Why Min-Based Conditioning

Salem Benferhat[1] and Vladik Kreinovich[2]
[1] *CRIL (Centre de Recherche en Informatique de Lens), CNRS – UMR 8188*
Université d'Artois, Faculté des sciences Jean Perrin
Rue Jean Souvraz, SP 18, F62307 Lens Cedex, France
benferhat@cril.univ-artois.fr
[2] *Department of Computer Science, University of Texas at El Paso*
500 W. University, El Paso, Texas 79968, USA
vladik@utep.edu

Abstract. In many practical situations, we do not have full information about which alternatives are possible and which are not. In such situations, an expert can estimate, for each alternative, the degree to which this alternative is possible. Sometimes, experts can produce numerical estimates of their degrees, but often, they can only provide us with qualitative estimates: they inform us which degrees are higher, but do not provide us with numerical values for these degrees. After we get these degrees from the experts, we often gain additional information, because of which some alternatives which were previously considered possible are now excluded. To take this new information into account, we need to appropriately update the corresponding possibility degrees. In this paper, we prove that under several reasonable requirements on such an update procedure, there is only one procedure that satisfies all these requirements – namely, the min-based conditioning.

Keywords: imprecise knowledge, possibility distribution, conditioning, knowledge update, invariance

1. Formulation of the Problem

Need for ordinal-scale possibility degrees. It is often useful to describe, for each theoretically possible alternative ω from the set of all *theoretically* possible alternatives Ω, to what extent this alternative is, in the expert's opinion, *actually* possible.

Often, the only information that we can extract from experts is the qualitative one: which alternatives have a higher degree of possibility and which have lower degree. In some cases, we have a linear order between possible degrees, so all we know is the order of different alternatives, from the least possible to the most possible.

In principle, we could just use this order to process this information, but computers have been designed to process numbers – and they are still much better in processing numbers. So, to speed up processing of this data, degrees of possibility are usually described by numbers $\pi(\omega)$ from the interval $[0, 1]$: the higher the degree of possibility of an alternative ω, the larger the value $\pi(\omega)$.

These numbers by themselves do not have an exact meaning, the only meaning is in the order. So, in principle, the same meaning can be described if we apply any strictly increasing transformation to the interval $[0, 1]$.

Usually, some of this freedom is eliminated by the convention that the largest degree of possibility is set to 1; we can always achieve this with an appropriate transformation. Such possibility degrees are known as *normalized*. Thus, we arrive at the following definition (see, e.g., (Dubois, Lang, and Prade, 1994; Dubois, Moral, and Prade, 1998; Dubois and Prade, 1998; Gutierrez et al., 2014)):

Definition 1. *Let Ω be a finite Universe of discourse. A* possibility distribution *is a function $\pi : \Omega \to [0, 1]$ for which*

$$\max_{\omega \in \Omega} \pi(\omega) = 1.$$

Need for conditioning and normalization. Often, after we have learned the possibility degrees $\pi(\omega)$, we acquire an additional information, that some of the alternatives that we originally thought to be possible are actually not possible.

For example, we had originally the set Ω of possible suspects, but it turned out that some of the original suspects have alibis. As a result, the new list of suspects Ψ is smaller: $\Psi \subset \Omega$.

Let us now assume that for each of the suspects ω we had a numerical degree of this possibility $\pi(\omega)$ that this person committed the crime. Then, once we know that the list of suspects has narrowed, how to we change these possibility degrees?

In general, if we had a possibility distribution on a set Ω of possible alternatives, and we now learned that only alternatives from a subset $\Psi \subset \Omega$ are possible, how does this new information change our possibility degrees?

Of course, now that we learn that only alternatives from the smaller set Ψ are possible, we should set $\pi'(\omega) = 0$ for all $\omega \notin \Psi$. For all other alternatives $\sigma \in \Psi$, at first glance, it may sound reasonable to just retain the original possibility degrees, i.e., to take $\pi'(\omega) = \pi(\omega)$. However, we have an additional requirement, that the largest possibility degree should always be 1, and the above procedure does not always guarantee this requirement.

For example, if:

- we started with $\pi(a) = 0.4$, $\pi(b) = 0.5$, and $\pi(c) = 1.0$, and
- we learned that $\omega \in \Psi = \{a, b\}$,

then:

- if we simply take $\pi'(a) = 0.4$, $\pi'(b) = 0.5$, and $\pi'(c) = 0$,
- the largest of the resulting three degrees is not equal to 1.

It is therefore necessary to *normalize* the resulting degrees $\pi'(\omega)$, i.e., to transform them into new degrees for which the largest is 1. This transformation – that we will call a *conditioning operator* – should:

- take, as input, the original possibility distribution π, and
- generate the new probability distribution π'' that we will denote by $(\pi \mid \Psi)$.

According to Definition 1, a possibility distribution π is a function that assigns, to every alternative ω from the finite set Ω of possible alternatives, a value $\pi(\omega)$ from the interval $[0,1]$ – the value that describe the degree to which, according to the available information, the alternative ω is possible.

A function from a finite set is nothing else but a tuple. Thus, if we denote the elements of the finite set Ω by $\omega_1,\ldots,\omega_n$, then a possibility distribution is simply a tuple consisting of n values from the interval $[0,1]$: $\pi=(\pi(\omega_1),\ldots,\pi(\omega_n))$. In these terms, a conditioning operator is a mapping that takes one such tuple π as an input and returns a new tuple $\pi''=(\pi\mid\Psi)$ as the output.

To get π'', one possibility is to divide all the values $\pi'(\omega)$ by the largest of these values. In the above example, the largest value is 0.5, so we get $\pi''(a)=\dfrac{0.4}{0.5}=0.8$ and $\pi''(b)=\dfrac{0.5}{0.5}=1$.

Another possibility is to replace the largest of the values $\pi'(\omega)$ by 1 and leave all other values unchanged. For the above example, this would mean that we take $\pi''(a)=0.4$ and $\pi''(b)=1$.

Let us describe a general definition of the corresponding operator.

Definition 2. *By a* conditioning operator*, we mean a mapping $(\pi\mid\Psi)$ that:*

- *inputs a possibility distribution π on a set Ω and a non-empty set $\Psi\subseteq\Omega$, and*
- *returns a new possibility distribution for which $(\pi\mid\Psi)(\omega)=0$ for all $\omega\notin\Psi$.*

What are the reasonable conditioning operators?

2. Analysis of the Problem

Let us describe the desired properties of the conditioning operator.

First property: it should not matter how we previously judged alternatives that we now know to be impossible. A first reasonable requirement is that since alternatives $\omega\notin\Psi$ are excluded, their original possibility degrees should not affect the resulting degrees. In other words, if two original possibility distributions π and π' differ only by their values outside Ψ, then the conditioning should be the same.

C1. If $\pi_{|\Psi}=\pi'_{|\Psi}$, i.e., if $\pi(\omega)=\pi'(\omega)$ for all $\omega\in\Psi$, then

$$(\pi\mid\Psi)=(\pi'\mid\Psi).$$

Second property: order between possibility degree of different alternatives should not change. Another reasonable condition is that while the numerical values of possibility degrees may change, the order between these degrees should not change:

C2. If $\pi(\omega)<\pi(\omega')$ for some $\omega,\omega'\in\Psi$, then

$$(\pi\mid\Psi)(\omega)<(\pi\mid\Psi)(\omega').$$

C3. If $\pi(\omega) = \pi(\omega')$ for some $\omega, \omega' \in \Psi$, then

$$(\pi \mid \Psi)(\omega) = (\pi \mid \Psi)(\omega').$$

Third property: it should not matter whether we learn the new knowledge right away or in two steps. Often, we first learn some information, based on which the set of possible alternatives is limited to a subset $\Psi \subset \Omega$, and then learn some additional information according to which the set of possible alternatives is even smaller $\Psi' \subset \Psi$. In this case:

– we first condition the original degrees of possibility π with respect to Ψ, resulting in $\pi' = (\pi \mid \Psi)$, and then

– we condition π' with respect to the new set Ψ', resulting in $\pi'' = (\pi' \mid \Psi') = ((\pi \mid \Psi) \mid \Psi')$.

Alternative, we could learn both pieces of the information at the same time. In this situation, our reaction to this new information would replace the original possibility distribution π with $(\pi \mid \Psi')$.

In both cases, we gain the exact same new information, so the resulting changes in possibility degrees should be the same:

C4. If $\Psi' \subset \Psi$, then $((\pi \mid \Psi) \mid \Psi') = (\pi \mid \Psi')$.

Fourth property: alternatives which were originally considered impossible should not matter. Another condition is that if had an alternative ω_0 which we originally believed to be impossible – i.e., whose degree of possibility is 0 – then:

– this alternative should remain impossible after conditioning, and

– the possibility degrees of all other alternatives $\omega \neq \omega_0$ should remain the same, whether we keep the alternative ω_0 in the remaining set Ψ or whether we explicitly delete ω_0 from the set Ψ.

This property can be described as follows:

C5. If $\pi(\omega_0) = 0$ for some $\omega_0 \in \Psi$, then

$$(\pi \mid \Psi)(\omega_0) = 0 \text{ and } (\pi_{|\Psi - \{\omega_0\}} \mid \Psi) = (\pi \mid \Psi)_{|\Psi - \{\omega_0\}}.$$

Final property: invariance. As we have mentioned earlier, often, the only information that we can extract from the experts is which alternatives have a higher degree of possibility and which have a lower degree of possibility. In other words, what matters is the order between the degrees, not the numerical values of the degrees. So, the situations should not change if we simply apply some re-scaling that does not change the order (such as $x \to x^2$), i.e., if we apply some increasing one-to-one function $T : [0,1] \to [0,1]$ to transform each degree $\pi(\omega)$ into a degree $T(\pi(\omega))$. The corresponding re-scaled tuple will be denoted by $T\pi$: $(T\pi)(\omega) \stackrel{\text{def}}{=} T(\pi(\omega))$.

It is reasonable to require that the result of applying the conditioning operator not change if we apply such a re-scaling. In other words, the following two operations should leads to the exact same result:

– either we apply the conditioning in the original scale, i.e., transform π into $(\pi \mid \Psi)$, and then, apply the re-scaling T, resulting in $T(\pi \mid \Psi)$;

– or we first apply the re-scaling, resulting in $T\pi$, and then apply the conditioning, resulting in $(T\pi \mid \Psi)$.

Thus, we arrive at the following requirement:

C6. For every increasing one-to-one function $T : [0,1] \to [0,1]$, we have

$$(T\pi \mid \Psi) = T(\pi \mid \Psi).$$

Now, we are ready to formulate our main result.

3. Main Result

Proposition. *The only conditioning operator that satisfies the properties* **C1**–**C6** *is the min-based operator (*Dubois and Prade, 1984*;* Hisdal, 1978*) for which:*

- $(\pi \mid \Psi)(\omega) = 1$ *when* $\omega \in \Omega$ *and* $\pi(\omega) = \max\limits_{\omega' \in \Omega} \pi(\omega')$*;*
- $(\pi \mid \Psi)(\omega) = \pi(\omega)$ *when* $\omega \in \Omega$ *and* $\pi(\omega) < \max\limits_{\omega' \in \Omega} \pi(\omega')$*; and*
- $(\pi \mid \Psi)(\omega) = 0$ *when* $\omega \notin \Psi$*.*

Discussion. The usual derivation of the min-based conditioning (see, e.g., (Dubois, Lang, and Prade, 1994)) is to interpret the degree $(A \mid B)$ as the maximal value for which $A \,\&\, B$ (with min as "and"-operation) has the same truth value as $(A \mid B) \,\&\, B$.

Our result shows that *maximality* can be replaced with *invariance* – which reflects the ordinal-scale character of the corresponding possibility degrees.

Proof.

1. It is easy to show that the min-based operator satisfies the properties **C1–C6**.

To complete the proof, we need to prove that, vice versa,

– every conditioning operator that satisfies these five properties

– is indeed the min-based operator.

To prove this statement, we will consider two possible cases:

– the case when the set Ψ contains some alternative ω for which $\pi(\omega) = 1$, and

– the case when the set Ψ does not contain any alternative ω for which $\pi(\omega) = 1$.

2. Let us first consider the case when the set Ψ contains some alternative ω for which $\pi(\omega) = 1$. In this case, the min-based formula leads to $(\pi \mid \Psi)(\omega) = \pi(\omega)$ for all $\omega \in \Psi$.

Let us show that this equality holds for all conditioning operators that satisfy the properties **C1**–**C6**.

2.1. If there is no $\omega_0 \in \Psi$ for which $\pi(\omega_0) = 0$, let us add such an element to our set Ω. According to Property **C5**, this will not change the result. Thus, without losing generality, we can safely assume that there is an element $\omega_0 \in \Psi$ for which $\pi(\omega_0) = 0$.

As for the values $\pi(\omega)$ for $\omega \notin \Psi$, we can use the property **C1** to replace them with zeros.

2.2. Let us sort values $\psi(\omega)$ corresponding to different alternatives $\omega \in \Psi$ in increasing order. We know that the resulting list of values includes 0 and 1, so this list has the form

$$v_1 = 0 < v_2 < \ldots < v_{k-1} < v_k = 1,$$

where k is the number of different values $\pi(\omega)$ corresponding to $\omega \in \Psi$.

Let us use property **C6** to prove that the values $(\pi \mid \Psi)$ should also be from this list. Indeed, let us consider the following strictly increasing function $T(v)$: for $v_i \le v \le v_{i+1}$, we take

$$T(v) = v_i + \left(\frac{v - v_i}{v_{i+1} - v_i}\right)^2 \cdot (v_{i+1} - v_i).$$

One can easily check that for this function, $T(v_i) = v_i$ for all i, so $T(\pi) = \pi$. Thus, the property **C6** implies that $T(\pi \mid \Psi) = (\pi \mid \Psi)$, i.e., that for each value $v = (\pi \mid \Psi)(\omega)$, we should have $T(v) = v$. But for the above function $T(v)$, the only such values are $v_1, \ldots, v_k$.

So, indeed, the values $v_1 < \ldots < v_k$ are mapped to the same k values. By properties **C2** and **C3**, equal values of $\pi(\omega)$ are mapped into equal values of $(\pi \mid \Psi)(\omega)$, and smaller values of $\pi(\omega)$ are mapped into smaller values of $(\pi \mid \Psi)(\omega)$. Thus, the values v'_i corresponding to v_i are also sorted in increasing order: $v'_1 < \ldots < v'_k$. Each new value v'_i must coincide with one of the original values v_j. So, in the increasing list $v_1 < \ldots < v_k$ of k values, we have k new values v'_i which have the same order. This implies that v'_1 must be the smallest of v_i, i.e., $v'_1 = v_1$, that v'_2 be the second smallest, i.e., $v'_2 = v_2$, and, in general, $v'_i = v_i$, i.e., indeed, $(\pi \mid \Psi)(\omega) = \pi(\omega)$ for all $\omega \in \Psi$.

3. Let us now consider the case when the set Ψ does not contain some alternative ω for which $\pi(\omega) = 1$.

In this case, we can also add (if needed) an element ω_0 for which $\pi(\omega_0) = 0$, and sort the values $\pi(\omega)$ corresponding to $\omega \in \Psi$ into an increasing sequence $v_1 = 0 < v_2 < \ldots < v_k < 1$; the only difference is that in this case, the largest value v_k in this increasing sequence is smaller than 1.

One of the new values should be equal to 1. So, due to Properties **C2** and **C3**, only the largest degree v_k should be mapped into 1.

Similarly to Part 2 of this proof, we can prove that each of the the values $v_1, \ldots, v_{k-1}$ maps into one of the values $v_1, \ldots, v_k$, and that if $v_i < v_j$, then $v'_i < v'_j$. By induction, we can prove that $v'_i \ge v_i$. Since we have only one additional value to move to, for each i, we have either $c'_i = v_i$ or $v'_i = v_{i+1}$. In other words, for each alternative, after conditioning, we will have either the same degree of possibility as before, or the next one.

Let use the Property **C4** to prove, by contradiction, that a value $v_i < v_k$ cannot be transformed into the next value v_{i+1}. Let us assume that, vice versa, there is an element $\omega_i \in \Psi$ for which $\pi(\omega_i) = v_i$ and $(\pi \mid \Omega)(\omega_i) = v_{i+1}$. To get a contradiction, let us consider the new set $\Omega^* = \Omega \cup \{\omega^*\}$, with a new element ω^*, and a new possibility distribution π^* for which:

- we have $v_i < \pi^*(\omega^*) < v_{i+1}$ and
- we have $\pi^*(\omega) = \pi(\omega)$ for all $\omega \neq \omega_i$.

Let us consider two conditionining paths from Ω^* to Ψ:

- in the first path, we go from Ω^* to Ω and then from Ω to Ψ;
- in the second path, we go from Ω^* to $\Psi^* \stackrel{\text{def}}{=} \Psi \cup \{\omega^*\}$ and then from Ψ^* to Ψ.

According to the Property **C4**, the resulting value $(\pi^* \mid \Psi)(\omega_i)$ should be the same for both paths.

In the first path, first, we go from Ω^* to Ω. This transition eliminates a single element ω^* for which $\pi^*(\omega^*) < 1$. Thus, according to Part 2 of this proof, all the possibility degrees of remaining elements remain unchanged. Thus, we have $(\pi^* \mid \Omega) = \pi$. We already know that $(\pi \mid \Psi)(\omega_i) = v_{i+1}$. Thus, due to Property **C4**, we have

$$(\pi^* \mid \Psi)(\omega_i) = ((\pi^* \mid \Omega) \mid \Psi)(\omega_i) = v_{i+1}.$$

On the other hand, in the second path, we first move from Ω^* to Ψ^*. In this transition, we have v_k transformed into 1, and the original value $\pi^*(\omega_i) = v_i$ can either remain the same, or it can be transformed to the next value which is now $\pi^*(\omega^*) < v_{i+1}$. In both cases, the new possibility degree is smaller than v_{i+1}: $\pi(\omega_i) < v_{i+1}$. When we then reduce the set Ψ^* to Ψ, then all the alternatives for which we originally had $\pi^*(\omega) = \pi(\omega) = v_k$ and now have $\pi'(\omega) = 1$ remain in the set. Thus, all other alternatives – including the alternative ω_i – according to Part 2 of this proof, retain their values. For ω_i, this implies that $(\pi' \mid \Psi)(\omega_i) = \pi'(\omega_i) < v_{i+1}$. Thus, we have $(\pi^* \mid \Psi)(\omega_i) = \pi'(\omega_i) < v_{i+1}$, which contradicts to the above equality $(\pi^* \mid \Psi)(\omega_i) = v_{i+1}$.

This contradiction shows that the transformation from v_i to v_{i+1} is indeed impossible. Thus, we have $v'_i = v_i$.

So, indeed, we have derive the min-based conditioning from the properties **C1**–**C6**. The proposition is proven.

Acknowledgements

This work was supported in part by the US National Science Foundation grants HRD-0734825, HRD-1242122, and DUE-0926721.

This work was performed when Salem Benferhat was visiting El Paso.

The authors are thankful to the anonymous referees for valuable suggestions.

References

Dubois, D., J. Lang, and H. Prade, Possibilistic logic, In: D. Gabbay, C. J. Hogger, and J. A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming. Vol. 3. Nonmonotonic Reasoning and Uncertain Reasoning*, pages 439–513, Oxford University Press, New York, 1994.

Dubois, D., S. Moral, and H. Prade, Belief change rules in ordinal and numerical uncertainty theories, In: D. Gabbay and P. Smets (eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 1. Quantified Representation of Uncertainty and Imprecision*, pages 311–392, Kluwer Academic Publ., Dordrecht, 1998.

Dubois, D., and H. Prade, Fuzzy logics and the generalized modus ponens revisited, *Cybernetics and Systems*, 15:293–331, 1984.

Dubois, D., and H. Prade, Possibility theory: qualitative and quantitative aspects. In: D. Gabbay and P. Smets (eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 1. Quantified Representation of Uncertainty and Imprecision*, pages 169–226, Kluwer Academic Publ. Dordrecht, 1998.

Gutierrez, L., M. Ceberio, V. Kreinovich, R. L. Gruver, M. Peña, M. J. Rister, A. Saldana, J. Vasquez, J. Ybarra, and S. Benferhat, From interval-valued probabilities to interval-valued possibilities: case studies of interval computation under constraints, *Proceedings of the 6th International Workshop on Reliable Engineering Computing REC'2014*, Chicago, Illinois, May 25–28, 2014, pp. 77–95.

Hisdal, E., Conditional possibilities: independence and non-interaction, *Fuzzy Sets and Systems*, 1:283–297, 1978.