Causal Basis for Probabilistic Belief Change: Distance *vs.* Closeness

Seemran Mishra[1] and Abhaya Nayak[2(✉)]

[1] National Institute of Technology, Rourkela, India
seemran.mishra1996@gmail.com
[2] Macquarie University, Sydney, Australia
abhaya.nayak@mq.edu.au

Abstract. In probabilistic accounts of belief change, traditionally Bayesian conditioning is employed when the received information is consistent with the current knowledge, and imaging is used otherwise. It is well recognised that imaging can be used even if the received information is consistent with the current knowledge. Imaging assumes, *inter alia*, a relational measure of similarity among worlds. In a recent work, Rens and Meyer have argued that when, in light of new evidence, we no longer consider a world ω to be a serious possibility, worlds more similar to it should be considered relatively less plausible, and hence more dissimilar (distant) a world is from ω, the larger should be its share in the original probability mass of ω. In this paper we argue that this approach leads to results that revolt against our causal intuition, and propose a converse account where a larger share of ω's mass move to worlds that are more similar (closer) to it instead.

1 Introduction

Our knowledge is often fallible, uncertain and incomplete. How this knowledge should evolve in light of observations made and evidence acquired is a subject of much research. Research in this area can be divided into two broad approaches. In the first, "deterministic approach", knowledge is assumed to be certain but fallible, and the crux of the problem is to devise a rational model for modifying it in light of evidence that contravenes this knowledge. Literature in this approach is directly or indirectly inspired by the AGM paradigm [1] and deals with issues such as the problem of repeated belief change [10,11], the problem of belief change when knowledge is finitely represented [5] and the problem of belief modification when the world described by the knowledge is dynamic [7]. A knowledge state (or belief state) in such approaches is represented as a set of sentences together with a mechanism to capture the firmness of beliefs, semantically underpinned by a plausibility ranking of worlds. In the second, "probabilistic approach", a knowledge state is often represented as a probability function,

This research has been partially supported by the Australian Research Council (ARC), Discovery Project: DP150104133.

C. Sombattheera et al. (Eds.): MIWAI 2016, LNAI 10053, pp. 112–125, 2016.
DOI: 10.1007/978-3-319-49397-8_10

with beliefs being interpreted as "full beliefs", meaning propositions assigned probability 1. Evidence with non-zero prior is then processed using standard Bayesian conditioning; however belief-contravening evidence, that is evidence with zero prior needs special treatment using "imaging" first introduced by David Lewis in the context of analysing conditionals [9].

It has been proposed that distance between worlds captured by a pseudo-distance function can be used to "implement" the notion of imaging and provide a construction of belief-contravening operations in a probabilistic framework. The idea is that if a world ω has to lose a portion (or all) of the probability apportioned to it, rationality dictates that this unaccounted for probability should be gained by the world that is closest to ω among the set of potentially deserving worlds.[1] In a recent work it has been suggested that such bias in the movement of probability may not be appropriate [13]. The rationale behind this suggestion is that a world that is losing probability is doing so because it is no longer deemed plausible, and hence, worlds similar to it would also suffer from a reduction in plausibility. Increasing probability of such worlds runs counter to this intuition, and so the probability salvaged should be distributed in proportion to their distance from the worlds losing that probability.

Out primary objective is to examine this suggestion. First of all, the suggestion that salvaged probabilities should be mostly contributed to worlds farthest from the "victim worlds" is based on abstract intuitions which are not much better than the counter-intuition that such probabilities should mostly be contributed to wolds closer to them. Consider for instance the counterfactual, *If Oswald had not killed Kennedy, someone else would have.* In order to evaluate such a counterfactual, it is natural to consider worlds that are very similar to the "real world" except that in them Oswald didn't kill Kennedy (but yet, it was the post-Cuban crisis period, JFK was a Democrat visiting Dallas on a reelection campaign, bullets are designed to kill people, and so on). If we take the suggestion in [13] seriously, we should instead consider worlds that are diametrically opposite to the real world in which, not only that Oswald didn't kill Kennedy, but also, presumably, JFK was a Republican visiting Alaska, bullets have salutary effects on humans, and so on which would make a complete mess of our understanding of counterfactuals. Secondly, while our intuitions about counterfactuals may not be very reliable,[2] we have quite strong causal intuitions. In the next section we look at a simple cause-effect scenario and examine what happens if, instead of doing Bayesian conditioning using the ratio principle we distribute probability along the way suggested by [13] and observe that it leads to counterintuitive consequences.

This leads us to explore the converse approach in causal domains. In Sect. 3 we define *closeness* between worlds based on the *distance* between them, and examine some properties of this measure. This notion of closeness is then used to develop an account of probabilistic reasoning, particularly probabilistic expansion, and

[1] Assuming such a unique closest world exists. Short of it, an appropriate distribution mechanism should be employed.

[2] Recall the standard refrain of the politicians, *I don't answer hypothetical questions.*

study its behaviour in our chosen example. Finally, in Sect. 4, we briefly outline the implications and limitations of our proposal, and our future research work.

2 Farther is not Better

We assume a propositional language and classical logic with standard notation. A world is defined to be a unique assignment of truth values $\{0, 1\}$ to all the atoms in the propositional language. The set of all the worlds is denoted by Ω. An agent's body of beliefs, denoted b, contains information as to different subjective probability it assigns to different worlds, and the set of all possible belief sets is denoted by B.

Definition 2.1. *A belief set b is the set of pairs $\bigcup_{\omega_i \in \Omega}\{(\omega_i, p_i)\}$ where world ω_i is allotted probability $p_i = P(\omega_i)$ such that $\Sigma_i p_i = 1$.*

It is understood that the probability that the agent assigns to a proposition Φ is given by $P(\Phi) = \Sigma_i\{p_i \mid \omega_i \models \Phi\}$.

The belief state of an individual is rationally modified in the light of new information Ψ that the agent receives. Received wisdom has it that the nature of this modification is sensitive to whether the knowledge domain is static or dynamic [7] and whether the new information contravenes current knowledge or not [4]. In this paper we will assume that the knowledge domain is static, and the received information is not knowledge contravening. In such case, as advocated in [4], the belief modification should be carried out using Bayesian conditioning, in other words, $b + \Psi = \bigcup_{\omega_i \in \Omega}\{(\omega_i, p'_i)\}$ where $p'_i = \frac{p_i}{P(\Psi)}$ if $\omega_i \models \Psi$ and 0 otherwise. Note that for this purpose no other machinery such as a plausibility ranking of worlds is used – if there is any implied notion of plausibility at all, it is presumably captured by probability.

Intuitively, similar scenarios are similarly plausible – that, at least, is the picture portrayed by plausibility rankings. However, this nice picture does not extend to probability calculation, as illustrated by the following example, and hence, arguably, *plausibility* and *probability* may not be reducible to each other.

Example 1. Three coins are tossed. We know that they are similarly biased: the odds of getting head are *same for each of the three coins* – either 9:1 or 1:9 – but we don't know which. If the bias favours heads, probabilities of the events $(HHH, HHT, HTH, HTT, THH, THT, TTH, TTT)$ are respectively $(0.729, 0.081, 0.081, 0.009, 0.081, 0.009, 0.009, 0.001)$. The respective probabilities are $(0.001, 0.009, 0.009, 0.081, 0.009, 0.081, 0.081, 0.729)$ on the other hand if the bias favours tails. Appealing to the principle of indifference we obtain $(0.365, 0.45, 0.045, 0.045, 0.45, 0.045, 0.045, 0.365)$ as the final probabilities. Clearly, the outcome HHH is more similar to HHT (and HTH and THH), and less similar to the outcome TTT. Yet, the outcomes HHH and TTT are equally probable, and that probability is very different from the probabilities of HHT!

Indeed, probability has been supplemented by an extraneous notion of plausibility in an account of probabilistic belief change advocated in [3]. In this approach it has been argued that *imaging* [9] which is designed to deal with knowledge update in dynamic worlds can also be used to capture probabilistic belief change in static worlds, and they use (pseudo-)distance between worlds [8] to compute the image of a world among a given set of worlds for this purpose. This approach can be used irrespective of whether the new information is belief-contravening or not. Nonetheless, this approach is rather restrictive in that given a world ω and a set X of worlds, the former has a unique image in the latter, that is, there is a unique world in X that is closer or more similar to ω than any other world in X. In a subsequent work [13] Rens and Meyer have developed a more general method that effectively allows multiple images of a world. We will now briefly outline how they deal with probabilistic belief change (called belief expansion) when the evidence does not contravene current knowledge before examining it. We assume *here onwards* that the new evidence is not belief-contravening.

In general, when evidence Ψ does not contravene existing knowledge, there are at least some worlds $\omega \models \Psi$ with non-zero prior. Evidence Ψ suggests that any world $\omega' \not\models \Psi$ is a scenario not longer deemed seriously possible and must be eliminated from the hypotheses space, that is it must be assigned a (posterior) probability 0, and the probability thus salvaged must be distributed among worlds $\omega \models \Psi$. For simplicity let us use the following notation:

Notation 1. *Removal set R and Acceptance set S.*

1. $R = \{\omega_{\in \Omega} \not\models \Psi\}$ *is the set of worlds that conflict with the evidence and should receive zero posterior.*
2. $S = \Omega \setminus R = \{\omega_{\in \Omega} \models \Psi\}$ *is the set of worlds consistent with the evidence and the probability salvaged from members of R should be distributed among its members.*

In the Bayesian conditioning, so to speak, the probabilities salvaged from members of R are all pooled together and then distributed among those in S in proportion to their prior probabilities. The worlds in S with zero prior will continue have zero posterior. In the approach described in [13], instead of using the priors in S as the basis of distribution, the distance between an individual world $\omega^{\times} \in R$ and the worlds $\omega \in S$ is used to determine the share of ω in the probability of $\omega^{\times}$. This distance between worlds is captured by a pseudo-distance function d.

Definition 2.2 *[8]. The pseudo-distance $d : \Omega \times \Omega \rightarrow Z$ between two worlds satisfies:*

1. $d(\omega, \omega') \geq 0$ *(Non-negativity)*
2. $d(\omega, \omega) = 0$ *(Identity)*
3. $d(\omega, \omega') = d(\omega', \omega)$ *(Symmetry)*
4. $d(\omega, \omega') + d(\omega', \omega'') \geq d(\omega, \omega'')$ *(Triangle Inequality)*

for all worlds ω, ω' and $\omega'' \in \Omega$, This distance function is then used to determine the different *weights* that members of $\omega \in S$ will carry while receiving their share of probability from any $\omega^\times \in R$. This weight, $\delta^{rem}(\omega^\times, \omega, S)$ is the distance $d(\omega^\times, \omega)$ normalised over the total distance of different worlds in S from $\omega^\times$.

Definition 2.3. *[13]* $\delta^{rem}(\omega^\times, \omega, S) = \dfrac{d(\omega^\times, \omega)}{\sum_{\omega' \in S} d(\omega^\times, \omega')}$

The weight $\delta^{rem}(\omega^\times, \omega, S)$ is used to compute total share of any $\omega \in S$ in the probability salvaged from R, denoted $\sigma(\omega, S, R)$:

Definition 2.4. $\sigma(\omega, S, R) = \sum_{\omega^\times \in R} P(\omega^\times) * \delta^{rem}(\omega^\times, \omega, S)$.

We note that the probability function P in Definition 2.4 is sensitive to the contextually fixed belief set b. The probability $P(\omega^\times)$ is the $p^\times$ extracted from the pair $(\omega,^\times p^\times) \in b$.

Finally, an operation $\langle prem \rangle : B \times 2^\Omega \to B$ is used to determine the result of modifying a belief set b in light of evidence Ψ by topping up the existing probability of each world in $\omega \in S$ by its share $\sigma(\omega, S, R)$.

Definition 2.5. $b\langle prem \rangle R = \bigcup_{\omega \in S} \{(\omega, p') \mid (\omega, p) \in b \text{ } and \text{ } p' = p + \sigma(\omega, S, R)\}$.

Let us illustrate the application of the proposed probabilistic reasoning using a simple example that also brings to the surface a problem with this approach.

> *Example 2.* The major causes of asthma are polluted air and stress. **R** is a factory town with bad air pollution. On a given day, the chance of **R** having high pollution index is 60 %. On the other hand, children in **R** have easy, stress-free life, and the probability that a child in this town suffers from stress is negligible, say 5 %. Both pollution and stress are equally efficacious in causing asthma, with a probability of 10 % each. In presence of both pollution and stress, there is a multiplier effect and the chance of asthma attack goes up to 25 %. On a particular day a little child who lives in **R** suffered from an asthma attack. Between pollution and stress, which factor should we blame?

This scenario is compactly represented as a Bayesian Network [12] as depicted in Fig. 1 below.

Intuitively, since pollution and stress are equally efficacious in causing asthma attack, and the prior of pollution is a lot higher, around 12 fold, than the prior of stress, pollution is more likely the cause of the asthma attack. Indeed this is indicated by Bayesian updating: the posterior probabilities, $P'(P) = P(P|A) = 0.91$ and $P'(S) = P(S|A) = 0.13$. Both pollution and stress contributed to the attack and accordingly both of their probabilities hiked.

Let us now consider what happens if, instead of Bayesian Conditioning, we employ the distance based approach outlined above. There are eight worlds here that we denote as APS, $AP\overline{S} \ldots \overline{APS}$ along expected lines. The A-worlds here represent the *accepted set* S and $\overline{A}$-worlds the *removal set* R. We use *Hamming Distance* as the pseudo-distance function d between different worlds.

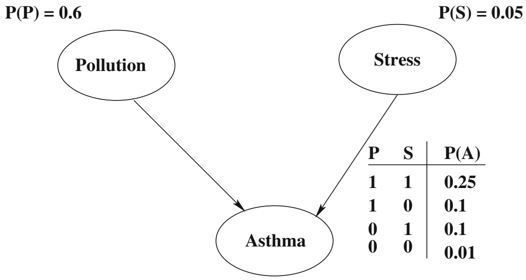

Fig. 1. A simple Bayesian Network depicting pollution and stress as the causes of asthma, with conditional probability tables (CPTs) given for each node. Pollution, stress and asthma are represented by P, S and A respectively.

Table 1 provides the distance between different worlds,[3] the weight of different A-worlds ω with respect to different $\overline{A}$-worlds $\omega^{\times}$,[4] and accordingly the share of ω in the probability of $\omega^{\times}$. The original probability of different $\overline{A}$-worlds is conveniently given beside their names in parentheses. For instance, the entry on top left says that $d(APS, \overline{A}PS) = 1$, since the total distance from $\overline{A}PS$ to different A-worlds is 8, the weight of APS with respect to $\overline{A}PS$ is $\frac{1}{8}$, and hence the share APS will claim in the probability of $\overline{A}PS$ (0.0225) is $\frac{0.0225}{8} \approx 0.0028$.

Table 1. Hamming distance, weight and share of an A-world with respect to different $\overline{A}$-worlds.

	$\overline{A}PS$ (0.0225)	$\overline{A}P\overline{S}$ (0.513)	$\overline{A}\overline{P}S$ (0.018)	$\overline{A}\overline{P}\overline{S}$ (0.3762)
APS	(1, 1/8, 0.0028)	(2, 2/8, 0.1282)	(2, 2/8, 0.0045)	(3, 3/8, 0.141)
$AP\overline{S}$	(2, 2/8, 0.0056)	(1, 1/8, 0.0641)	(3, 3/8, 0.0066)	(2, 2/8, 0.094)
$A\overline{P}S$	(2, 2/8, 0.0056)	(3, 3/8, 0.1923)	(1, 1/8, 0.0022)	(2, 2/8, 0.094)
$A\overline{P}\overline{S}$	(3, 3/8, 0.0084)	(2, 2/8, 0.1282)	(2, 2/8, 0.0044)	(1, 1/8, 0.047)

Now, the total probability share that an A-world receives from $\overline{A}$-worlds can be computed by simply adding up the third figures for each entry in a row. For instance, the total share that the world APS receives is $0.0028 + 0.1282 + 0.0045 + 0.141 \approx 0.2766$. These are shown in the third column in Table 2. It is easily noted that among the $\overline{A}$-worlds, the worlds $\overline{A}P\overline{S}$ and $\overline{A}\overline{P}\overline{S}$ account for most of the initial probability, and consequently the two A-worlds APS and $A\overline{P}S$ benefit the most from the probabilities initially allotted to the $\overline{A}$-worlds since they are *both* relatively "far away" from *each* of $\overline{A}P\overline{S}$ and $\overline{A}\overline{P}\overline{S}$.

[3] The distance between different A-worlds, or between different $\overline{A}$-worlds is not shown since they will not be used in the calculation.

[4] The more distant/different a world is from a target world, the higher is its relative weight.

The result of topping up the old probabilities by these shares gives the new probability of the A-worlds, as shown in the fourth column. Now, to compute the posterior probability of *Pollution*, say $P''(P)$, following this distance based approach, we simply add up the new probabilities of APS and $AP\overline{S}$, and get $0.2841 + 0.2275 \equiv 0.51$.[5] Similarly, the posterior $P''(S) = 0.2841 + 0.2963 \equiv 0.58$. In other words, stress is the more likely cause of the asthma attack than pollution. This is rather unexpected! Pollution and stress are equally effective in causing an asthma attack, the chance of pollution is very high, and the presence stress was assessed to be unlikely. And yet, when an asthma attack happens we conclude that stress was the likely cause of the attack. This is akin to believing in miracles.

Table 2. Old probabilities of A-worlds, their total probability share from $\overline{A}$-worlds, and the new probabilities.

	Old probability	Share from $\overline{A}$	New probability
APS	0.0075	0.2766	0.2841
$AP\overline{S}$	0.0570	0.1705	0.2275
$A\overline{P}S$	0.0020	0.2943	0.2963
$A\overline{PS}$	0.0038	0.1882	0.1920

So we consider closeness rather than distance as the basis for probabilistic reasoning instead.

3 From Distance to Closeness

In this section we will develop an account of probabilistic belief change that uses a closeness measure between worlds. Intuitively, the smaller the distance between two worlds is, the closer they are, and this measure will exploit this conviction. In many ways it will mimic the approach in [13] but trail imaging by moving the probability of a "discredited" world to similar worlds. This process makes use of the pseudo-distance function but dispenses with the unique closest world assumption, and in that it generalises the approach developed in [3].

We denote by $c(\omega, \omega')$ the closeness or similarity between two worlds ω and ω', and let the minimum closeness between two worlds to be 0 and the maximum to be 1. Intuitively, $c(\omega, \omega')$ attains the value 0 when ω and ω' are at farthest distance from one another. On the other hand, it is standard practice in the modal semantics to assume that each world is most similar to itself. This is indeed indicated by the fact that $d(\omega, \omega) = 0$ in case of pseudo-distance. Hence we would want that the closeness between any world and itself should be maximum

[5] We need not worry about adding the new probabilities of relevant $\overline{A}$-worlds since they are all zero.

and equal to 1, that is $c(\omega, \omega) = 1$. We will also assume that the closeness or similarity is symmetric, that is $c(\omega, \omega') = c(\omega', \omega)$ for all worlds ω and ω'.[6]

These properties are very similar to those of a pseudo-distance function. Where these two functions diverge is the *Triangle Inequality* which is satisfied by the pseudo-distance function but not satisfied by closeness function. It is easily seen that this is not a desirable property for closeness function. For instance, consider the major cities in Australia. Sydney is not very close to Perth, say the $c(Sydney, Perth) \approx 0$, and hence by *Symmetry* we also have $c(Perth, Sydney) \approx 0$. On the other hand, by *Identity* we have $c(Sydney, Sydney) = 1$. However, *Triangle Inequality* mandates that $c(Sydney, Perth) + c(Perth, Sydney) \approx 0 \geq c(Sydney, Sydney) = 1$.

In order to limit the range of $c(\omega, \omega')$ to $[0, 1]$, apart from the pseudo distance, we will need the maximum distance between any two worlds in any set X that we call its diameter $\Delta(X)$.

Definition 3.1. $\Delta(X) = max\{d(\omega, \omega') \mid \omega, \omega' \in X\}$ *for any set* X *of worlds.*

Closeness, being the complementary concept of distance, it is natural to assume that the closeness between two worlds would correspond to the gap between the distance between them and the maximum distance possible between any two worlds. We define the closeness function $c : \Omega \times \Omega \to [0, 1]$, parametrised to a relevant set $X \subseteq \Omega$ as:

Definition 3.2. $c_X(\omega, \omega') = \dfrac{\Delta(X) - d(\omega, \omega')}{\Delta(X)}$ *for* ω *and* $\omega' \in X \subseteq \Omega$.

Note that when we set X to be Ω in this definition, we get the absolute closeness between two worlds, and in that case we drop the subscript Ω in $c_\Omega(\cdot, \cdot)$. In Fig. 2 below we graphically illustrate the notion of this parametrised closeness (or comparative similarity). Two sets of worlds are represented as two spheres with diameters X and X'. We want to capture the comparative similarity between two pairs of worlds, (a, b) in the first set and (a', b') in the second. The actual distances between the two pairs of worlds is given by $d(a, b) = x$ and $d(a', b') = x'$. The "raw similarity" between the pairs is given by y and y'. The corresponding degrees of comparative closeness are given by $\frac{y}{x+y}$ and $\frac{y'}{x'+y'}$. Assuming $\frac{y}{x+y} < \frac{y'}{x'+y'}$, the worlds a' and b' are comparatively closer to each other in comparison to the worlds a and b.

It is easily shown that the closeness measure as defined in Definition 3.2 has the desirable properties we discussed earlier.

Observation 1. *The closeness function* c *(appropriately parametrised to a set* X*) satisfies the following conditions:*

1. $0 \leq c(\omega, \omega) \leq 1$ *(Range)*

[6] Arguably the use of similarity in common parlance is non-symmetric. For instance, if John is non-violent, we would say *John is like Gandhi*. But saying *Gandhi is like John* would mean a very different thing. Capturing such asymmetry in our simple framework may not be quite feasible.

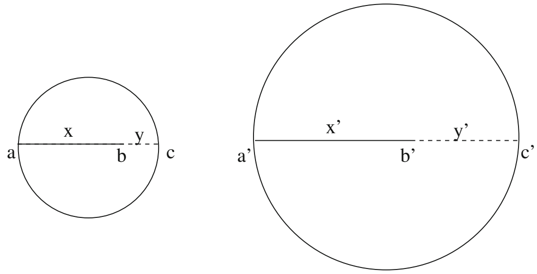

Fig. 2. Worlds a' and b' are closer to each other than a and b are since $\frac{y'}{x'+y'} > \frac{y}{x+y}$.

2. $c(\omega, \omega) = 1$ *(Identity)*
3. $c(\omega, \omega') = c(\omega', \omega)$ *(Symmetry)*

Now, let us see how we can distribute the probability of a world $\omega^\times \in R$ among worlds $\omega \in S$ with $R, S \subseteq \Omega$. One way would be to use the absolute closeness $c(\omega^\times, \omega)$ for different $\omega \in S$ and then compute the weight of different worlds in S with respect to $\omega^\times$ based on it analogous to the approach in [13]. This would entail that *all* worlds in S except the farthest ones will receive some non-zero share from the probability of $\omega^\times$. It would appear rather ad hoc since, if every other world in S benefits from the probability of $\omega^\times$, there is no reason why the least close ones should be deprived of this benefit. Hence we suggest that only those worlds in S that are in the "neighbourhood" of $\omega^\times$ should receive a portion of the latter's probability. We view *neighbourhood* as a very flexible concept – ranging from a very tight neighbourhood encompassing only those worlds in S closest to $\omega^\times$, to a very liberal one that includes almost the whole of S – and it can be adapted to suit the needs as necessary. (It is easily seen that classical imaging can be thus modeled by suitably choosing the distance function and required closeness of neighbours.) Here we provide only a particular interpretation of neighbourhood via what we call *mean proximity* below. Also, we use absolute closeness for convenience since it will not make any difference at the end. The mean proximity of a world $\omega^\times \in R$ with respect to set S is given as:

Definition 3.3. *Mean proximity of* $\omega^\times$ *wrt* S*:* $\pi(\omega^\times, S) = \frac{\sum_{\omega' \in S} c(\omega^\times, \omega')}{|S|}$

Intuitively, mean proximity gives us the boundary of the neighbourhood of $\omega^\times$ in which its probability will be distributed. We define the neighbourhood of a world $\omega^\times$ in a set S, denoted by $\nu(\omega^\times, S)$, as those worlds in S that are at most mean-proximity away from $\omega^\times$.

Definition 3.4. $\nu(\omega^\times, S) = \{\omega \in S \mid c(\omega^\times, \omega) \geq \pi(\omega^\times, S)\}$

Now we define the closeness-based weights δ^{cl} of the worlds in the neighbourhood of $\omega^\times$ along the expected lines. The δ^{cl} of ω to $\omega^\times$ is its closeness to $\omega^\times$ appropriately normalised. The closeness-weight of ω wrt $\omega^\times$ is given by:

Definition 3.5. *For all* $\omega^\times \in R, \omega' \in \nu(\omega^\times, S)$,

$$\delta^{cl}(\omega^\times, \omega, S) = \begin{cases} \dfrac{c(\omega^\times, \omega)}{\sum_{\omega' \in \nu(\omega^\times, S)} c(\omega^\times, \omega')} & \textit{if} \quad \omega \in \nu(\omega^\times, S) \\ 0 & \textit{otherwise.} \end{cases}$$

Now, analogous to Definition 2.4, we use the weights $\delta^{cl}(\omega^\times, \omega, S)$ to compute the total share of any $\omega \in S$ in the overall probability salvaged from R, denoted $\sigma^{cl}(\omega, S, R)$:

Definition 3.6. $\sigma^{cl}(\omega, S, R) = \sum_{\omega^\times \in R} P(\omega^\times) * \delta^{cl}(\omega^\times, \omega, S)$.

Subsequently, we use the operation $\langle prem^{cl} \rangle : B \times 2^{\Omega} \to B$ to determine the result of modifying a belief set b by new evidence Ψ through supplementing the existing probability of each world in $\omega \in S$ by its share $\sigma^{cl}(\omega, S, R)$:

Definition 3.7. $b\langle prem^{cl} \rangle R = \bigcup_{\omega \in S} \{(\omega, p'') \mid (\omega, p) \in b,\ p'' = p + \sigma^{cl}(\omega, S, R)\}$.

The closeness based probabilistic belief expansion operation $\langle prem^{cl} \rangle$ described above may be operationalised by the algorithm displayed below.

Algorithm: Closeness based Expansion

Input:b: belief-state, R: set of worlds that conflict with evidence
Output: new belief-state b'; R has total probability 0

foreach $\omega^\times \in R$ **do**
  foreach $\omega \in S$ **do**
    Calculate closeness c(ω, ω')
  endfor
  Calculate mean proximity $\pi(\omega^\times, S)$
  Determine neighborhood $\nu(\omega^\times, S)$
  foreach $\omega' \in \nu(\omega^\times, S)$ **do**
    $\delta^{cl}(\omega^\times, \omega', S) \leftarrow \dfrac{c(\omega^\times, \omega')}{\sum_{\omega' \in \nu(\omega^\times, S)} c(\omega^\times, \omega')}$
    $\mathrm{p}_{new} \leftarrow p_{old} + p^\times * \delta^{cl}(\omega^\times, \omega, S)$
  endfor
endfor

Let us now see how our proposal fares *vis-a-vis* Example 2. As before, we use the Hamming distance displayed in Table 1 as the pseudo-distance. As the parameter X for computing closeness we will use the space Ω, and hence, for this purpose, use as diameter $\Delta(\Omega) = 3$. Closeness function is availed to calculate the closeness between pairs of worlds as shown in Table 3. For instance,

$$c(APS, \overline{A}PS) = \frac{\Delta(\Omega) - d(APS, \overline{A}PS)}{\Delta(\Omega)} = \frac{3-1}{3} \approx 0.67.$$

Table 3. Closeness between A-worlds and $\overline{A}$-worlds.

	$\overline{A}PS$	$\overline{A}P\overline{S}$	$\overline{AP}S$	$\overline{APS}$
APS	0.67	0.33	0.33	0
$AP\overline{S}$	0.33	0.67	0	0.33
$A\overline{P}S$	0.33	0	0.67	0.33
$A\overline{PS}$	0	0.33	0.33	0.67

Now, when we learn A: *the child has had an asthma attack*, the $\overline{A}$-worlds will lose their probabilities which would be distributed in their respective neighbourhoods. We provide in Table 4 below for each $\overline{A}$-world its mean-proximity to the set of A-worlds, its neighbourhood among A-worlds, and the respective weight of each such neighbour. As in the case of distance based approach, we can compute the total probability share that an A-world receives from $\overline{A}$-worlds by adding up the weighted probabilities of each $\overline{A}$-world in whose neighbourhood it resides. For instance, APS is in the neighbourhoods of three $\overline{A}$-worlds: $\overline{A}PS$, $\overline{A}P\overline{S}$, and $\overline{AP}S$. Furthermore, its claim to the probabilities of these $\overline{A}$-worlds receives the respective weights of 0.5, 0.25 and 0.25 (the first entries in the last column for the three relevant rows of Table 4). The total share of APS in the probabilities of $\overline{A}$-worlds is then the weighted sum of their probabilities, that is, $0.0225 * 0.5 + 0.513 * 0.25 + 0.018 * 0.25 \approx 0.144$. These total shares that different A-worlds receive from $\overline{A}$-worlds are shown in the third column in Table 5. The result of topping up the A-worlds' old probabilities by these shares gives the new probability of the A-worlds, as shown in the fourth column in Table 5.

Table 4. Mean proximity and neighbourhood of an $\overline{A}$-world and the respective weights of its neighbours. Probabilities of the $\overline{A}$-worlds are also provided for convenience.

$\overline{A}$-worlds	Mean proximity	Neighborhood	Weights
$\overline{A}PS$ (0.0225)	0.33	$\{APS, AP\overline{S}, A\overline{P}S\}$	(0.5, 0.25, 0.25)
$\overline{A}P\overline{S}$ (0.513)	0.33	$\{APS, AP\overline{S}, A\overline{PS}\}$	(0.25, 0.5, 0.25)
$\overline{AP}S$ (0.018)	0.33	$\{APS, A\overline{P}S, A\overline{PS}\}$	(0.25, 0.5, 0.25)
$\overline{APS}$ (0.3762)	0.33	$\{AP\overline{S}, A\overline{P}S, A\overline{PS}\}$	(0.25, 0.25, 0.5)

Now, we compute the posterior probability of *Pollution*, say $P''_{cl}(P)$, following this closeness based approach by adding up the new probabilities of

APS and $AP\overline{S}$, and get $0.1515 + 0.4131 \equiv 0.564$. Similarly, the posterior $P''_{cl}(S) = 0.1515 + 0.1106 \equiv 0.262$. The observation of asthma attack has resulted in slight decrease in probability of pollution, and some increase in the probability of stress, and yet pollution remains the major causal contributor. This is more in alignment with our causal intuition.

Table 5. Old probabilities of A-worlds, their total probability share from $\overline{A}$-worlds, and the new probabilities using closeness based approach.

	Old probability	Share from $\overline{A}$	New probability
APS	0.0075	0.144	0.1515
$AP\overline{S}$	0.0570	0.3561	0.4131
$A\overline{P}S$	0.0020	0.1086	0.1106
$A\overline{P}\overline{S}$	0.0038	0.3208	0.3246

4 Discussion and Future Work

We looked at a simple scenario from the causal domain, described as Example 2 and sought to examine the propagation of probabilities when an effect is observed. The example was a case of multiple causes with a single effect. The two causes, stress and pollution, are equally efficacious as far as effecting an asthma attack is concerned. However, the pollution is a lot more prevalent than stress. When a child gets an asthma attack, which factor should we causally attribute it to?

Bayesian conditioning leads to results that are pleasing to the intuition. However since Bayesian conditioning cannot handle belief-contravening evidence we sought to explore unified methods that are more comprehensive.

A method advocated by Rens and Meyer [13] is one such more general approach. This approach exploits the distance between worlds to determine how the probabilities of the worlds "eliminated" by evidence should be distributed among those that "survive". It shifts the probability of a world to be eliminated to other worlds in the proportion of their distance from it. The rationale behind this approach is that if a world is considered implausible, then the worlds similar to it should also be considered implausible as well. The major flaw inherent in this method is that, as illustrated in Example 2, it produces counter-intuitive results. Given multiple causes of an effect, if the causes have equal efficacy, intuition demands that updating in light of the effect should not reverse the profile of the causal priors – if the prior of one cause dominates the prior of the other, their posteriors should exhibit that trend as well. In Example 2, both pollution and stress are taken to be equally good in effecting asthma attack, and prior of pollution is a lot higher than the prior of stress. However, the distanced based approach using Hamming distance shows a reversal of this trend – the posterior of stress is higher than the posterior of pollution.

The approach developed in this paper is based on closeness instead of distance. It may be considered to be a generalisation of imaging [9], as well as of an approach proposed in [3], in that it dispenses with the *unique closest world assumption* and distributes the probability of an eliminated world ω among those not eliminated roughly based on their similarity to ω. It is shown that when applied to Example 2, the posteriors of the causes do not reverse the direction of their priors. This indicates that the intuition behind imaging is basically sound, and probabilities of worlds eliminated by observed evidence should be moved to worlds similar to them, not to worlds dissimilar to them.

This paper is based on our preliminary exploration to probabilistic belief change based on closeness. There are a number of issues to be addressed down the road:

1. In this paper the closeness measure we employed is based on Hamming distance. The properties of Hamming distance that primarily contributed to the desirability of the outcome need to be formally captured.
2. This paper is based on a single, simple example. It does not show if the approach will work in other examples. We have tried it with a few other examples, and when there is substantial conflict between the distance based approach and the closeness based approach, the result of the latter appeared more intuitive. That, however, is no substitute for a formal proof.
3. The scope of this paper was restricted to belief non-contravening evidence. It will be interesting to see how the closeness based reasoning behaves when the evidence is belief-contravening.
4. It is likely that in different problem domains different approaches to probabilistic reasoning will be more appropriate. It will be fruitful to comprehensively explore this issue and compile the findings.
5. In the literature on logic of action there have been attempts to marry qualitative and quantitative approaches to causality and diagnosis (see, e.g., [2,6]). It will be interesting to see how our proposal sits in with such approaches.

We seek to address these issues in our future work.

References

1. Alchourrón, C.E., Gärdenfors, P., Makinson, D.: On the logic of theory change: partial meet contraction and revision functions. J. Symb. Log. **50**(2), 510–530 (1985)
2. Baral, C., Tran, N., Tuan, L.: Reasoning about actions in a probabilistic setting. In: Proceedings of the AAAI, pp. 507–512 (2002)
3. Chhogyal, K., Nayak, A., Schwitter, R., Sattar, A.: Probabilistic belief revision via imaging. In: Pham, D.-N., Park, S.-B. (eds.) PRICAI 2014. LNCS(LNAI), vol. 8862, pp. 694–707. Springer, Heidelberg (2014). doi:10.1007/978-3-319-13560-1_55
4. Gärdenfors, P.: Knowledge in Flux. MIT Press, Cambridge (1988)
5. Hansson, S.O.: Knowledge-level analysis of belief base operations. Artif. Intell. **82**(1–2), 215–235 (1996)

6. Iocchi, L., Lukasiewicz, T., Nardi, D., Rosati, R.: Reasoning about actions with sensing under qualitative and probabilistic uncertainty. ACM Trans. Comput. Log. 10(1) (2009). doi:10.1145/1459010.1459015
7. Katsuno, H., Mendelzon, A.O.: On the difference between updating a knowledge base and revising it. In: Proceedings of the KR, pp. 387–394 (1991)
8. Lehmann, D.J., Magidor, M., Schlechta, K.: Distance semantics for belief revision. J. Symb. Log. **66**(1), 295–317 (2001)
9. Lewis, D.: Probabilities of conditionals and conditional probabilities. Philos. Rev. **85**(3), 297–315 (1976)
10. Nayak, A.C.: Iterated belief change based on epistemic entrenchment. Erkenntnis **41**, 353–390 (1994)
11. Nayak, A., Goebel, R., Orgun, M., Pham, T.: Taking Levi Identity seriously: a plea for iterated belief contraction. In: Lehner, F., Fteimi, N. (eds.) KSEM 2016. LNCS(LNAI), vol. 9983, pp. 305–317. Springer, Heidelberg (2006). doi:10.1007/11811220_26
12. Pearl, J.: Probabilistic Reasoning in Intelligent Systems - Networks of Plausible Inference. Morgan Kaufmann, Burlington (1989)
13. Rens, G.B., Meyer, T.: A new approach to probabilistic belief change. In: Proceedings of the FLAIRS, pp. 581–587 (2015)