mode of explanatory inference that has been considered at great length in both artificial intelligence and philosophy of science [1–4]. While the term 'abduction' is sometimes used to refer only to how explanatory hypotheses are generated, it is also used more generally to include the the evaluation of these hypotheses, which is the focus in this paper. In this case it is often referred to as 'inference to the best explanation' [4], where the general idea is to compare how well various competing hypotheses explain the available evidence and then make an inference to the hypothesis that does best. One of its main attractions is that it seems to capture aspects of scientific inference as well as reasoning in everyday life. The relationship between abduction and probability has been a frequent theme in the literature, with abduction often being spelt out in probabilistic terms. Questions here concern the compatibility or otherwise between abduction and Bayesian inference, which is to be preferred if they are incompatible and how 'best' is to be quantified [3, 5–9].

✉ David H. Glass
    dh.glass@ulster.ac.uk

[1]   School of Computing, Ulster University, Newtownabbey, Co. Antrim, BT37 0QB, UK

Logic-based approaches to abduction have received a lot of attention in the artificial intelligence literature, with a lot of focus on abductive logic programming [10, 11]. From this perspective, abduction involves finding an explanation which along with a background logical theory entails an observation. A challenge in such approaches is how to handle the large numbers of explanations that can exist. While various techniques for achieving this have been proposed, of particular interest here is the use of probability such as *probabilistic horn abduction* [12], which specifies probabilities over hypotheses and then uses explanations to compute posterior probabilities of interest. This work also showed that the knowledge in a discrete Bayesian network can be represented within this framework and vice versa. This approach was extended into the *independent choice logic*, which provides a richer logic and allows for multi-agent decision-making under uncertainty [13, 14].

There has also been a lot of research on abductive inference in the context of Bayesian networks more generally [15]. Given a set of observations or evidence represented as an instantiation of a subset of nodes (evidence nodes) within the network, the general idea is to identify the best explanation in terms of a configuration of other nodes (explanatory nodes) in the network. *Most Probable Explanation* (MPE) is a version of abductive inference that includes all the non-evidence nodes in the set of explanatory nodes [16]. Clearly, the explanatory hypotheses under consideration are mutually exclusive since they differ in the configuration of one or more of the explanatory nodes. A problem with this approach is that it often produces explanations that are overspecified since they include many variables that are irrelevant to the evidence and so an alternative approach is *Maximum a Posteriori* (MAP) which instead includes only a subset of non-evidence nodes in the set of explanatory nodes [17]. A variety of approaches as well as approximation techniques have been proposed to address this issue [17–22]. Irrespective of what selection of explanatory nodes is made, typically these approaches still involve comparing mutually exclusive explanatory hypotheses. Addressing concerns not only about overspecification, but also how to quantify 'best explanation' in probabilistic terms [7, 23], an approach based not on posterior probability but on generalized Bayes factors, which allows the set of explanatory variables to be varied and does not restrict the comparison to mutually exclusive hypotheses, is very relevant here [24].

The current paper focusses on the nature of competition between the hypotheses and how this relates to inference. In many presentations of abduction the hypotheses are assumed to be mutually exclusive, but while it is certainly true that mutually exclusive hypotheses compete, it seems just as clear that hypotheses can compete even if they are compatible. For example, in medical diagnosis suppose that two conditions are being considered to explain a patient's symptoms. Typically these conditions will be compatible and could both contribute to the symptoms, but in many cases it will be reasonable to treat them as competing nevertheless since it would be unnecessary to infer two conditions if one accounts for all the evidence. For an insightful paper that raises the problem concerning the nature of competition, see Harman [25]. This topic is closely related to that of explaining away and intercausal reasoning which have been explored in the artificial intelligence literature [26, 27]. The relationship between explanatory hypotheses could also be relevant to evaluating the goodness of an explanation, what Lipton calls its 'loveliness' [3]. Returning to the case of diagnosis, sometimes two (or more) conditions will provide a better explanation than just one condition, which suggests 'best explanation' should not be quantified in terms of posterior probability. Also, it could be the case that there is merit to preferring hypotheses that are compatible with other hypotheses that are less informative and/or have lower probability. These issues will not be pursued here, however.

In the next section of the paper, several features of a particular account of competition are explored, particularly insofar as it differs from negative dependence between hypotheses, which might initially seem to provide a promising way to generalize competition beyond mutual exclusion. After exploring the nature of competition in Section 2, the rest of the paper investigates a further question. If the relationship between hypotheses is modelled incorrectly, how does that affect the results of abduction? This question is addressed by means of experiments which involve generating a probability model (designated the correct probability model) that has a specified degree of competition between hypotheses according to a recently proposed measure of competition [28]. Three modifications of this model, two that treat the hypotheses as mutually exclusive and a third that treats them as independent, are then used for abductive inference and the results compared against those obtained with the correct probability model. These experiments were designed to identify the limitations of treating the hypotheses as mutually exclusive, for example, when in actual fact they may just be competing to some extent or perhaps not competing at all.

2 Competing hypotheses

2.1 Preliminaries

Let $\mathcal{B}$ be a Boolean algebra closed under the usual logical operations (&, $\vee$, $\neg$) and P be a probability measure defined over $\mathcal{B}$. Let $E \in \mathcal{B}$ represent the evidence and $H, H' \in \mathcal{B}$ hypotheses for E. The catchall hypotheses, denoted H^c, is defined as $H^c = \neg H \& \neg H'$. Probabilities are taken to represent degrees of belief of an agent relevant to background knowledge, which is omitted in the notation for simplicity. Unless otherwise state, $P(H)$, $P(H')$, $P(E)$, $P(H|E)$, $P(H'|E)$, $P(H \& \neg H')$ and $P(\neg H \& H')$ are assumed to be non-extreme (neither zero nor one) representing the idea that the hypotheses and evidence are not considered by the agent to be either certain or impossible and that neither hypothesis is considered to entail the other.

2.2 An account of competition

Suppose a detective has two main suspects in a murder inquiry, Smith and Jones. Treating the suspects as two competing hypotheses and reasoning abductively, the detective tries to determine which hypothesis best explains all the relevant evidence. We can represent the hypotheses as follows:

H: Smith committed the murder
H': Jones committed the murder

Note that in general the hypotheses need not be assumed to be mutually exclusive since both Smith and Jones could have colluded in committing the murder and hence it cannot be assumed that $P(H \& H') = 0$. However, if the two hypotheses are known to be mutually exclusive so that—if Smith and Jones could not have colluded—then clearly they are competing hypotheses. In reality, it might be difficult to establish mutual exclusion, but in many cases it would still be reasonable to treat them as competing hypotheses. After all, if either suspect's guilt on its own would explain all the relevant evidence, it would seem like a violation of Ockham's razor to infer the guilt of both parties. However, in other cases where, for example, the guilt of both parties is needed to explain the evidence, it may well

be appropriate to make such an inference even though $P(H\&H'|E)$ can be no greater than $P(H|E)$ or $P(H'|E)$.[1]

Since hypotheses could compete without being mutually exclusive, it might be tempting to propose a more adequate account of competition in terms of negative dependence between the hypotheses with respect to background knowledge only. For example, suppose Smith and Jones are known to be criminal rivals and so very unlikely to have colluded. In such a scenario, Smith's guilt would reduce the probability of Jones's guilt and vice versa, so perhaps they could be considered as competing hypotheses. The difficulty with this is that it fails to take into account the influence of the evidence on the relationship between the hypotheses. For example, two hypotheses, H and H' could be negatively dependent i.e. $P(H|H') < P(H|\neg H')$ and yet become positively dependent when evidence is taken into account. This would occur if they account for the evidence much better when combined than either would on its own. This suggests that competition between hypotheses should be defined with respect to evidence, which motivates the following definition:

Definition 1 Let each of H and H' be hypotheses and E evidence under consideration and suppose that $P(H\&E)$ and $P(H'\&E)$ are greater than zero. Hypotheses H and H' are said to be **competing hypotheses** with respect to evidence E if and only if $P(H|H'\&E) < P(H|E)$.

Since the relevant condition can also be expressed as $P(H'|H\&E) < P(H'|E)$, it is clear that competition is a symmetric concept. This definition of competition can be expressed in terms of Bayesian confirmation by noting that H and H' are in competition with respect to E if and only if H' disconfirms H given E. It can be shown that an equivalent condition for competition can be stated as follows:[2]

$$P(E|H\&H') \cdot P(E|\neg H\&\neg H') \cdot P(H'|H) \cdot P(\neg H'|\neg H)$$
$$< P(E|\neg H\&H') \cdot P(E|H\&\neg H') \cdot P(H'|\neg H) \cdot P(\neg H'|H). \quad (1)$$

In recent work, Schupbach and Glass [28] defined a measure of the degree of competition between two hypotheses, H and H', with respect to evidence E as the average degree to which H and H' disconfirm each other given E [28]:

$$\mathrm{Comp}(H, H'|E) = \frac{C_l(H, \neg H'|E) + C_l(H', \neg H|E)}{2} \quad (2)$$

where C_l is the likelihood ratio measure of confirmation conditioned on E, that is,

$$C_l(H, H'|E) = \log\left[\frac{P(H|H'\&E)}{P(H|\neg H'\&E)}\right]. \quad (3)$$

So defined, Comp takes increasingly positive values to the extent that H and H' disconfirm one another given E, decreasingly negative values to the extent that H and H' confirm one another given E, and zero when H and H' are are probabilistically independent given E. Note that if H and H' are assumed to be mutually exclusive, then $P(H|H'\&E) = 0$

---

[1] While it would be interesting to consider when the conjunctive hypothesis $H\&H'$ (or for that matter the disjunctive hypothesis $H \vee H'$ or the exclusive disjunctive hypothesis $H \veebar H'$) is to be preferred to either H or H', our focus here is on whether H and H' compete with each other.

[2] From $P(H|H'\&E) < P(H|E)$, it follows that $P(H|H'\&E) < P(H|\neg H'\&E)$ and applying Bayes' theorem gives $P(E|H\&H')P(H|H')/P(E|H') < P(E|H\&\neg H')P(H|\neg H')/P(E|\neg H')$. Expanding the term $P(E|H')$ as $P(E|H\&H')P(H|H') + P(E|\neg H\&H')P(\neg H|H')$ and similarly for $P(E|\neg H')$ gives the desired result after some rearrangement.
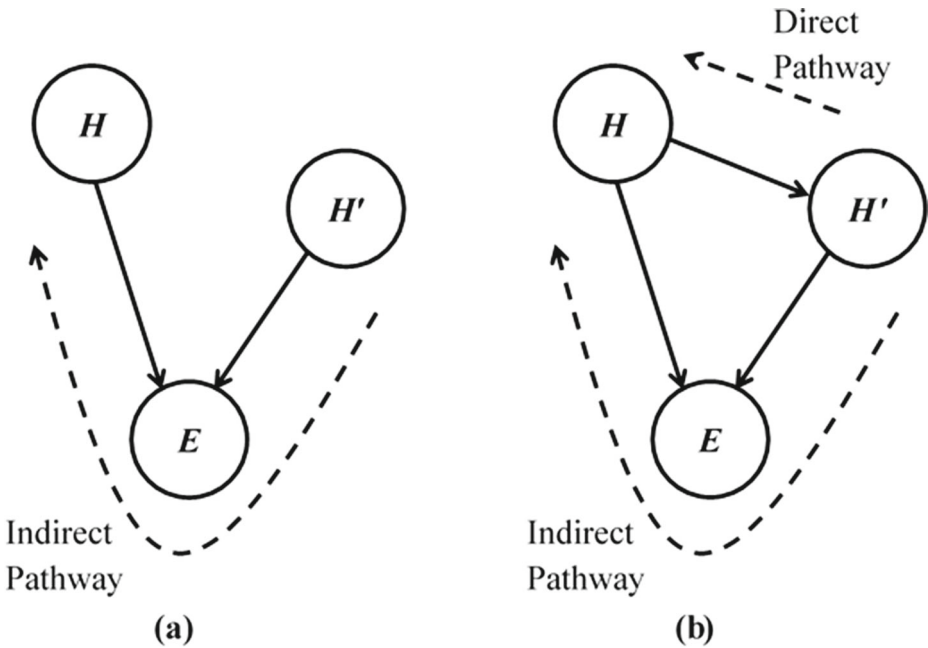
Fig. 1 If H and H' are unconditionally independent, they can nevertheless come into competition along an indirect pathway via the evidence (**a**), but otherwise they can also compete along a direct pathway (**b**) [28]

and hence this would be a case with the highest possible degree of competition.[3] H and H' can be said to compete with respect to E if Comp > 0. It can then be shown that H and H' compete with respect to E if the condition in Definition 1 is met. Schupbach and Glass also demonstrate that Comp can be expressed in terms of two pathways: a direct pathway between H and H', which depends on the extent to which H and H' disconfirm each other unconditionally, and hence does not depend on the evidence E, and an indirect pathway via the evidence E (see Fig. 1).[4] Since the measure of competition lies in the range $[-\infty, \infty]$ their alternative measure, which lies in the range $[-1, 1]$, will be used for convenience in Sections 3 and 4. It is given by [28]

$$\text{Comp}_k(H, H'|E) = \frac{C_k(H, \neg H'|E) + C_k(H', \neg H|E)}{2} \tag{4}$$

where C_k is the confirmation measure proposed by Kemeny and Oppenheim [29] when conditioned on E,

$$C_k(H, H'|E) = \left[\frac{P(H|H'\&E) - P(H|\neg H'\&E)}{P(H|H'\&E) + P(H|\neg H'\&E)} \right]. \tag{5}$$

---

[3]Strictly speaking it is not defined since there is a division by zero in this case. However, the likelihood ratio measure, C_l, is often taken to be infinite in this case and so the same approach could be adopted for Comp. More straightforwardly, Comp_k in (4) will take on the maximum value of one when H and H' are mutually exclusive.

[4]$\text{Comp}(H, H'|E) = \text{Comp}_D(H, H') + \text{Comp}_I(H, H'|E)$, where the direct competition, $\text{Comp}_D(H, H') = \left[C_l(H', \neg H) + C_l(H, \neg H') \right]/2$ and the indirect competition $\text{Comp}_I(H, H'|E) = \left[C_l(E, H) - C_l(E, H|H') + C_l(E, H') - C_l(E, H'|H) \right]/2$.

2.3 Features of the account

Motivation and justification for the above approach can be found in [28], but here it is worth highlighting the importance of the fact that competition according to this account depends not only on the hypotheses in question but also on the evidence under consideration, which means that hypotheses could be competing with respect to one piece of evidence and non-competing with respect to another. As noted earlier, merely considering negative dependence between the hypotheses would be inadequate, but it is still relevant to competition via the direct pathway. Here some features of the account relating to the interplay between direct and indirect pathways contributing to competition are explored.

Since expression (1) needs to be false for non-competing hypotheses and it is easy to see that $P(H'|H) \cdot P(\neg H'|\neg H) < P(H'|\neg H) \cdot P(\neg H'|H)$ in the case of negative dependence, then provided $P(E|H\&H') \cdot P(E|\neg H\&\neg H')$ is sufficiently greater than $P(E|H\&\neg H') \cdot P(E|\neg H\&H')$ this would be sufficient to ensure non-competition. In terms of the different pathways that contribute to competition, this means that a negative dependence along the direct pathway can be outweighed by a positive dependence along the indirect pathway, which depends on the evidence. However, a stronger result shows that this can be achieved even if both hypotheses are confirmed by the evidence E, i.e.,

$$P(E|H) > P(E|\neg H) \text{ and } P(E|H') > P(E|\neg H'). \tag{6}$$

The following result shows that this can be achieved irrespective of how negatively dependent H and H' are, provided they are not mutually exclusive. Returning to our murder example, this means that even if there is very good reason to think that Smith and Jones do not collaborate in general and so there is negative dependence between them, this could be overturned by appropriate evidence that could best be explained by their joint involvement, in which case the hypotheses should not be considered to be competing. To simplify notation, it will be useful to let $x = P(H'|H)$, $y = P(H'|\neg H)$, $x' = P(H|H')$ and $y' = P(H|\neg H')$ and for the likelihoods $a = P(E|H\&H')$, $b = P(E|H\&\neg H')$, $c = P(E|\neg H\&H')$ and $d = P(E|\neg H\&\neg H')$.

Theorem 1 *Let H and H' be hypotheses and E evidence under consideration such that E confirms both H and H' and let H and H' be negatively dependent, i.e. $P(H'|H) < P(H'|\neg H)$. Unless H and H' are mutually exclusive, H and H' are not necessarily in competition with respect to E.*

Proof Using the notation provided above consider a scenario where $a > d = b = c$, i.e. a probability model where the likelihood is greater for $H\&H'$ than it is for the other conjunctions of hypotheses or their negations. The first condition in (6) can be expressed as

$$P(E|H\&H') \cdot P(H'|H) + P(E|H\&\neg H') \cdot P(\neg H'|H)$$
$$> P(E|\neg H\&H') \cdot P(H'|\neg H) + P(E|\neg H\&\neg H') \cdot P(\neg H'|\neg H).$$

Given the definition and assumption that $b = c = d$ this condition can in turn be expressed as $ax + d(1-x) > dy + d(1-y)$ which becomes $(a-d)x > 0$ and since $a > d$ this condition is satisfied. Similarly, the second condition in (6) is also satisfied. The condition of H and H' to be non-competing is that expression (1) be false, i.e

$$adx(1-y) \geq bcy(1-x).$$

Since $a > d = b = c$, this gives

$$\frac{a}{d} \geq \frac{y(1-x)}{x(1-y)}.$$

If $a > y(1-x)$ and $0 < d < x(1-y)$ this condition will be satisfied and in such a case H and H' will be non-competing irrespective of how much smaller $y = P(H_2|\neg H_1)$ is than $x = P(H'|H)$ provided $y \neq 0$. $\square$

Just as negatively dependent hypotheses need not be competing, so positively dependent hypotheses can be competing. Contrary to the earlier supposition, if it is known that Smith and Jones often work together in their criminal activity and hence are likely to have colluded in this case, this could be overturned if there is compelling evidence that the murder was committed by one person working alone. This means that while there is a positive dependence along the direct pathway, it can be outweighed by a sufficiently strong negative dependence along the indirect pathway. This is captured in the following result.

Theorem 2 *Let H and H' be hypotheses and E evidence under consideration such that E confirms both H and H' and let H and H' be positively dependent, i.e. $P(H'|H) > P(H'|\neg H)$. Unless H entails H' or vice versa, it is possible that H and H' will be in competition with respect to E.*

Proof Using the notation provided earlier consider a probability model where $a > \max(b, c)$ and $\min(b, c) > d = 0$. Since the necessary condition for H and H' to be competing in expression (1) can be expressed as $adx(1-y) < bcy(1-x)$, then this condition is satisfied provided $y > 0$ and $x < 1$. However, since E confirms both H and H', the conditions in (6) must be satisfied. The first of these conditions can be expressed as $ax + b(1-x) > cy + d(1-y)$ and since $d = 0$, $a > c$ and $x > y$ (since there is a positive dependence between H and H') then this condition is satisfied. Similarly, the second of these conditions can be expressed as $ax' + c(1-x') > by' + d(1-y')$ and since $d = 0$, $a > b$ and $x' > y'$ (since there is a positive dependence between H and H' and so $P(H|H') > P(H|\neg H')$) then this condition is also satisfied. $\square$

Suppose that each of the hypotheses in conjunction with the negation of the other fails to provide an explanation that raises the probability of the evidence, but that the conjunction of both hypotheses does achieve this. As noted earlier, this seems like an example where the hypotheses would not be in competition since they need to work together to account for the evidence, but as the following result shows this need not be the case if there is a sufficiently strong negative dependence between the two hypotheses before the evidence is taken into account. This result can be seen as a counterpart to Theorem 1. In Theorem 1 a negative dependence along the direct pathway could be outweighed by a positive dependence along the indirect pathway, but here the converse is true where a positive dependence along the indirect pathway can be outweighed if there is a strong enough negative dependence along the direct pathway. In the murder investigation, it could be the case that the evidence is better explained by Smith and Jones having colluded, but this positive dependence along the indirect pathway is insufficient to outweigh previous evidence of their rivalry and hence they are considered as competing hypotheses.

Theorem 3 *Let H and H' be hypotheses and E evidence under consideration such that E confirms both H and H' and let the likelihood for the conjunction of one hypothesis and the negation of the other be equal to that of the catchall hypothesis, H^c, which is assumed to*

be non-zero, i.e. $P(E|H\&\neg H') = P(E|\neg H\&H') = P(E|H^c) > 0$, while the likelihood for the conjunction of the two hypotheses is greater than that of the catchall. If H and H' are independent or positively dependent, i.e. $P(H'|H) \geq P(H'|\neg H)$, then they are not in competition with respect to E, whereas if they are negatively dependent they can be competing.

Proof Using the notation provided earlier, $ad > bc$ which in the independent case corresponds to expression (1) being false and so they are not competing. If they are positively dependent, then $x > y$ and so $x(1 - y) > y(1 - x)$. This gives $adx(1 - y) > bcy(1 - x)$ which corresponds to expression (1) being false and so again they are not competing.

In the proof of theorem 1 it was shown that negative dependence between the hypotheses could be outweighed in the case where the two hypotheses need each other to account for the evidence, i.e. where $a > d = b = c$ as in the current case, so that the two hypotheses are not competing. Here the goal is to show that there are scenarios where the same condition holds, i.e. $a > d = b = c$, but the negative dependence outweighs the fact that they need each other and results in their being in competition. For this to be the case the following condition must be satisfied

$$\frac{a}{d} < \frac{y(1 - x)}{x(1 - y)}.$$

If $x = d$ and $y = a$ this condition will be satisfied since $a > d$. Furthermore, it is easy to show that the conditions in (6) will be satisfied provided $x > 0$ and $x' > 0$. Hence, H and H' will be competing hypotheses. □

The results presented so far demonstrate some plausible features of this account of competition between hypotheses, which goes beyond mutual exclusion. In light of our murder example, the results highlight the fact that whether suspects should be viewed as competing hypotheses depends on a subtle interplay between the direct pathway, which relates to previous knowledge about whether they are rivals or work together, and the indirect pathway, which relates to how well their guilt either individually or together would account for the current evidence.

The following result highlights a further important feature of the account.[5]

Theorem 4 *Let H and H' be hypotheses, E evidence under consideration and suppose that $\neg H$ and $\neg H'$ can also be considered as hypotheses for E. H and H' are competing (non-competing) with respect to E if and only if $\neg H$ and $\neg H'$ are also competing (non-competing) with respect to E.*

Proof Necessity. Suppose H and H' compete with respect to E. From Definition 1, this means that $P(H|H'\&E) < P(H|E)$. From this it follows that $P(H|H'\&E) < P(H|\neg H'\&E)$ and hence that $P(\neg H|H'\&E) > P(\neg H|\neg H'\&E)$. It in turn follows that $P(\neg H|E) > P(\neg H|\neg H'\&E)$ and hence from Definition 1 that $\neg H$ and $\neg H'$ compete with respect to E. Sufficiency can be established similarly and the result for non-competition follows trivially. □

---

[5]It might be questioned whether $\neg H$ and $\neg H'$ should be considered as hypotheses, but there is nothing in the definition or foregoing discussion to preclude doing so. However, even if they were not considered as hypotheses the relevant probabilistic relations would still apply.

The definition of competition tells us that if H and H' are competing with respect to E, then learning that H' is true will reduce the probability of H. Suppose that H and H' are non-competing with respect to E. Theorem 4 tells us that learning that H' is false cannot make it more likely that H is true. In fact, if H and H' have a negative degree of competition, then learning that H' is false makes it more likely that H is false as well.

This suggests that the account sets the bar quite high for two hypotheses to be non-competing (or quite low for them to be competing). Two hypotheses might have some explanatory merit on their own, but intuitively provide a much better explanation when combined together as a 'conjunctive explanation' and both might be well-supported by the evidence. Yet, if they are also the only plausible explanations on offer such that if one is discovered to be false, it becomes very likely that the other one is true, then according to Theorem 4 they must be in competition. Returning to our murder suspects, Smith and Jones, consider the following example.

Example 1 Suppose that before any evidence is considered, the hypotheses that Smith committed the murder (H) and that Jones did so (H') are probabilistically independent and each has a prior probability of $P(H) = P(H') = 0.2$. Suppose also that the probability of the evidence available is much higher if Smith and Jones worked together, $P(E|H\&H') = 0.9$, than it is if either committed the murder alone, $P(E|H\&\neg H') = P(E|\neg H\&H') = 0.2$, which in turn is much higher than if neither was involved, $P(E|\neg H\&\neg H') = 0.01$. The posterior probability of Smith's guilt is $P(H|E) = 0.639$. If subsequent evidence comes to light showing that Jones was not involved (but otherwise the evidence is conditionally independent of Smith's guilt), then it becomes even more probable that Smith is guilty, $P(H|\neg H'\&E) = 0.833$, which implies that H and H' are in competition with respect to E. This can be confirmed by noting that $P(H|E) > P(H|H'\&E) = 0.529$.

Even though the two hypotheses seem to work very well together to provide a better account of the evidence, they are nevertheless in competition according to the account of competition presented earlier. Is this a problem for the account? In a *broad sense of competition* it might be argued that intuitively the hypotheses are not competing in this case since they work together effectively and are both supported by the evidence. Alternatively, it could be argued that in a *narrow, but well-defined, sense* they are competing since they are negatively dependent given the evidence.

It would certainly be worth exploring this issue further, but it will not be pursued here since for current purposes, the narrow sense of competition provided captures key features of the relationship between the hypotheses, including the two pathways which affect their interdependence. As such, it provides a suitable framework for the rest of the paper, where the goal is to explore how treating hypotheses as mutually exclusive or independent affects abductive inference as the interdependence between the hypotheses varies.

3 Abductive inference

Having considered the nature of competition, this section investigates what bearing this has on abductive inference. In particular, if incorrect assumptions are made about the hypotheses, what effect does that have on the reliability of the inferences made? Returning to our murder example, it might be difficult to establish the mutual exclusivity of the suspects' guilt, but would mutual exclusion be a reasonable assumption nevertheless? How would the reliability of inferences made depend on such an assumption, and how does this

vary with the degree of competition? Alternatively, if the guilt of the suspects were incorrectly assumed to be probabilistically independent, so that prior to considering the evidence, Smith's guilt was assumed to have no bearing positively or negatively on Jones's guilt and vice versa, how would this affect inference?

In order to investigate the impact of failing to take into account the nature of competition between two hypotheses properly, three experiments were carried out on abductive inference. Each of the experiments involves generating a probability model involving evidence E and hypotheses H, H' and a catchall hypothesis, $H^c = \neg H \& \neg H'$, with a specified degree of competition between H and H' given E. This model is stipulated to be the correct model and three different incorrect models are used for inference in each of the experiments. One reason for including a catchall hypothesis is that in most cases we cannot assume we have an exhaustive list of hypotheses and so the catchall provides a way of representing our ignorance. Also, since in some cases the probability of neither H nor H' will be greater than the probability of either, allowing the catchall to be inferred in the inference process ensures that the inference procedure is not forced into inferring a false hypothesis. Hence its inclusion provides a fairer way to evaluate the different approaches. In the first two experiments, abductive inference is performed based on a modified probability model in which H and H' are treated as mutually exclusive and the success of this approach is evaluated by comparing the inference with an inference made using the correct model. The third experiment proceeds in the same way, but uses a modified probability model in which H and H' are treated as independent before the evidence is taken into account.

3.1 Generating a probability model

The measure of competition given in (4) is used so that its value lies in the interval $[-1, 1]$. There are many different probability models that can give rise to the same value of competition, so the idea in the simulations is to sample this space of models multiple times and then repeat the process for different values of competition. For a given value of competition, d, a probability model is generated as follows. First, note that d can be expressed as $(d_1 + d_2)/2$, where $d_1 = (b_1 - a_1)/(b_1 + a_1)$ and $a_1 = P(H|H' \& E)$, $b_1 = P(H|\neg H' \& E)$, and $d_2 = (b_2 - a_2)/(b_2 + a_2)$, where $a_2 = P(H'|H \& E)$ and $b_2 = P(H'|\neg H \& E)$. For a positive value of d, (a similar approach is adopted for negative values), d_1 can then be selected randomly from a uniform distribution over the interval $(d - \epsilon, d + \epsilon)$, where ϵ is set to 0.1, and d_2 is then set to $2d - d_1$. This ensures values for d_1 and d_2 are selected from intervals of the same size for each value of d and so allows for a meaningful comparison of results. Given this choice, only values of d between -0.9 and 0.9 are considered. The following equations can then be used to generate a probability model based on these values of d_1 and d_2:

$$P(H|E) = a_1 \cdot P(H'|E) + b_1 \cdot P(\neg H'|E), \tag{7}$$
$$P(H'|E) = a_2 \cdot P(H|E) + b_2 \cdot P(\neg H|E), \tag{8}$$

and

$$a_1 \cdot P(H'|E) = a_2 \cdot P(H|E). \tag{9}$$

Since $b_1 = a_1(1 + d_1)/(1 - d_1)$ and $b_2 = a_2(1 + d_2)/(1 - d_2)$, b_1 and b_2 can be replaced in (7) and (8). $P(H|E)$ is selected randomly and then the three equations above can be used to determine values for three unknowns, $P(H'|E)$, a_1 and a_2.[6]

---

[6]In the case where the degree of competition is zero, things are slightly different. Given that competition is defined in terms of confirmation conditional on E, if $d = 0$ then H and H' are probabilistically independent

The foregoing is sufficient to ensure that the degree of competition between H and H' given E is specified by d. However, a full probability model is needed so there is more to be done. $P(E)$ can be set randomly on the interval $(0, 1)$, while $P(H|\neg E)$ can be set randomly on the interval $(0, P(H|E))$ and $P(H'|\neg E)$ on the interval $(0, P(H'|E))$, which ensures that E confirms both H and H'. Finally given that

$$P(H'|\neg E) = P(H'|H\&\neg E)P(H|\neg E) + P(H'|\neg H\&\neg E)P(\neg H|\neg E), \quad (10)$$

and letting $x = P(H|\neg E)$ and $y = P(H'|\neg E)$, $P(H'|\neg H\&\neg E)$ can be set randomly on the interval $(\max[0, (y - x)/(1 - x)], \min[y/(1 - x), 1])$ and then $P(H'|H\&\neg E)$ obtained by rearranging (10).

In specifying the probability model a number of random assignments have been made. To get meaningful results the experiments are run multiple times (10^6) and average results obtained, as will be discussed later.

3.2 Mutually exclusive models

As noted earlier, the first two experiments treat H and H' as mutually exclusive hypotheses. In the first experiment, which will be referred to as MEx1, a probability model, denoted P_1, is obtained from the original model, P, essentially by replacing H with $H\&\neg H'$ and H' with $H'\&\neg H$. More precisely, P_1 is obtained by setting $P_1(H) = P(H\&\neg H')$, $P_1(H') = P(H'\&\neg H)$, $P_1(E|H) = P(E|H\&\neg H')$, $P_1(E|H') = P(E|\neg H\&H')$, $P_1(H^c) = 1 - P_1(H) - P_1(H')$ since this is equal to $P_1(\neg H\&\neg H') = P_1(\neg(H \vee H'))$ when $P_1(H\&H') = 0$, and $P_1(E|H^c) = P(E|H^c)$ (which is just $P(E|\neg H\&\neg H')$).

However, this is not the only way to obtain a mutually exclusive model from the original model, so in the second experiment, which will be referred to as MEx2, a different strategy is adopted. A probability model, denoted P_2, is obtained from the original model, P, by basing the probabilities of $P_2(H)$ and $P_2(H')$ on $P(H)$ and $P(H')$ rather than on $P(H\&\neg H')$ and $P(\neg H\&H')$. This is done in such a way that $P_2(H \vee H') = P_2(H) + P_2(H') = P(H \vee H')$, which means that $P_2(H\&H') = 0$ and the total area of the probability space taken up by the hypotheses H and H' remains the same as it was in the original model. This is achieved by a normalizing factor, $P(H \vee H')/(P(H) + P(H'))$, which is multiplied by $P(H)$ and $P(H')$ to yield $P_2(H)$ and $P_2(H')$ respectively. Apart from that, the probabilities for the likelihood terms are set in the same way as for P_1 and the probability for the catchall hypothesis, H^c, is similarly set to $P_2(H^c) = 1 - P_2(H) - P_2(H')$ since H and H' are assumed to be mutually exclusive.

3.3 Independence model

In contrast to the first two models, the third experiment, which will be referred to as IND, treats H and H' as independent. One reason for considering a model of this kind is that it provides a compromise between incorrectly treating hypotheses as mutually exclusive on the one hand and fully taking into account the dependence between them on the other, so it is interesting to see how it performs. Another reason is that in some cases hypotheses are modelled as being independent. For example, in probabilistic horn abduction, hypotheses, which can be represented as root nodes in a Bayesian network, are probabilistically

---

given E and from this it follows that d_1 and d_2 must also be zero as well. In the general case the value of $P(H'|E)$ can be determined from $P(H|E)$, d_1 and d_2, but this is not the case when $d = d_1 = d_2 = 0$. In this case a value for $P(H'|E)$ can be assigned randomly.

Table 1 Summary of probability models used in the experiments

	$P_1(\cdot)$, MEx1	$P_2(\cdot)$, MEx2	$P_3(\cdot)$, IND
$P_i(H)$	$P(H\&\neg H')$	$P(H)\cdot\frac{P(H\vee H')}{P(H)+P(H')}$	$P(H)\cdot\frac{P(H\vee H')}{P(H)+P(H')}$
$P_i(H\&H')$	0	0	$P_3(H)\cdot P_3(H')$
$P_i(H^c)$	$1-P_1(H)-P_1(H')$	$1-P_2(H)-P_2(H')$	$1-P_3(H\vee H')$
$P_i(E\|H)$	$P(E\|H\&\neg H')$	$P(E\|H\&\neg H')$	$P(E\|H\&H')\cdot P_3(H')$
			$+P(E\|H\&\neg H')\cdot P_3(\neg H')$

Note that for a given probability model $P_i(\cdot)$ some of the expressions are based on the original probability model, $P(\cdot)$, and some on terms already defined in the current model, $P_i(\cdot)$

independent, though minimal explanations consisting of multiple hypotheses are typically mutually exclusive [12]. More generally, independence assumptions are relevant to any context in which root nodes in a Bayesian network constitute hypotheses. The probability model in this case, denoted P_3, is obtained from the original model, P, by first of all obtaining $P_3(H)$ and $P_3(H')$ in the same way as used for P_2, except that now the joint probability is set to $P_3(H\&H') = P_3(H)P_3(H')$. Furthermore, since H and H' are not treated as being mutually exclusive in this case, the likelihoods are treated differently since the term $P_3(E|H\&H')$ cannot be ignored. This is done by setting $P_3(E|H\&H') = P(E|H\&H')$ and similarly for the other likelihood terms. $P_3(E|H)$ is then obtained via the expression $P_3(E|H) = P_3(E|H\&H') \times P_3(H') + P_3(E|H\&\neg H') \times P_3(\neg H')$ and $P_3(E|H')$ is obtained in the same way. In this case the probability for the catchall hypothesis is given by $P_3(H^c) = 1 - P_3(H) - P_3(H') + P_3(H)P_3(H')$. Key features of both the mutually exclusive and independence models are presented in Table 1.

3.4 Experiments

With the models in hand, the experiments were carried out as follows (described here for MEx1):

Experiment MEx1

1. For a specified value of the degree of competition, generate the correct probability model, P.
2. From the correct model, generate the incorrect model P_1.
3. Perform abductive inference for both the correct and incorrect models, i.e. for the correct model find the hypothesis, H, H' or H^c which maximizes the posterior probability given evidence, E, and then do the same for the incorrect probability model, P_1.
4. If the hypothesis identified by abduction using the incorrect model, P_1, is the same as that identified by the correct model, P, count that as a success for P_1.
5. Repeat 1–5 multiple (10^6) times to determine the accuracy (percentage success) for P_1 for this value of the degree of competition.
6. Repeat 1–6 for a range of values of the degree of competition between -1 and 1.

Exactly the same procedure was followed for MEx2 and IND. Note that the approach here is to maximize the posterior probability, but this could be questioned as an approach to abductive inference [7, 23] and instead other measures could be used for comparing

hypotheses [8, 9, 24]. Nevertheless, inference based on maximizing posterior probability is frequently used in abductive inference in the artificial intelligence literature through approaches such as *Maximum a Posteriori* (MAP) and *Most Probable Explanation* (MPE) and so it provides a standard approach which is appropriate for investigating how the degree of competition affects inference.

Two further modifications of the above experiments were also carried out. According to the correct model, the two hypotheses, H and H' are not mutually exclusive and so could both have a probability greater than 0.5 in some cases. Clearly, this is not possible in the case of MEx1 or MEx2. Even in cases where they correctly infer the same hypothesis as the correct model, they will not infer a second hypothesis even if that hypothesis is more likely to be true than false according to the correct model. How much of a weakness is this? Does this scenario occur frequently to the detriment of MEx1 and MEx2? Similarly, how well does IND fare in this respect?

As a first step, the above experiments were repeated, but only cases where either H or H' (or both) have a posterior probability greater than 0.5 according to the correct model, P, were taken into account. That is, step 4 in the above procedure is modified as follows to give another experiment MEx1-0.5:

Experiment MEx1-0.5
4′. If the hypothesis identified by abduction using the incorrect model, P_1, is H or H', and if it is the same as that identified by the correct model, P, and if the posterior probability of that hypothesis is greater than 0.5 according to P, count that as a success for MEx1. If the posterior probability of this hypothesis according to P is less than or equal to 0.5 it is not taken into account in determining the accuracy.

Corresponding experiments MEx2-0.5 and IND-0.5 were also carried out. In these experiments, the accuracy is defined to be the percentage success in identifying whichever of H and H' has the greatest posterior probability according to the correct model, P, in cases where this probability is greater than 0.5.

Following this, further experiments were carried in which, if there is a second hypothesis with probability greater than 0.5 (according to the correct model), it is taken into account in determining the accuracy. That is, step 4′ is used again, but now step 5 is modified to give another experiment MEx1-2:

Experiment MEx1-2
5′. Repeat 1–5 multiple (10^6) times to determine the accuracy (percentage success) for P_1 for this value of the degree of competition taking into account the total number of hypotheses with probability greater than 0.5 according to the correct model.

That is, the accuracy is defined as

$$\frac{\text{Total number of correctly identified hypotheses}}{\text{Total number of hypotheses with } P(.|E) > 0.5} \times 100\%, \tag{11}$$

where this is restricted to the hypotheses H and H' and as before correctly identified hypotheses are those where the inference made using the incorrect probability model (P_1 in the case of MEx1-2) matches the most probable hypothesis given the evidence according to the correct model, P. Again, corresponding experiments MEx2-2 and IND-2 were also carried out. Since there could be two hypotheses with posterior probability greater than 0.5 in the case of the independence model, a second hypothesis can be included in its successes in 4′ when this occurs.
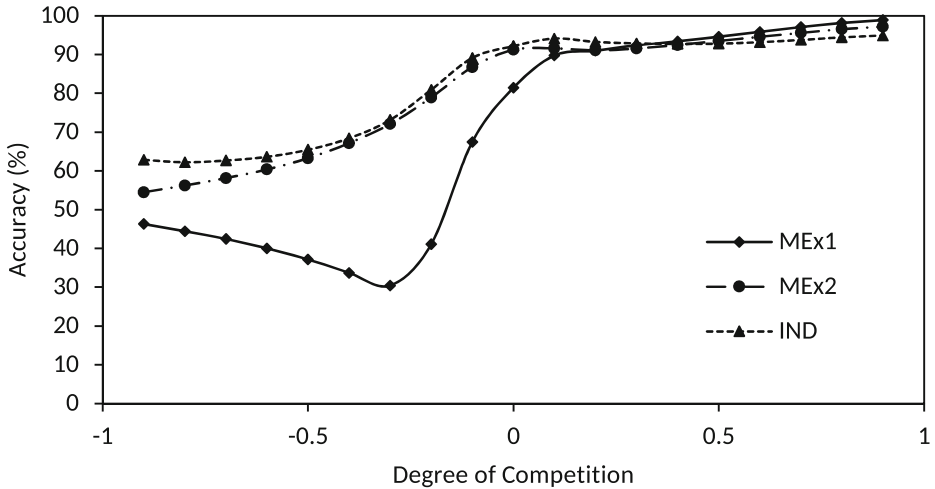
4 Results

The results for experiments MEx1, MEx2 and IND are presented in Fig. 2. All of the approaches achieve a high accuracy for high degrees of competition. This is not surprising. Recall that when the degree of competition is one this corresponds to the case where the probability of one hypothesis given the other, conditional on the evidence E, is zero according to the correct model. Hence, given E, H and H' are mutually exclusive, which of course is guaranteed by the mutually exclusive models (MEx1 and MEx2). The independence model (IND) does not guarantee this, but by representing the likelihoods more accurately than the other two models it is able to model the indirect negative dependence between H and H' conditional on E and so still achieves good results for high degrees of competition.

Interestingly, all three models perform well for positive degrees of competition, achieving an accuracy of over 90% in all cases for degrees of competition greater than 0.1. Hence, despite the rather drastic modifications of the correct probability model, particular for MEx1 and MEx2, where the hypotheses are treated as if they were mutually exclusive, these models all perform well in terms of abductive inference provided there is some positive degree of competition between the hypotheses given E.

For lower values of the degree of competition, the performance of MEx1 deteriorates markedly. When the degree of competition is zero, its accuracy has already dropped to 81% and it falls further to a value of 30% when the degree of competition is -0.3. Perhaps surprisingly, it then increases for still lower values of the degree of competition. The reason for this is that these are mostly cases in which MEx1 correctly identifies the catchall hypothesis, H^c. Overall, the poor performance of MEx1 for negative degrees of competition can be explained by the fact that it in effect represents H by $H\&\neg H'$ and H' by $H'\&\neg H$ and so discards the overlap between H and H'. By comparison with the correct probability model P, the probability model used in MEx1, ignores $P(H\&H')$ by simply assigning $P(H\&\neg H')$ to $P_1(H)$ and similarly for $P_1(H')$. By contrast, MEx2 does not ignore the overlap between H and H', but essentially reassigns $P(H\&H')$ to $P_2(H)$ and $P_2(H')$. This difference between MEx1 and MEx2 becomes much more important for negative degrees of competition because in these cases the overlap between H and H' plays a much more significant role. Indeed, in the extreme case where the degree of competition is -1, the probability of one hypothesis given the other conditional on the evidence is one.

The results for MEx2 and IND in Fig. 2 are very similar across the range of degrees of competition, with MEx2 performing slightly better at high values and IND slightly better at low values. The accuracy is lower for lower degrees of competition and it begins to drop off quite quickly for degrees of competition below zero, although not nearly as fast as it drops off for MEx1.
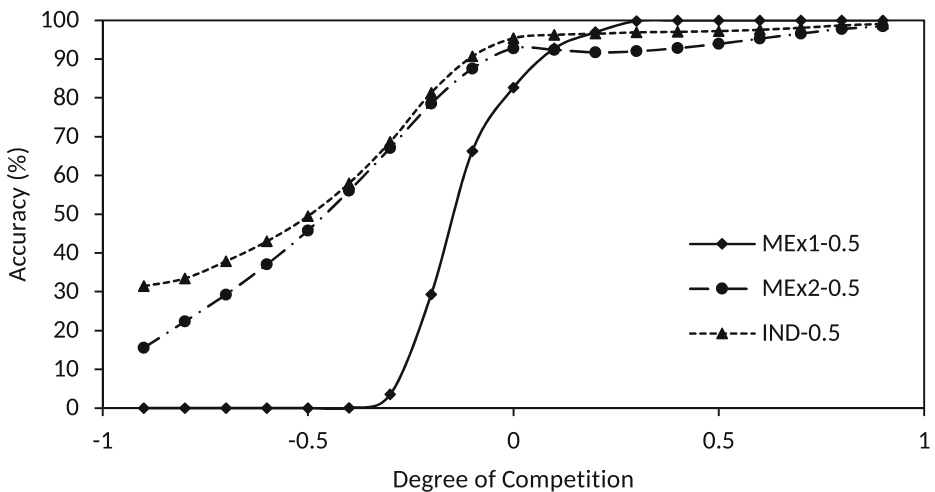
Figure 3 shows the results for experiments MEx1-0.5, MEx2-0.5 and IND-0.5. Recall that for these experiments the focus is on cases where either H or H' has the greatest posterior probability according to the correct model, P, and this probability is greater than 0.5. The results are similar to Fig. 2 for positive degrees of competition, but the accuracy drops off much more quickly for negative values in all three models. This is particularly so for MEx1-0.5 where the accuracy drops to almost zero at a degree of competition of -0.4 and then remains at zero for all lower values. This highlights the point that its higher accuracy in Fig. 2 was due to the catchall hypothesis. These low values can be attributed to the fact that a) MEx1-0.5 essentially ignores the overlap between H and H' in the correct probability model as noted earlier, and b) when there is a large negative degree of competition there
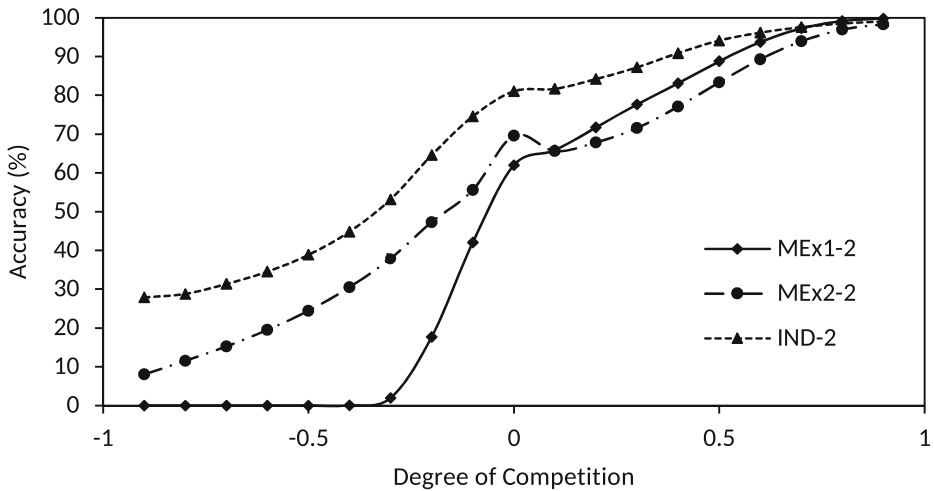
Fig. 2 Results for experiments MEx1, MEx2 and IND, where the accuracy is defined as percentage success in identifying the most probable hypothesis according to the correct probability model, P

can be a large overlap between these hypotheses, which results in small priors for H and H' in the probability model, P_1, used in MEx1-0.5, and hence makes it likely that an inference to the catchall hypothesis will be made.

The results for MEx2-0.5 and IND-0.5 in Fig. 3 are very similar across the range of degrees of competition, although in contrast to Fig. 2, IND-0.5 now does slightly better for both positive and negative degrees of competition. At very low degrees of competition the difference between these approaches becomes more apparent, with IND-0.5 outperforming MEx2-0.5. Like MEx1-0.5, MEx2-0.5 treats H and H' as mutually exclusive, but this is

Fig. 3 Results for experiments MEx1-0.5, MEx2-0.5 and IND-0.5, where the accuracy only considers cases where either H or H' have posterior probabilities greater than 0.5 according to the correct probability model

Fig. 4 Results for experiments MEx1-2, MEx2-2 and IND-2, where the accuracy takes into account the fact that both H and H' could have posterior probabilities greater than 0.5 according to the correct probability model

very inaccurate for very low degrees of competition, where there can be significant overlap between them, indeed complete overlap given the evidence when the degree of competition is -1. So while MEx2-0.5 does better than MEx1-0.5 in terms of the priors it assigns to H and H', it eventually runs into similar problems for very low degrees of competition.

Figure 4 presents the results for experiments MEx1-2, MEx2-2 and IND-2. Recall that for these experiments the fact that both H and H' could have a posterior probability greater than 0.5 is taken into account. Hence, even if any of the three approaches is successful in making an inference to the hypothesis with the greatest posterior probability according to the correct model, it may not identify when it is appropriate to make an inference to more than one hypothesis. Indeed, this is ruled out by the two mutually exclusive approaches since they cannot have two hypotheses with probability greater than 0.5. By contrast, IND-2 does allow for this. Its successes are counted in the normal way, but if its probability model, P_3 correctly identifies two hypotheses with posterior probability greater than 0.5 a second success is counted.

Unlike the previous two figures, the accuracy drops off more quickly even for positive degrees of competition, so that when the degree of competition is zero the accuracies are 62%, 70% and 81% for MEx1-2, MEx2-2 and IND-2 respectively. This result is significant because while Figs. 2 and 3 show that all three models perform quite well in terms of identifying the correct hypothesis for positive degrees of competition, Fig. 4 shows they are limited in that they fail to identify a second hypothesis which may also be true and this effect is non-negligible even for positive degrees of competition. Not surprisingly this limitation is even more evident for negative degrees of competition.

In the previous figures, there was very little to distinguish the MEx2 and IND approaches, but in Fig. 4 it becomes clear that the independence model has a significant advantage. While IND-2 is not immune to the problem of failing to identify a second hypothesis with probability greater than 0.5, the fact that it is able to do so in some cases means that its drop in performance between Figs. 3 and 4 is less marked than it is for MEx2.

5 Conclusion

Competition between hypotheses is fundamental to abductive inference as it is usually understood, but the nature of that competition and how it impinges upon inference have not received much attention. This paper has considered both of these topics by exploring a number of features of a particular account of hypothesis competition and the effect of failing to model this relationship between hypotheses correctly when performing abductive inference. In the former case, this involved studying the interplay between direct and indirect pathways, which contribute to competition. For example, pre-theoretically, it might be thought that hypotheses that are negatively dependent on each other would be competing, but this need not be the case as they may become positively dependent when evidence is taken into account. By exploring various features of the account, it was argued that it possesses a number of characteristics that make it suitable at least as an account of competition in a narrow sense suitable for the current work, although there is scope for further work on the nature of competition.

In terms of how this relates to inference, experimental results showed, rather surprisingly, that modelling the relationship between two hypotheses as mutually exclusive or independent can still generate good results provided the hypotheses are competing to some extent (a positive degree of competition), at least for the problems considered here. This would suggest that in applications to problems such as diagnosis, if hypotheses are modelled as mutually exclusive or independent, the results will still be accurate provided the hypotheses are negatively dependent given the evidence, though it would be interesting to explore this in the context of real data. Related to this, it would be worthwhile exploring the feasibility of using degrees of competition to discriminate between hypotheses so as to reduce the set of hypotheses that needs to be considered.

Less surprisingly, the performance of these models was much worse when the hypotheses were not competing (a negative degree of competition), with a model that completely ignores the overlap between the hypotheses performing much worse than the other two models. Despite the success of all three models for positive degrees of competition, an important failing is their inability to identify a second hypothesis that may well be true, and this is a significant effect even when the hypotheses are competing. This failing is due to the focus on maximizing posterior probability in the experiments carried out here. As noted earlier, this approach is generally assumed in the context of abductive inference in Bayesian networks. The results suggest merit in exploring approaches that allow the number of variables to vary and maximize quantities other than posterior probability [24, 30].

Directions for future research include carrying out similar experiments on a wider range of inference problems, exploring how inference is affected when other measures are used to compare explanations instead of maximizing posterior probability, and investigating how abduction might best be formulated and implemented in light of the approach to hypothesis competition considered here.

References

1. Josephson, J.R., Josephson, S.G.: Abductive Inference: Computation, Philosophy and Technology. Cambridge University Press, Cambridge (1994)
2. Flach, P.A., Kakas, A.C.: On the relation between abduction and inductive learning. In: Gabbay, D.M., Kruse, R. (eds.) Abductive Reasoning and Learning, pp. 1–33. Springer, Netherlands (2000)
3. Lipton, P.: Inference to the Best Explanation, 2nd edn. Routledge, London (2004)
4. Douven, I.: Abduction. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, winter 2016 edition (2016)
5. van Fraassen, B.C.: Laws and Symmetry. Clarendon Press, Oxford (1989)
6. Douven, I.: Inference to the best explanation, dutch books, and inaccuracy minimisation. Philos. Q. **63**(252), 428–444 (2013)
7. Chajewska, U., Halpern, J.Y.: Defining explanation in probabilistic systems. In: Proceedings of the 13th Conference on Uncertainty in AI, pp. 62–71 (1997)
8. Schupbach, J.N.: Comparing probabilistic measures of explanatory power. Philos. Sci. **78**(5), 813–829 (2011)
9. Glass, D.H.: Inference to the best explanation: does it track truth? Synthese **185**(3), 411–427 (2012)
10. Kakas, A.C., Kowalski, R.A., Toni, F.: Abductive logic programming. J. Log. Comput. **2**(6), 719–770, 12 (1992)
11. Denecker, M., Kakas, A.C.: Abduction in logic programming. In: Computational Logic: Logic Programming and Beyond, Essays in Honour of Robert A. Kowalski, Part I, pp. 402–436. Springer, London (2002)
12. Poole, D.: Probabilistic Horn abduction and Bayesian networks. Artif. Intell. **64**(1), 81–129 (1993)
13. Poole, D.: The independent choice logic for modelling multiple agents under uncertainty. Artif. Intell. **94**(1), 7–56 (1997)
14. Poole, D.: The independent choice logic and beyond. In: De Raedt, L., Frasconi, P., Kersting, K., Muggleton, S. (eds.) Probabilistic Inductive Logic Programming, pp. 222–243. Springer, Berlin (2008)
15. Lacave, C., Diez, F.J.: A review of explanation methods for Bayesian networks. Knowl. Eng. Rev. **17**(2), 107–127 (2002)
16. Pearl, J.: Probabilistic Reasoning in Intelligent Systems. Morgan Kaufman, San Mateo (1988)
17. Shimony, A.E.: Explanation, irrelevance and statistical independence. In: Proceedings of AAAI, vol. 1991, pp. 482–487 (1991)
18. Flores, M.J., Gámez, J.A., Moral, S.: Abductive inference in bayesian networks: finding a partition of the explanation space. In: ECSQARU (2005)
19. Park, J., Darwiche, A.: Complexity results and approximation strategies for map explanations. J. Artif. Intell. Res. **21**, 101–133 (2004)
20. Nielsen, U.H., Pellet, J.-P., Elisseeff, A.: Explanation trees for causal bayesian networks. In: Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence, UAI'08, pp. 427–434. AUAI Press, Arlington (2008)
21. Kwisthout, J.: Most probable explanations in Bayesian networks: complexity and tractability. Int. J. Approx. Reason. **52**(9), 1452–1469 (2011)
22. Kwisthout, J.: Most frugal explanations in Bayesian networks. Artif. Intell. **218**, 56–73 (2015)
23. Glass, D.H.: Coherence measures and inference to the best explanation. Synthese **157**, 275–296 (2007)
24. Yuan, C., Lim, H., Lu, T.-C.: Most relevant explanation in Bayesian networks. J. Artif. Intell. Res. **42**(1), 309–352 (2011)
25. Harman, G.: Competition for evidential support. In: Proceedings of the Eleventh Annual Conference of the Cognitive Science Society, pp. 220–225. NJ (1989)
26. Druzdzel, M.J., Henrion, M.: Intercausal reasoning with uninstantiated ancestor nodes. In: UAI '93: Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence, pp. 317–325 (1993)
27. Wellman, M.P., Henrion, M.: Explaining "explaining away". IEEE Trans. Pattern Anal. Mach. Intell. **15**, 287–292 (1993)
28. Schupbach, J.N., Glass, D.H.: Hypothesis competition beyond mutual exclusivity. Philos. Sci. **84**(5), 810–824 (2017)
29. Kemeny, J., Oppenheim, P.: Degrees of factual support. Philos. Sci. **19**, 307–324 (1952)
30. Glass, D.H.: An evaluation of probabilistic approaches to inference to the best explanation. Int. J. Approx. Reason. **103**, 184–194 (2018)