Development of Possibilistic Causal Model from Data

Koichi YAMADA
Department of Management and Information Systems Science
Nagaoka University of Technology
1603-1, Kami-tomioka, Nagaoka, Niigata, 940-2188 , JAPAN

Abstract: - Possibilistic causal models have been proposed as an approach for prediction and diagnosis based on uncertain causal relations. However, the only way to develop the causal models is to acquire the possibilistic knowledge from the experts. The paper proposes an approach to develop the models from a dataset including causes and effects. It first develops a probabilistic causal model, then transform it into a possibilistic one. The points which should be discussed in the approach are 1) the way to transform multiple probabilistic distributions consistently into possibilistic ones, and 2) the merits of the transformation from a probabilistic model into a possibilistic one.

Key-Words: - Causal model, possibility theory, probability-possibility transformation, machine learning.

1 Introduction
Uncertainty of a causal relation is one in the situation where an event is literally caused by another. Peng and Reggia studied a diagnostic reasoning with the uncertainty called conditional causal probabilities [1]. It was also studied in psychology and called causal power [2]. There are also some studies which use the Possibility theory instead of probability [3,4].

Both probability and possibility are scales of uncertainty. The differences between them are mathematically clear based on the Evidence theory [5]. It is also indicated that possibility is a qualitative, relative and ordinal scale of uncertainty, while probability is a ratio scale that has a quantitative and absolute nature [6,7,8]. However, the practical merits of using possibility instead of probability are not confirmed except the cases where possibility is used as membership functions of fuzzy sets.

Limiting the discussion to possibilistic causal models, the only way to develop them is to acquire the possibilistic knowledge from experts. However, acquiring possibilistic knowledge is not an easy task, because there is no standard gauge to measure possibility unlike probability based on frequency.

An alternative is the way to use transformation from probability into possibility, after a probabilistic model is developed. However, current transformation techniques only deal with transformation of a distribution [8-11]. When the model consists of multiple distributions, it must be transformed with consistency among the distributions. The reason why attention must be paid to the consistency is the relative nature of possibility distributions. In addition, practical merits of using possibility must be clarified, because probability is replaced by possibility.

The paper proposes an approach to develop a possibilistic causal model from a dataset including causes and effects. It first develops a probabilistic causal model from the data [12], then transform it into a possibilistic one. The important contributions of the paper are 1) to propose the way to transform multiple probabilistic distributions consistently into possibilistic ones, and 2) to show the practical merits of transformation of the model.

2 Causal Models
2.1 Causal Event [1]
Let $U = \{u_i \mid i = 1, \ldots, N\}$ and $V = \{v_j \mid j = 1, \ldots, M\}$ be two disjoint sets of primitive events. The events in the both sets are not necessarily exclusive nor exhaustive in the paper. The occurrence of $u_i$ and $v_j$ is represented by $u_i^+$ and $v_j^+$, respectively. $u_i^-$ and $v_j^-$ represent absence. Any $v_j$ is never present without being caused by one of $u_i$. Thus, $u_i$ and $v_j$ are called causes and effects, respectively.

$c_{ij}$ defined by the next is called causation event.

$c_{ij}^+ \rightarrow u_i^+ \land v_j^+,$

(1)

where $\rightarrow$ is an implication and $\land$ is a conjunction.

$c_{ij}$ represents an event that $u_i$ is present and that $u_i$ causes $v_j$. It is easily shown that the following hold.
Proceedings of the 10th WSEAS International Conference on SYSTEMS, Vouliagmeni, Athens, Greece, July 10-12, 2006 (pp336-341)

\[ c_j^x \leftrightarrow c_j^{x'} \land u_i^x \land v_j^x, \]  
\[ c_j^y \land u_i^y \leftrightarrow u_i^y, \]  
\[ c_j^x \land v_j^y \leftrightarrow v_j^y, \]

where \( \leftrightarrow \) is an equivalence.

Then, the following formula is assumed.

\[ v_j^y \leftrightarrow \bigvee_i^c c_j^{u_i^y}, \]

where \( \bigvee \) is a disjunction.

2.2 Probabilistic Causal Model [1]

Let \( x_i, y_j, \) and \( z_j \) be generic values in \( U_i = \{u_i^x, u_i^y\}, \)
\( V_j = \{v_j^x, v_j^y\} \) and \( C_{ij} = \{c_{ij}^x, c_{ij}^y\}, \) respectively.
\( \{P(x_i) \mid x_i \in U_i\}, \{P(y_j) \mid y_j \in V_j\} \) and \( \{P(z_j) \mid z_j \in C_{ij}\} \) are probability distributions on \( U_i, V_j \) and \( C_{ij}. \)

\( \PrCM \) is called conditional causal probabilities. Eq. (2) and (3) derive \( \PrCM(u_i^x) = 0 \) and \( \PrCM(u_i^y) = 1. \)

Let \( \Delta_j \) be a conjunction of arbitrarily chosen \( x_i (i \neq j) \) and \( z_j (j \neq i) \). If eq. (6) holds, \( z_j \) is called (probabilistically) causation independent.

\[ P(z_j \mid x_i) = P(z_j \mid x_i \land \Delta_j), \]  
where \( x_i \land \Delta_j \) must be consistent.

Probabilistic causal model (PrCM) is defined by

\[ \PrCM = (U, V, P^m, P^c), \]

where \( P^m = \{P(u_i^x) \mid i = 1, ..., N\} \) and \( P^c = \{P(c_{ij}^x \mid u_i^x) \mid i = 1, ..., N, j = 1, ..., M\}. \)

If eq. (6) holds, \( z_j \) is called (probabilistically) causation independent.

In this case, the next equations hold [1].

\[ P(c_{ij}^x \mid u_i^x) = P(v_j^y \mid u_i^x \land u_i^{y-1} \land u_i^{y+1} \land u_i^N) \]  
\[ P(v_j^y) = \prod_i (\neg P(u_i^x)P(c_{ij}^x \mid u_i^x) ) \]

2.3 Possibilistic Causal Model [3,4]

Similarly to section 2.2, possibility distributions on \( U_i, V_j \) and \( C_{ij} \) are given as \( \{\Pi(x_i) \mid x_i \in U_i\}, \)
\( \{\Pi(y_j) \mid y_j \in V_j\} \) and \( \{\Pi(z_j) \mid z_j \in C_{ij}\}. \)

\( \Pi(z_j \mid x_i) \) is called conditional causal possibilities. The paper employs the Hisdal's definition of conditional possibilities [13]. \( \Pi(c_{ij}^x \mid u_i^x) = 0 \) and \( \Pi(c_{ij}^y \mid u_i^y) = 1 \) hold. If eq. (10) holds, \( z_j \) is called (possibilistically) causation independent.

\[ \Pi(z_j \mid x_i) = \Pi(z_j \mid x_i \land \Delta_j), \]

where \( x_i \land \Delta_j \) must be consistent.

Possibilistic causal model (PoCM) is defined by

\[ \text{PoCM} = (U, V, \Pi^m, \Pi^c). \]

\( \Pi^m = \{\{\Pi(u_i^x) \mid i = 1, N\} \) and \( \Pi^c = \{\{\Pi(c_{ij}^x \mid u_i^x) \mid i = 1, ..., N, j = 1, ..., M\}. \)

Since causes in \( U \) are not necessarily exhaustive, \( \max_i \Pi(u_i^y) \leq 1. \)

We assume the following:

i) \( x_i \) are possibilistically independent of one another.

ii) \( z_j \) are possibilistically causation independent.

In this case, the next equations hold [3,4],

\[ \Pi(c_{ij}^x \mid u_i^x) = \Pi(v_j^y \mid u_i^x \land u_i^{y-1} \land u_i^{y+1} \land u_i^N), \]  
\[ \Pi(v_j^y) = \Pi(c_{ij}^x \mid u_i^x) \land \Pi(u_i^x) \]

where \( \lor \) and \( \land \) denote max and min, respectively, when used in the calculation of possibilities.

3 Development of Probabilistic Model

3.1 Data and Model

Let \( D \) be the dataset, from which a \( PrCM \) is developed. \( U \) and \( V \) are given in advance. Each element of \( D \) is a record including causes and effects, and is represented by \( d_k = (x_1, ..., x_N, y_1, ..., y_M). \)

We consider two different situations where the data are collected; one is the situation where all of \( d_k \) have \( x_i = u_i^x \) at least one of \( i \) (in this case, causes in \( U \) are exhaustive, but still not exclusive), and the other is the situation where \( d_k \) may have \( x_i = u_i^x \) for all of \( i. \)

The former assumes that the data are collected in the condition (denoted by \( \Omega \)) that one or more causes are present. So, it could be called trouble model (\( \Omega \)-model), assuming that the causes are troubles. The latter does not assume the condition (denoted by \( G \)).

So, it could be called general model (G-model).

The paper deals with the following four cases.

a) Case-1: develop \( G \)-model from data collected in the condition \( G \).

b) Case-2: develop \( \Omega \)-model from data in \( G \).

c) Case-3: develop \( G \)-model from data in \( \Omega \).

d) Case-4: develop \( \Omega \)-model from data in \( \Omega \).

Let \( P_G(\bullet) \) and \( P_{\Omega}(\bullet) \) be probabilities in \( G \)-model and \( \Omega \)-model, respectively. From the definition, \( P_G(u_i^x) + ... + P_G(u_i^N) \geq 1 \) and \( P_{\Omega}(u_i^x) + ... + P_{\Omega}(u_i^N) \geq 1 \), because \( U \) is not exclusive.
In cases of 1 and 2, probabilities obtained by counting data are represented by \( P_G(\bullet) \). In cases of 3 and 4, they are \( P_D(\bullet) \). On the other hand, the developed model must be represented by \( P_G(\bullet) \) in cases of 1 and 3, and by \( P_D(\bullet) \) in cases of 2 and 4.

Now, the following is derived from the definitions of \( \Omega \) and the causation event.

\[
\Omega \leftrightarrow \bigvee_{i,j} u_i 
\leftrightarrow \left( \bigvee_{i,j} c_{ij} \right) \left( \bigvee_{i,j} v_i \right)
\]

Thus, the following are obtained.

\[
P_G(c_{ij} | u_i) = P_G(c_{ij} | \Omega) = P_G(c_{ij} | u_i).
\]

\[
P_D(v_i | u_i) = P_D(v_i | \Omega) = P_D(v_i | u_i).
\]

\[
P_G(u_i) = P_G(u_i | \Omega) = P_G(u_i | \nu_i).
\]

\[
P_D(v_j) = P_D(v_j | \Omega) = P_D(v_j | P_G(\Omega).
\]

From the definition of \( \Omega \), the next equation holds.

\[
P_G(\Omega) = 1 - P_G(u_i) \cdots P_G(u_n).
\]

3.2 Probabilities of Causal Model

In both cases of 1 and 4, probabilities of \( P^m \) (\( P_G(u_i) \)) in Case-1, \( P_D(u_i) \) in Case-4 are easily obtained by counting data. In Case-2, the probability \( P_D(u_i) \) is derived using eq. (17) and (19). In Case-3, the probability \( P_D(u_i) \) can be calculated numerically, using the way in Appendix.

For probabilities of \( P^c \), we do not have to care about the difference between \( P_G(\bullet) \) and \( P_D(\bullet) \), because \( P_G(c_{ij} | u_i) = P_G(c_{ij} | u_i) \). Thus, we use the expression \( P(\bullet) \) in the rest of this subsection.

The simplest way to calculate \( P(c_{ij} | u_i) \) is to use eq. (8). However the way is not realistic, because the combination of \( u_i | u_i \) \((i=1, \ldots, N)\) amounts to \( 2^N \) or \( 2^{N-1} \). Thus, the simple average number of data which could be used to calculate eq. (8) is just \( K / 2^N \).

The paper uses the next equation proposed in [12].

\[
P(c_{ij} | u_i) = \frac{P(v_j | u_i) - P(v_j) - P(u_i)P(v_j)}{1 - P(u_i) - P(v_j) + P(u_i)P(v_j) | u_i}.
\]

where \( P(u_i) \), \( P(v_j) \) and \( P(v_j | u_i) \) can be calculated by counting data in \( D \).

4 Development of Possibilistic Model

When there are an enough number of data to obtain reliable probabilities, the natural choice of uncertainty expression is probability. The \( PrCM \) developed in the previous section are reliable in the sense that probabilities (for example, \( P(v_j) \)) calculated with the model are in the right order, if the number of records is more than several hundreds [12]. However, in the case where the number is less than 100, the derived model is not reliable.

The paper proposes to transform a \( PrCM \) into a \( PoCM \) by transforming \( P^m \) and \( P^c \) into \( P^m \) and \( P^c \), respectively. Though it is impossible to obtain a reliable model from an unreliable one without additional information, the qualitative and ordinal nature of possibility would cover the ordinal errors which the probabilistic models make due to an insufficient number of data. Possibilistic models are expected to derive moderate results of reasoning, which means that they prevent unnecessary total ordering unavoidable for probability, a ratio scale of uncertainty.

4.1 Probability-Possibility Transformation

Many methods were proposed for the transformation between probability and possibility. The paper chooses one proposed in [10] as the start of discussion, because it is supported by another paper [8] from the view point of evidence theory.

Let \( E = \{ e_1, \ldots, e_n \} \) be an exclusive and exhaustive set of events unlike our case. The probability that \( e_i \) is present is represented by \( P(e_i) \), \( \{ P(e_i), \ldots, P(e_n) \} \) is a probability distribution. Then, possibility of \( e_i \) is given by the next equation.

\[
\Pi(e_i) = \sum_{k=1}^{n} (P(e_i) \land P(e_i)).
\]

\( \Pi(e_i) \) is given by the maximum of \( \Pi(e_i), (k \neq i) \).

4.2 Transformation of Causal Model

4.2.1 Problem of simple transformation

\( P^m \) and \( P^c \) are not probability distributions. Thus, the transformation by eq. (21) cannot be applied directly to them. Instead, it could be applied to \( \{ P(u_i), \ldots, P(u_i) \} \) and \( \{ P(c_{ij} | u_i), P(c_{ij} | u_i) \} \) . However, the simple transformation would cause a problem.

When \( PoCM \) is obtained by the transformation, the reasoning results by eq. (13) tend to be determined only by the values of \( \Pi(u_i) \), especially when it is a \( G \)-model and/or the model has many possible causes. In those cases, \( P(u_i) \) are usually very small, and so as \( \Pi(u_i, \ldots, u_i) = 2P(u_i) \) in this case). Thus, the results of eq. (13a) tend to be \( \Pi(v_j) = \max \Pi(u_i) \) for most of \( v_j \). In the case \( \Pi(v_j) = 1.0 \) is obtained.
The problem comes from the relative nature of possibility. Looking at eq. (21), the value of \( \Pi(e_i^j) \) is dependent not only on \( P(e_i^j) \) but also on probabilities of the other events and the number of events. The dependence on probabilities of the other events is also observed in other transformation methods between probability and possibility.

The relativity could also be explained by evidence theory [5]. That is, possibility is a plausibility measure where all of its focal elements are nested. The nest structure clearly gives a very restrictive relativity among the elements of the universal set.

On the other hand, probability is a measure where cardinality of every focal element is 1. That is, every piece of evidence is independent of one another. Thus, probability could be determined independently of probabilities of other events, for example only by its frequency.

As the results, probability is an absolute index of uncertainty connected to the concept of frequency, while possibility is a relative index of uncertainty affected by the other elements in the universal set. Then, we must come to a conclusion that there is no rationale to transform a model with multiple probability distributions defined on different universal sets to a possibilistic model by transforming each probability distribution into a possibility distribution separately. This is because possibility distributions derived in the way take into account only the relativity in their universal set, but not relativity among all events in all the universal sets.

The only sound approach derived from the above discussion seems the way to define the joint probability distribution on the cartesian product of all sets to a possibilistic model by transforming each possibility distribution into a possibility distribution. This is because possibility distributions derived in the way take into account only the relativity in their universal set, but not relativity among all events in all the universal sets.

4.2.2 Proposed Transformation

We start from the question of what we want to know from \( \Pi(v_i^j) \) and \( \Pi(v_j^i) \) calculated by eq. (13).

Since the effects in \( V \) are not exclusive nor exhaustive, \( \{\Pi(v_i^j),...\} \) is not a possibility distribution, but \( \{\Pi(v_j^j),\Pi(v_i^j)\} \) is. Thus, \( \Pi(v_j^j) \) and \( \Pi(v_i^j) \) are relative uncertainties between \( v_j \) and \( v_i \). In order to obtain the relative possibilities by eq. (13), \( \Pi(u_i^j) \) and \( \Pi(u_j^i) \) as well as \( \Pi(c_i^j | u_i) \) and \( \Pi(c_j^i | u_j) \) should also be relative uncertainties between \( u_i^j \) and \( u_j^i \), and between \( c_i^j \) and \( c_j^i \) conditioned by \( u_i^j \).

However, from the view point of application, what we want to know more is the relative uncertainty among \( v_j^i \), \( (j = 1,...,M) \), namely, which \( v_j^i \) is more possible than the others. In order to give the semantics to \( \Pi(v_j^i) \), \( \Pi(u_j^i) \) should also have relativity among \( u_j^i \) \( (i=1,...,N) \). In addition, the magnitude of \( \Pi(u_j^i) \) must be comparable to the one of \( \Pi(c_j^i | u_j) \).

Based on the discussion above, we propose the next transformation, if the model is a \( \Omega \)-model.

\[
\Pi(u_j^i) = \theta(u_j^i) \left( \frac{1}{\sqrt{\sum_{k=1}^{N} \theta(u_k^i)}} \right), \quad (22)
\]

\[
\theta(u_j^i) = \sum_{k=1}^{N} (P_G(u_k^i) \land P_G(u_j^i)), \quad (23)
\]

\[
\Pi(c_j^i | u_j^i) = \left\{ P_G(c_j^i \land u_j^i) \right\} + P_G(c_j^i | u_j^i), \quad (24)
\]

\[
\Pi(c_i^j | u_j^i) = \left\{ P_G(c_i^j \land u_j^i) \right\} + P_G(c_i^j | u_j^i), \quad (25)
\]

where \( \sqrt{\sum_{k=1}^{N} \theta(u_k^i)} \geq 1 \), because \( P_G(u_1^i) + ... + P_G(u_N^i) \geq 1 \).

Since causes in \( U \) are considered exhaustive in \( \Omega \)-model, \( \Pi(u_j^i) \) is derived by the following:

\[
\Pi(u_j^i) = \sqrt{\Pi(u_j^i)}. \quad (26)
\]

The possibilities given by eq. (22) clearly express the relative uncertainties among \( u_j^i \) \( (i=1,...,N) \). The maximum value is always 1.0. So, they are comparable to the possibilities given by eq. (24).

In the case of \( G \)-model, neither exclusiveness nor exhaustiveness holds for causes in \( U \). Thus, we introduce a new event \( u_{N+1} \), which represents the event meaning nothing in \( U \) is present. \( u_{N+1} \leftrightarrow \overline{\Omega} \), where \( \overline{\Omega} \) is the negation of \( \Omega \) defined by formula (14). Then, the transformation of \( P_G(u_j^i) \) is conducted by the following:

\[
\Pi(u_j^i) = \sigma(u_j^i) \left( \frac{1}{\sum_{k=1}^{N+1} \sigma(u_k^i)} \right), \quad (27)
\]

\[
\sigma(u_j^i) = \sum_{k=1}^{N+1} (P_G(u_k^i) \land P_G(u_j^i)), \quad (28)
\]

\[
\Pi(u_j^i) = \sqrt{\Pi(u_j^i)} \quad (29)
\]

the transformation of \( P_G(c_j^i | u_j^i) \) is done using eq. (24) and (25), because \( P_G(c_j^i | u_j^i) = P_G(c_j^i | u_j^i) \).
5 Numerical Experiments
Numerical experiments are conducted to evaluate the proposed approach, mainly the effects of transformation from PrCM to PoCM. First, many datasets $D_h = \{d_h^1, ..., d_h^K\}$, $(h = 1, ..., H)$ are generated from a given PrCM named ProbCM. The size of each dataset $K$ must be small so that the moderate ordering effect of possibility could be observed. Then, PrCMs named Pro(h) are developed from each $D_h$ in the way of section 3, and Pro(h) are transformed into PoCMs named Pos(h) as in section 4. Case-3 in section 3.1 is assumed. Parameters used are $N=10$ (number of causes), $M=5$ (effects), $H=20$ (datasets) and $K=30$ (size of each dataset).

The transformation is evaluated by comparing uncertainties of $v_j$ obtained from the developed models. Uncertainties of $v_j$ are reasoned by eq. (9) for ProbCM and Pro(h), and by eq. (13) for Pos(h). Probabilities for ProCM, average probabilities for Pro(h) and average values of $\Pi(v_j^1)-\Pi(v_j)$ for Pos(h) are shown in Table 1. The uncertainties are sorted in the decreasing order for each model, and the orders for ProCM and average orders for Pro(h) and Pos(h) are shown in Table 2.

Table 3 shows average degrees of compatibility and incompatibility of the ordered sequence of $v_j$ between Pro(h) (or Pos(h)) and ProbCM. They are defined by

$$Com(md) = \sum_{h<j} Com_{md}(v_h, v_j),$$

$$Com_{md}(v_h, v_j) = \begin{cases} 1, & \text{if } d(v_h, v_j) > 0, \\ 0, & \text{otherwise}, \end{cases}$$

$$Inc(md) = \sum_{h<j} Inc_{md}(v_h, v_j),$$

$$Inc_{md}(v_h, v_j) = \begin{cases} 1, & \text{if } d(v_h, v_j) < 0, \\ 0, & \text{otherwise}, \end{cases}$$

where $md$ denotes Pro(h) or Pos(h), and $Com(md)$ ($Inc(md)$) means the degree of compatibility (incompatibility) between $md$ and the given model ProbCM. $Com_{md}(v_h, v_j)$, ($Inc_{md}(v_h, v_j)$) takes 1, if the orders of $v_h$ and $v_j$ for uncertainties are compatible (incompatible) between $md$ and ProbCM, and 0, otherwise.

<p>| Table 1 Average of Reasoned Uncertainties of $v_j$ |
|------------------|------------------|------------------|------------------|------------------|</p>
<table>
<thead>
<tr>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$v_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProbCM</td>
<td>0.477</td>
<td>0.523</td>
<td>0.455</td>
<td>0.261</td>
</tr>
<tr>
<td>Ave. of Pro(h)</td>
<td>0.556</td>
<td>0.581</td>
<td>0.529</td>
<td>0.338</td>
</tr>
<tr>
<td>Ave. of Pos(h)</td>
<td>-0.043</td>
<td>0.005</td>
<td>-0.005</td>
<td>-0.124</td>
</tr>
</tbody>
</table>

<p>| Table 2 Averaged Orders of Reasoned Uncertainties |
|------------------|------------------|------------------|------------------|</p>
<table>
<thead>
<tr>
<th>$v_1$</th>
<th>$v_2$</th>
<th>$v_3$</th>
<th>$v_4$</th>
<th>$v_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ProbCM</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Ave. of Pro(h)</td>
<td>2.3</td>
<td>1.8</td>
<td>2.4</td>
<td>4.7</td>
</tr>
<tr>
<td>Ave. of Pos(h)</td>
<td>2.4</td>
<td>1.3</td>
<td>1.4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Table 3 Average of Compatibility / Incompatibility

<table>
<thead>
<tr>
<th>quantization num.</th>
<th>Average for Pro(h)</th>
<th>Average for Pos(h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Com$</td>
<td>$Inc$</td>
<td>$Inc/Com$</td>
</tr>
<tr>
<td>101</td>
<td>8.0</td>
<td>1.6</td>
</tr>
<tr>
<td>21</td>
<td>8.0</td>
<td>1.8</td>
</tr>
<tr>
<td>11</td>
<td>6.7</td>
<td>1.0</td>
</tr>
<tr>
<td>6</td>
<td>7.1</td>
<td>1.0</td>
</tr>
</tbody>
</table>
other words, if the effect of partial ordering should be large or the reasoning results should be moderate, possibilistic models might be a better choice than probabilistic ones.

6 Conclusion

The paper proposes a method to develop a possibilistic causal model from a dataset. It first develops a probabilistic one, then transforms it into possibilistic one. The paper indicates that the transformation should take into account the overall relatively of events in all the universal sets, and proposes a way to transform a PrCM into a PoCM.

The proposed approach are evaluated with a numerical example. It shows a practical merit of using PoCMs; PoCMs derive moderate reasoning results than PrCMs. Thus, possibilistic models might be a better choice, if it is preferable, typically in the case where the there is no enough number of data.

References:


Appendix: Since the causes are mutually independent,  
\[ 1 - P_C(\Omega) = (1 - P_C(u_1)) \cdot \ldots \cdot (1 - P_C(u_N)) \]  
Substituting  
P_C(u_i)P_C(\Omega) \quad \text{for} \quad P_C(u_i) ,
\[ \prod_i (1 - P_C(u_i)) P_C(\Omega) + P_C(\Omega) - 1 = 0 \]  
Let  
q_i = P_C(u_i) , \quad p = P_C(\Omega) , \quad \text{and} \quad f(p) = \prod_i (1 - q_i p) + p - 1  
for simplicity. Then, 
\[ f^{(0)}(p) = \partial f(p)/\partial p = -\sum q_i \prod_{i \neq i} (1 - q_i) + 1 \]  
Considering  
f^{(0)}(0) = 0 , \quad f^{(0)} = \prod_i (1 - q_i) \geq 0 , \quad f^{(0)}(0) = 1 - \sum q_i \leq 0 , \quad f^{(0)}(1) = 1 - \sum q_i \prod_{i \neq i} (1 - q_i) \geq f^{(0)}(0) \quad \text{and} \quad f^{(0)}(p) \geq 0 , \quad \text{only a solution of} \ f(p) = 0 \text{exists clearly in} \ (0,1] , \quad \text{except for the case where} \ \exists i \forall j \neq i, q_i = 1 \wedge q_j = 0 \]  
Since the exceptional case does not appear in practical situations, the solution  
\[ p = P_C(\Omega) \]  
satisfying  
\[ f(p) = 0 \]  
is derived by a numerical calculation.