Probabilistic Estimation for Generalized Rough Modus Ponens and Rough Modus Tollens

Ning Yao[1,2], Duoqian Miao[1,2](✉), Zhifei Zhang[1,2], and Guangming Lang[1,2]

[1] Department of Computer Science and Technology, Tongji University, Shanghai 201804, China
dqmiao@tongji.edu.cn

[2] The Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 201804, China
xinhuperfect@163.com

Abstract. We review concepts and principles of Modus Ponens and Modus Tollens in the areas of rough set theory and probabilistic inference. Based on the upper and the lower approximation of a set as well as the existing probabilistic results, we establish a generalized version of rough Modus Ponens and rough Modus Tollens with a new fact different from the premise (or the conclusion) of "if ... then ..." rule, and address the problem of computing the conditional probability of the conclusion given the new fact (or of the premise given the new fact) from the probability of the new fact and the certainty factor of the rule. The solutions come down to the corresponding interval for the conditional probabilities, which are more appropriate than the exact values in the environment full of uncertainty due to errors and inconsistency existed in measurement, judgement, management, etc., plus illustration analysis.

Keywords: Rough modus ponens · Rough modus tollens · The lower approximation · Conditional probability · Rough sets

1 Introduction

At the center of human intelligence and reasoning lies common sense, gained from experience of life or common knowledge. Knowledge is often acquired from data such as observations and measurements in the form of numbers, words, or images, usually represented in an organized manner with a level of granularity, and pervaded by imprecision or vagueness. However, data, collected for use, are generally disorganized and contain useless details. Therefore, how to obtain the available knowledge or information from data is a central point in data analysis whose goal is finding patterns or regularity hidden in the data. The utilization of statistics was only realizable in the early period of data analysis, then followed by fuzzy sets, rough sets, neural networks, genetic algorithms, cluster analysis and other analysis tools.

Typically encoded as the rule of "if ... then ...", hidden patterns or regularity in data can enable us to make decisions, do prediction and management

V. Flores et al. (Eds.): IJCRS 2016, LNAI 9920, pp. 166–176, 2016.
DOI: 10.1007/978-3-319-47160-0_15

activities, or other reasoning activities in everyday life, and the highly influential comes down to Modus Ponens inference rule (Modus Ponens, for short), as the basis of classical deductive reasoning and an universal rule of inference valid in any logical system. Modus Ponens has the form that, given that a formula or a new fact ϕ is true and the rule "if ϕ then ψ" is also true, then the formula or the conclusion ψ would also be true. To estimate the truth value of ψ relates closely to the formalization of the conditional "if ... then ... ", and in turn, to formalize "if ... then ... " has led to the keen competition between material implication and conditional probability. For the material implication of Modus Ponens, the representative work is Compositional Rule of Inference as an approximate extension of Modus Ponens [17], proposed by Zadeh based on fuzzy sets, where fuzzy relations derived from fuzzy implication operators are employed to compute the truth value of the conclusion ψ and the new fact is allowed to different from the premise of the rule. However, considering lack of the ability to treat the exceptions or counterfactuals for material implication and its inherent paradoxes, for example, the false of the premise does infer the true of material implication only if the conclusion is true, the probabilistic interpretation of "if ... then ... " is more plausible in the reasoning process [5,13].

In addition, due to the uncertainty in the represented data or the knowledge, inconsistency in information systems, and the limited number of available knowledge obtained for use, directly characterizing the truth values of formulas is not feasible because of many difficulties in the construction of the truth function and can often be influenced by subjective factors such as assumption intervention. The idea of replacing truth values with probabilities was first proposed by Łukasiewicz, who advocates multivalued logic as probability logic and assigns each of indefinite proposition $\phi(x)$ the ratio $\pi(\phi(x))$ of the number of all values of the variable x satisfying $\phi(x)$ to the number of all possible values of x as the truth value of $\phi(x)$. And later, Pawlak, the founder of rough sets with the aim of finding the dependencies or cause-effect relations in data, pursued this idea and introduced Rough Modus Ponens, a generalized version of Modus Ponens in the context of rough set theory where the new fact is of the same form as the premise of the rule [10]. Fuzzy-Rough version of Modus Ponens [4] presented the characterization of the conclusion through gradual decision rules extracted from decision table based on fuzzy rough set theory, plus the fuzzy-rough version for modus tollens (i.e., given that a formula $\neg\psi$ is true and the rule "if ϕ then ψ" is also true, then the formula $\neg\phi$ would also be true), without using any fuzzy logical connectives. Although this approach is successful in the treatment of the difference between the new fact and the premise, it still involves the selection of fuzzy membership function influenced by subjective factors. Probabilistic counterpart of Modus Ponens yields the best possible bounds for the probability of the conclution ψ and even for the update of the bounds on new-found uncertain evidence as well as the bounds for modus tollens [12,15].

Reasoning based on rough set theory obeys data collected and the inferences stem from the data. Empowered by these motivations and analysis, the central goal of this paper is to investigate the generalized version of Rough Modus

Ponens permitting the new fact different from the premise, and try to allow the solution to make the plausible responds to the new evidence even when the evidence is contradictory or irrelevant to the premise of the rule, as well as the study for Rough Modus Tollens. Section 2 exhibits some definitions about rough set theory as well as some results on probable Modus Ponens and Modus Tollens. In Sect. 3 the generalized rough Modus Ponens and rough Modus Tollens are developed through the consideration for the relations between the new fact and the premise based on the concept of lower approximation of the set, together with some illustration studies depicted. The detailed comments on the present approach are explored in Sect. 4, along with a brief sketch of further research.

2 Basic Concepts on Rough Modus Ponens and Rough Modus Tollens

Subject to measurability requirements, one is led to consider upper and lower approximations defined over any set as follows:

Definition 1 ([8]). *Given an information system $S = (U, A)$ with U a nonempty finite set called the Universe and A a nonempty finite set called the set of Attributes, and let $X \subseteq U$, $B \subseteq A$. The upper approximation $\overline{B}(X)$ and the lower approximation $\underline{B}(X)$ of any set X in terms of attributes B can be defined respectively by $\overline{B}(X) = \bigcup_{x \in U} \{x \in U : [x]_B \cap X \neq \emptyset\}$, $\underline{B}(X) = \bigcup_{x \in U} \{x \in U : [x]_B \subseteq X\}$, where $[x]_B$ (i.e., the set of $\{y \in U : y\ I(B)\ x\}$) denotes the equivalence class of x with respect to the indiscernibility relation $I(B)$(i.e., $\{(x, y) \in U^2 | a(x) = a(y)$ for every $a \in B\}$, $a(x)$ denotes the value of attribute a for element x), which means the object y and x are indiscernible in terms of attributes in B.*

Definition 2 ([9,11]). *Given a decision table $S = (U, C, D)$ with the attributes A of the system classified into disjoint sets of condition attributes C and decision attributes D, and let $\phi \to \psi$ be a decision rule with ϕ and ψ as logical formulas representing conditions and decisions, respectively. Define the certainty factor $\mu(\phi, \psi)$ of the rule as a number, namely, $\mu(\phi, \psi) = \pi(\psi|\phi) = \frac{\pi(\phi \wedge \psi)}{\pi(\phi)} = \frac{card(\|\phi\| \cap \|\psi\|)}{card(\|\phi\|)}$, where $\|\phi\|$ denote the set of all objects satisfying ϕ in S, $card(\cdot)$ denotes the cardinality or the number of elements in a given set, and $\pi(\cdot)$ represents the corresponding probability (the purpose of using this notation as probability is only to accord with the ones in the rough set literature), $\pi(\phi) = \frac{card(\|\phi\|)}{card(U)}$ and $card(\|\phi\|) \neq 0$.*

The rough modus ponens [10] may be formed from

$$\begin{array}{lll} \text{if} & \phi \to \psi \text{ is true with probability} & \pi(\psi|\phi) \\ \text{and} & \phi \quad\quad \text{is true with probability} & \pi(\phi) \\ \hline \text{then} & \quad\quad \psi \text{ is true at least with probability} & \pi(\psi) \end{array}$$

where $\pi(\psi) = \pi(\neg\phi \wedge \psi) + \pi(\phi) \cdot \pi(\psi|\phi)$. This formula can be taken as a generalization (e.g., for $\pi(\phi \rightarrow \psi) \neq 0$) of Łukasiewicz' axiom 3 (i.e., if $\pi(\phi \rightarrow \psi) = 1$, then $\pi(\psi) = \pi(\neg\phi \wedge \psi) + \pi(\phi)$).

From a probabilized point of view, as to modus ponens, one has the best possible bounds [12,15] for $\pi(\psi)$, namely, $\pi(\phi)\pi(\psi|\phi) \leq \pi(\psi) \leq \pi(\phi)\pi(\psi|\phi) + 1 - \pi(\phi)$ with $0 < \pi(\phi) \leq 1$ and $0 \leq \pi(\psi|\phi) \leq 1$; as to modus tollens, given $\neg\psi$ and the rule $\phi \rightarrow \psi$ to infer $\neg\phi$, the solution with the best possible bounds is (see the theorem on p. 751 of [12]: 'ϕ' and 'ψ' for 'H' and 'E'; $\neg$ for over-bars; '$\pi(\psi|\phi)$'and '$\pi(\neg\psi)$' for 'a' and 'b')

$$\begin{aligned}
&if\ \ 0 < \pi(\neg\psi), \pi(\psi|\phi) < 1,\ then \\
&\qquad \max\left\{\tfrac{1-\pi(\psi|\phi)-\pi(\neg\psi)}{1-\pi(\psi|\phi)}, \tfrac{\pi(\psi|\phi)+\pi(\neg\psi)-1}{\pi(\psi|\phi)}\right\} \leq \pi(\neg\phi) < 1 \\
&if\ \ 0 < \pi(\neg\psi) \leq 1, \pi(\psi|\phi) = 0,\ then\ 1 - \pi(\neg\psi) \leq \pi(\neg\phi) < 1 \\
&if\ \ 0 \leq \pi(\neg\psi) < 1, \pi(\psi|\phi) = 1,\ then\ \pi(\neg\psi) \leq \pi(\neg\phi) < 1.
\end{aligned}$$

Moreover, concerning the update of the probability for ψ on new-found possibly uncertain evidence, the solution has been obtained as follows [15]:

*Let time $-t$ be, for a person, just before time t probabilistically speaking. Assume that this person is not certain of $\neg\phi$ at t, that is, $\pi_t(\phi) > 0$ and $\pi_t(\phi) \neq \pi_{-t}(\phi)$. Then this person does update his probability for ψ subject to the bounds $\pi_t(\phi)\pi_{-t}(\psi|\phi)$ and $\pi_t(\phi)\pi_{-t}(\psi|\phi) + 1 - \pi_t(\phi)$, **if and only if**, the rigidity-condition for ψ on ϕ, i.e., $\pi_t(\psi|\phi) = \pi_{-t}(\psi|\phi)$, is satisfied.*

Analogous to the update of the probability for ψ, updating $\neg\phi$ on new-found possibly uncertain evidence $\neg\psi$ has been solved to yield [15].

*Let time $-t$ be, for a person, just before time t probabilistically speaking. Assume that $\pi_t(\neg\psi) \neq \pi_{-t}(\neg\psi)$. Then this person does update his probability for $\neg\phi$ on the evidence $\neg\psi$, **if and only if**, the rigidity-condition $\pi_t(\psi|\phi) = \pi_{-t}(\psi|\phi)$ is satisfied. If this condition is satisfied, there is*

$$\begin{aligned}
&if\ \ 0 < \pi_t(\neg\psi), \pi_t(\psi|\phi) < 1,\ then \\
&\qquad \max\left\{\tfrac{1-\pi_t(\psi|\phi)-\pi_t(\neg\psi)}{1-\pi_t(\psi|\phi)}, \tfrac{\pi_t(\psi|\phi)+\pi_t(\neg\psi)-1}{\pi_t(\psi|\phi)}\right\} \leq \pi_t(\neg\phi) < 1 \\
&if\ \ 0 < \pi_t(\neg\psi) \leq 1, \pi_t(\psi|\phi) = 0,\ then\ 1 - \pi_t(\neg\psi) \leq \pi_t(\neg\phi) < 1 \\
&if\ \ 0 \leq \pi_t(\neg\psi) < 1, \pi_t(\psi|\phi) = 1,\ then\ \pi_t(\neg\psi) \leq \pi_t(\neg\phi) < 1.
\end{aligned}$$

3 Generalized Versions of Rough Modus Ponens and Rough Modus Tollens

In this section we continue Pawlak's work [3,10,11] and it is convenient to begin with the case where $\phi^\diamond$ takes a different form of ϕ but the same rule "if ϕ then ψ" as the case of Modus Ponens, associating this rule with a conditional probability $\pi(\psi|\phi) = \frac{\pi(\phi\wedge\psi)}{\pi(\phi)}$ and likewise the formulas ϕ and ψ with their respective unconditional probabilities $\pi(\phi)$ and $\pi(\psi)$.

Consider that, in practice, the new observation is rarely identical to the sample data but to some extent is of particular relevance to the observed sample (i.e., they describe different states of the same attributes or different attributes of different attributes). Here we denote the new fact or observation by $\phi^\diamond$ and mainly deal with the case when $\phi^\diamond$ is not the same as the sample ϕ. The above Rough Modus Ponens may be regarded as the special case of the proposed generalized version:

$$\begin{array}{lll} \text{if} \quad \phi \to \psi & \text{is true with probability} & \pi(\psi|\phi) \\ \text{and} \ \phi^\diamond & \text{is true with probability} & \pi(\phi^\diamond) \\ \hline \text{then} \qquad ?\psi & \text{is true with probability} & ?\pi_t(\psi) \ ?\pi(\psi|\phi^\diamond) \end{array}$$

where the notation $\pi_t(\psi)$ means the probability of ψ when the new fact $\phi^\diamond$ is observed, which is identical to the meaning of the conditional probability of ψ given $\phi^\diamond$. The subscript t is used only to distinguish the probability of ψ when the new fact $\phi^\diamond$ is observed from the prior probability of ψ as well as the posterior probability of ψ when the evidence ϕ is observed.

Lemma 1. *Let $0 < \pi(\psi|\phi) \leq 1$ and $0 < \pi(\phi^\diamond) \leq 1$. Then the probability of ψ with $\phi^\diamond$ known satisfies*

$$\pi(\phi^\diamond)\pi(\psi|\phi) \leq \pi_t(\psi) \leq \pi(\phi^\diamond)\pi(\psi|\phi) + 1 - \pi(\phi^\diamond).$$

The relation between the new fact $\phi^\diamond$ and the occurrence of ψ is connected closely to the relation of the new fact $\phi^\diamond$ and the premise ϕ. Our solution lies in the detailed description of the relation between $\phi^\diamond$ and ϕ, more specifically, the lower approximation of $\phi^\diamond$ and the lower approximation of ϕ.

Lemma 2. *For $\underline{\phi^\diamond} \subseteq \underline{\phi}$, the probability $\pi(\psi|\phi^\diamond)$ satisfies*

$$\frac{\pi(\psi|\phi)\pi(\phi)}{\pi(\phi^\diamond)} = \frac{\pi(\phi \wedge \psi)}{\pi(\phi^\diamond)} \leq \pi(\psi|\phi^\diamond) = \frac{\pi(\phi^\diamond \wedge \psi)}{\pi(\phi^\diamond)} \leq \frac{\pi(\phi^\diamond)}{\pi(\phi^\diamond)} = 1.$$

Proof. This result follows immediately from the fact that the frequency of the occurrence of one event is usually greater and equal to the frequency of the simultaneous occurrence of this event together with other events. Let $\|\phi^\diamond\|$ and $\|\phi\|$ represent the sets of all objects satisfying respectively $\phi^\diamond$ and ϕ in S, there exist $card(\|\phi^\diamond\|) \geq card(\|\phi\|)$ and $card(\|\phi^\diamond \wedge \psi\|) \geq card(\|\phi \wedge \psi\|)$, plus $\pi(\phi^\diamond \wedge \psi) = \frac{card(\|\phi^\diamond \wedge \psi\|)}{card(U)}$, $\pi(\phi \wedge \psi) = \frac{card(\|\phi \wedge \psi\|)}{card(U)}$ and $\pi(\phi^\diamond \wedge \psi) \leq \pi(\phi^\diamond)$. □

Theorem 1. *If $\underline{\phi^\diamond} \subseteq \underline{\phi}$, the probability of ψ given $\phi^\diamond$ can be solved by*

$$\max\left\{\frac{\pi(\psi|\phi)\pi(\phi)}{\pi(\phi^\diamond)}, \pi(\phi^\diamond)\pi(\psi|\phi)\right\} \leq \pi(\psi|\phi^\diamond) \leq \pi(\phi^\diamond)\pi(\psi|\phi) + 1 - \pi(\phi^\diamond)\ .$$

Example 1. From Table 1 (see [10]), we have the rule "if $\phi = (Headache, yes)$ *and* $(Muscle - pain, no)$ *and* $(Temperature, high)$, then $\psi = (Flu, yes)$" with the probability $\pi(\psi|\phi) = \frac{1}{2}$ and $\pi(\phi) = \frac{1}{3}$, and a new fact

Table 1. Characterization of flu

Patient	Headache	Muscle-pain	Temperature	Flu
p1	no	yes	high	yes
p2	yes	no	high	yes
p3	yes	yes	very high	yes
p4	no	yes	normal	no
p5	yes	no	high	no
p6	no	yes	very high	yes

$\phi^\diamond = (Headache, yes)$ with $\pi(\phi^\diamond) = \frac{1}{2}$. Because $\underline{\phi^\diamond} = \{(Headache, yes)\} \subseteq \underline{\phi}$, the probability of $\psi = (Flu, yes)$ given $\phi^\diamond = (Headache, yes)$ lies within the interval $[\frac{1}{3}, \frac{3}{4}]$ by Theorem 1 (if possible, based on rough set theory from the table one has $\pi(\psi|\phi^\diamond) = \frac{2}{3}$).

When $\underline{\phi^\diamond} \subseteq U - \underline{\phi}$ but $\underline{\phi^\diamond} \cap \underline{\phi} \neq \emptyset$, $\pi(\psi|\phi^\diamond)$ depends on whether or not the element of $\underline{\phi^\diamond} - \underline{\phi^\diamond} \cap \underline{\phi}$ is the description of the same attribute with the different states from the one of $\underline{\phi} - \underline{\phi^\diamond} \cap \underline{\phi}$.

Theorem 2. *If the elements of* $\underline{\phi^\diamond} - \underline{\phi^\diamond} \cap \underline{\phi}$ *and* $\underline{\phi} - \underline{\phi^\diamond} \cap \underline{\phi}$ *depict different states of the same attribute, then the probability* $\pi(\psi|\phi^\diamond)$ *can be determined by the interval* $[\pi(\phi^\diamond)\pi(\psi|\phi), \pi(\phi^\diamond)\pi(\psi|\phi) + 1 - \pi(\phi^\diamond)]$*, more specifically, for* $0 < \pi(\psi|\phi) \leq 1$*,* $\pi(\psi|\phi^\diamond)$ *can be located inside or outside this interval, which corresponds to the degrees of beliefs for the attribute in the language of* $\underline{\phi^\diamond} - \underline{\phi^\diamond} \cap \underline{\phi}$ *and* $\underline{\phi} - \underline{\phi^\diamond} \cap \underline{\phi}$*. By contrast, if the elements of* $\underline{\phi^\diamond} - \underline{\phi^\diamond} \cap \underline{\phi}$ *and* $\underline{\phi} - \underline{\phi^\diamond} \cap \underline{\phi}$ *depict different states of different attributes, then the probability of* ψ *given* $\psi^\diamond$*, namely,* $\pi(\psi|\phi^\diamond)$ *is generally situated in* $[\pi(\phi^\diamond)\pi(\psi|\phi), \pi(\phi^\diamond)\pi(\psi|\phi) + 1 - \pi(\phi^\diamond)]$.

Example 2. From Table 1, one have the rule "if $\phi = (Headache, no)$ then $\psi = (Flu, yes)$" with $\pi(\psi|\phi) = \frac{2}{3}$ as well as a new fact $\phi^\diamond = (Headache, no)$ *and* $(Temperature, normal)$ with $\pi(\phi^\diamond) = \frac{1}{6}$. Additionally we have known that the probability of 'if Temperature is normal then Flu is yes' is 0 and the probability of 'if Temperature is very high then Flu is yes' is 1. According to Theorem 2, one can determine that the value $\pi(\psi|\phi^\diamond)$ is outside $[\frac{1}{9}, \frac{17}{18}]$ and specifically in $[0, \frac{1}{9}]$. (From Table 1 one has $\pi(\psi|\phi^\diamond) = 0 \in [0, \frac{1}{9}]$, which means that the obtained result acts in accordance with our common sense.)

When $\underline{\phi^\diamond} \cap \underline{\phi} = \emptyset$, similarly $\pi(\psi|\phi^\diamond)$ is associated with the fact whether $\phi^\diamond$ depicts the same attributes as ϕ does or as the elements of ϕ do. In more details, if they do, then the range of $\pi(\psi|\phi^\diamond)$ closely relates to the degrees of beliefs for these attributes of $\phi^\diamond$ and ϕ, specified in the following result.

Theorem 3. *Let* $\underline{\phi^\diamond} \cap \underline{\phi} = \emptyset$*. If the elements of* $\phi^\diamond$ *and* ϕ *describe different states of the same attribute, then for* $\pi(\psi|\phi) = 1$*,* $\pi(\psi|\phi^\diamond)$ *might smaller than or equal to* 1 *and the specific value will be inside or outside* $[\pi(\phi^\diamond)\pi(\psi|\phi), 1]$ *with a trend of moving from right to left on the horizontal axis according to the degrees of beliefs* *for the attribute; for* $0 < \pi(\psi|\phi) < 1$, $\pi(\psi|\phi^\diamond)$ *might be* 0 *or* 1 *and the specific value also might be inside or outside the interval* $[\pi(\phi^\diamond)\pi(\psi|\phi), \pi(\phi^\diamond)\pi(\psi|\phi) + 1 - \pi(\phi^\diamond)]$ *according to the degrees. If they depict different states of different attributes,* $\pi(\psi|\phi^\diamond)$ *generally lies in* $[\pi(\phi^\diamond)\pi(\psi|\phi), \pi(\phi^\diamond)\pi(\psi|\phi) + 1 - \pi(\phi^\diamond)]$.

Example 3. From Table 1, one can get the rule "if $\phi = (Headache, no)$ *and* $(Muscle - pain, yes)$ *and* $(Temperature, high)$, then $\psi = (Flu, yes)$" with the probability $\pi(\psi|\phi) = 1$ and $\pi(\phi) = \frac{1}{6}$, and the new fact $\phi^\diamond = (Temperature,$ *very high*$)$ with $\pi(\phi^\diamond) = \frac{1}{3}$. By Theorem 3, $\pi(\psi|\phi^\diamond)$ falls into $[\frac{1}{3}, 1]$ and from the table one might get $\pi(\psi|\phi^\diamond) = 1 \in [\frac{1}{3}, 1]$.

Obviously notice that the case of $\pi(\psi|\phi) = 1$ is the special case of the above result. Clearly in this case there is $\pi(\phi^\diamond) \leq \pi_t(\psi) \leq 1$. In particular, let $C(\psi)$ denote the set of all conditions of ψ in the data table about the domain of interest, and $C_*(\psi)$ denote the lower approximation of $C(\psi)$ defined by [11] $C_*(\psi) = \bigcup_{\phi^\diamond \in C(\psi), \pi(\psi|\phi^\diamond)=1} \|\phi^\diamond\| = \|\bigvee_{\phi^\diamond \in C(\psi), \pi(\psi|\phi^\diamond)=1} \phi^\diamond\|$. When $\phi^\diamond \in C_*(\psi)$, we have $\pi(\psi|\phi^\diamond) = 1$ and furthermore, if $\phi^\diamond \notin C_*(\psi)$, we need to study the relation between $\phi^\diamond$ and ϕ. If $\underline{\phi^\diamond} \subseteq \underline{\phi}$, then $\max\left\{\pi(\phi^\diamond), \frac{\pi(\phi)}{\pi(\phi^\diamond)}\right\} \leq \pi(\psi|\phi^\diamond) = \frac{\pi(\phi^\diamond \wedge \psi)}{\pi(\phi^\diamond)} \leq 1$, otherwise the probability of ψ given $\phi^\diamond$ will be from the inside or the outside of the interval $[\pi(\phi^\diamond), 1]$.

Analogous to the discussion of the rough modus ponens, consider the Rough Modus Tollens, formed from [3,11]

$$\begin{array}{llll} \text{if} & \phi \to \psi & \text{is true with probability} & \pi(\phi|\psi) \\ \text{and} & \quad\;\; \psi & \text{is true with probability} & \pi(\psi) \\ \hline \text{then } \phi & & \text{is true with probability} & \pi(\phi) \end{array}$$

where $\pi(\phi) = \pi(\phi \wedge \neg\psi) + \pi(\psi)\pi(\phi|\psi)$.
From the conditional probability point of view, there is

$$\begin{array}{ll} if & 0 < \pi(\psi|\phi), \pi(\psi) < 1, \qquad then\ 0 < \pi(\phi) \leq \min\left\{\frac{1-\pi(\psi)}{1-\pi(\psi|\phi)}, \frac{\pi(\psi)}{\pi(\psi|\phi)}\right\} \\ if & \pi(\psi|\phi) = 0, 0 \leq \pi(\psi) < 1, then\ 0 < \pi(\phi) \leq 1 - \pi(\psi) \\ if & \pi(\psi|\phi) = 1, 0 < \pi(\psi) \leq 1, then\ 0 < \pi(\phi) \leq \pi(\psi) \end{array}$$

To put it in another way, one has

$$\pi(\psi)\pi(\phi|\psi) \leq \pi(\phi) \leq \pi(\psi)\pi(\phi|\psi) + 1 - \pi(\psi)\ with\ \pi(\phi|\psi) > 0.$$

The following attention in the remaining part of this section will be given to the case when the fact $\psi^\diamond$ is not always the same as the conclusion ψ of the rule $\phi \to \psi$ but can be regarded as the characterization of ψ with different beliefs such as 'if it rained then it was cold' and 'it is very cold', defined by

$$\begin{array}{lllll} \text{if} & \phi \to \psi & \text{is true with probability} & \pi(\phi|\psi) & \\ \text{and} & \qquad \psi^\diamond & \text{is true with probability} & \pi_t(\psi^\diamond) & \\ \hline \text{then } ?\phi & & \text{is true with probability} & ?\pi_t(\phi) & ?\ \pi_t(\phi|\psi^\diamond) \end{array}$$

It is worth mentioning that, $\psi^\diamond$ takes the different form from the one of ψ, $\pi_t(\psi^\diamond)$ denotes the prior probability of $\psi^\diamond$ which is different from the probability of ψ, $\pi_t(\phi|\psi^\diamond)$ represents the conditional probability of ϕ given $\psi^\diamond$ and is identical to the posterior probability $\pi_t(\phi)$ of ϕ when the new fact ψ is observed. The subscript t means the derived conclusions is inferred under the condition that a new fact $\psi^\diamond$ occurs.

Theorem 4. *The estimation for the probability of ϕ given $\psi^\diamond$ can be formed in*

$$\left[\pi_t(\psi^\diamond)\pi(\phi|\psi), \pi_t(\psi^\diamond)\pi(\phi|\psi)+1-\pi_t(\psi^\diamond)\right] \bigcap \left(0, \min\left\{\tfrac{1-\pi_t(\psi^\diamond)}{1-\pi(\psi|\phi)}, \tfrac{\pi_t(\psi^\diamond)}{\pi(\psi|\phi)}\right\}\right] \bigcap \left[\pi(\phi|\psi)\pi(\psi^\diamond|\phi), 1\right]$$

Proof. Given the fact $\psi^\diamond$ and the rule $\phi \rightarrow \psi$, the calculation of $\pi_t(\phi)$ follows from the intersection of

$$\pi_t(\psi^\diamond)\pi(\phi|\psi) \leq \pi_t(\phi) \leq \pi_t(\psi^\diamond)\pi(\phi|\psi)+1-\pi_t(\psi^\diamond)$$

and $0 < \pi_t(\phi) \leq \min\left\{\frac{1-\pi_t(\psi^\diamond)}{1-\pi(\psi|\phi)}, \frac{\pi_t(\psi^\diamond)}{\pi(\psi|\phi)}\right\}$, where $\pi(\phi|\psi)$ and $\pi(\psi|\phi)$ can be estimated by the definition of certainty factor of the rule, that is, $\pi(\phi|\psi) = \frac{card(\|\phi\wedge\psi\|)}{card(\|\psi\|)}$ and $\pi(\psi|\phi) = \frac{card(\|\phi\wedge\psi\|)}{card(\|\phi\|)}$, here we postulate that the sizes of data tables or information systems in the domain of interest do not change. In addition, as for $\pi_t(\phi|\psi^\diamond)$, we shall get $\pi(\phi|\psi)\pi(\psi^\diamond|\phi) \leq \pi_t(\phi|\psi^\diamond) \leq 1$, which follows from $\frac{1}{\pi_t(\phi)} \leq \frac{1}{\pi_t(\psi^\diamond)\pi(\phi|\psi)}$ and $\frac{\pi(\phi|\psi)\pi_t(\phi\wedge\psi^\diamond)}{\pi_t(\phi)} \leq \frac{\pi_t(\phi\wedge\psi^\diamond)}{\pi_t(\psi^\diamond)}$ with $\pi(\phi|\psi) > 0$. By means of the results of $\pi_t(\phi|\psi^\diamond)$ and $\pi_t(\phi)$, the estimation can be obtained. □

Example 4. From Table 1, given the new fact of $\psi^\diamond = (Flu, no)$ and the rule "if $\phi = (Headache, yes)$ *and* $(Muscle-pain, no)$ *and* $(Temperature, high)$, then $\psi = (Flu, yes)$" with the probability $\pi(\psi|\phi) = \frac{1}{2}$, one has $\frac{1}{8} \leq \pi_t(\phi) \leq \frac{8}{12}$, which follows from $\pi_t(\psi^\diamond) = \frac{card(\|(Flu,no)\|)}{card(U)} = \frac{1}{3}$, $\pi(\phi|\psi) = \frac{1}{4}$ and the intersection of $[\frac{1}{12}, \frac{9}{12}]$ and $(0, \frac{2}{3}]$ and $[\frac{1}{8}, 1]$ according to Theorem 4. If possible, the probability $\pi_t(\phi|\psi^\diamond) = \frac{1}{2}$ in the light of rough set theory.

Example 5. Given the new fact $\psi^\diamond = (Nationality, Swede)$ and the rule "if $\phi = (Height, medium)$ *and* $(Hair, dark)$, then $\psi = (Nationality, German)$" with the probability $\pi(\psi|\phi) = \frac{90}{135} = 0.67$ (from the characterization of nationalities in [11]), then it can happen that $0.08 \leq \pi_t(\phi) \leq 0.63$, which follows from $\pi_t(\psi^\diamond) = \frac{card(\|(Nationality,Swede)\|)}{card(U)} = \frac{405}{900} = 0.45$, $\pi(\phi|\psi) = \frac{90}{495} = 0.18$ and the intersection of $[0.08, 0.63] \bigcap (0, 0.67] \bigcap [0.06, 1]$ (if possible, $\pi_t(\phi|\psi^\diamond) = \frac{45}{405} = 0.11$).

Example 6. If we have known that a new fact $\psi^\diamond = (Fly, yes)$ and the rule "if $\phi = (Bird, yes)$ *and* $(Gregarious, yes)$, then $\psi = (Fly, no)$" with the probability $\pi(\psi|\phi) = \frac{2}{7}$ (from the characterization of birds in [6]), the probability of $\phi = (Bird, yes)$ *and* $(Gregarious, yes)$ under the condition of (Fly,yes) can be solved by $\pi_t(\psi^\diamond) = \frac{8}{20}$, $\pi(\phi|\psi) = \frac{2}{12}$, $\pi(\psi^\diamond|\phi) = 1 - \frac{2}{7} = \frac{5}{7}$ and $[\frac{1}{15}, \frac{10}{15}] \bigcap (0, \frac{21}{25}] \bigcap [\frac{5}{42}, 1]$, thereby $\frac{5}{42} \leq \pi_t(\phi|\psi^\diamond) \leq \frac{10}{15}$ (if possible, $\pi_t(\phi|\psi^\diamond) = \frac{5}{8}$).

Due to space limitation, the descriptions of decision tables in Examples 5 and 6 have been omitted, and case analysis of other data tables can be taken as exercises on top of the illustrations displayed in this paper. Also it is noted that, based on rough set theory, if the data table or information system is available to "if ϕ then ψ" and $\psi^{\diamond}$, then $\pi_t(\phi|\psi^{\diamond}) = \frac{card(\|\phi \wedge \psi^{\diamond}\|)}{card(\|\psi\|)}$ with $\pi(\phi|\psi) > 0$. Moreover, if $\pi_t(\phi|\psi^{\diamond}) \neq 0$ and $\pi(\phi|\psi) \neq 0$, it means this information system is inconsistent.

4 Conclusion

In this paper, we started with the relationship or dependency between the new fact and the premise (or if clause) of the rules crystallized by human wisdom, and then presented the solutions for every different relations in the cases of rough Modus Ponens generalized by new-found possibly evidence related to the premise, finally turning to the case of rough Modus Tollens.

In light of the difficulty of gaining the exact value of $\pi_t(\phi|\psi^{\diamond})$ or $\pi(\psi|\phi^{\diamond})$, we got the interval for the possible values on the basis of the available data source, which is relatively believable compared with the subjective judgement of the fuzzy membership functions except that the reasoner is one of the experts or authorities in the domain of interest, but the expert might make false decisions or inconsistent opinions. Of course the hypothesis ensuring the validity of being believable is that the data source gathered is sound and representative so as to preserve the accuracy of the probability estimated in the process of reasoning. As can be seen from the results of examples in Sect. 3, sometimes we can obtain the exact value of the probability $\pi_t(\phi|\psi^{\diamond})$ or $\pi(\psi|\phi^{\diamond})$ through computing the corresponding certainty factors, but this is not always the case, for instance, the probability of $\phi^{\diamond} = (Height, short)$ *and* $(Hair, dark)$ as well as the conditional probability $\pi(\psi|\phi^{\diamond})$ from data table in [11] where there is no simultaneous occurrence for$(Height, short)$ and $(Hair, dark)$. The root cause of this problem lies in the incompleteness of the data source, which is an inevitable factor even in a big data environment.

Another comment in need is that by comparison with the assessment of the rough probability [7] or the measurement of the observability for the new fact in the event involved in the premise, the direct comparing between the elements of the new fact and the premise is clearer and sharper, although the rough probability has the advantage of the uncertainty measure for an event. Besides, nonmonotonic reasoning is of the center tasks of uncertainty reasoning and human reasoning has been proved to be nonmonotonic [5]. Hence the proposed solution in this paper can be viewed as an initial alternative of solving the nonmonotonic reasoning based on Modus Ponens and Modus Tollens inference patterns from the viewpoint of rough set theory. The causal effects [2] among the data collected (or the events considered) perform a crucial role in human thinking. The direct or indirect causal relationships among the data or events closely affect the treatment of the contrary facts or the irrelevant facts in human reasoning. Moreover, probabilistic rough set models such as variable precision rough set model and Bayesian rough set model [14,18], together with game-theoretic rough sets, have showed great strength in analyzing uncertainties [1,16]. Further research will be put on the relations of causal effects, the approximate characterization of sets [19] and probabilistic rough set approach.

Acknowledgments. The authors would like to thank the reviewers for their comments that help improve the manuscript. This research was supported in parts by the National Natural Science Foundation of China (No. 61273304, 61202170, 61573255, 61573259), the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20130072130004), and the program of Further Accelerating the Development of Chinese Medicine Three Year Action of Shanghai (2014–2016) (No. ZY3-CCCX-3-6002).

References

1. Azam, N., Yao, J.T.: Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets. Int. J. Approximate Reason. **55**(1), 142–155 (2014)
2. Bareinboim, E., Pearl, J.: Causal inference from big data: theoretical foundations and the data-fusion problem. UCLA Cognitive Systems Laboratory, Technical Report (R-450) (2015). Proceedings of the National Academy of Sciences (2016)
3. Greco, S., Pawlak, Z., Słowiński, R.: Generalized decision algorithms, rough inference rules, and flow graphs. In: Alpigini, J.J., Peters, J.F., Skowron, A., Zhong, N. (eds.) RSCTC 2002. LNCS (LNAI), vol. 2475, pp. 93–104. Springer, Heidelberg (2002)
4. Inuiguchi, M., Greco, S., Słowiński, R.: Fuzzy-rough modus ponens and modus tollens as a basis for approximate reasoning. In: Tsumoto, S., Słowiński, R., Komorowski, J., Grzymała-Busse, J.W. (eds.) RSCTC 2004. LNCS (LNAI), vol. 3066, pp. 84–94. Springer, Heidelberg (2004)
5. Johnson-Laird, P.N., Khemlani, S.S., Goodwin, G.P.: Logic, probability, and human reasoning. Trends Cogn. Sci. **19**(4), 201–214 (2015)
6. Jue, W., Yiyu, Y., Feiyue, W.: 'Rule plus Exception' learning based on reduct. Chin. J. Comput. (11), 1778–1789 (2005)
7. Pawlak, Z.: Rough probability. Bull. Polish Acad. Sci. Tech. **33**(9–10), 499–504 (1985)
8. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht (1991)
9. Pawlak, Z., Skowron, A.: Rough membership functions: a tool for reasoning with uncertainty. In: Algebraic Methods in Logic and in Computer Science, Banach Center Publications, vol. 28, pp. 135–150 (1993)
10. Pawlak, Z.: Rough Modus Ponens. In: Traitement d'information et gestion d'incertitudes dans les systèmesà base de connaissances. Conférence Internationale, pp. 1162–1166 (1998)
11. Pawlak, Z.: Rough sets, decision algorithms and Bayes' theorem. Eur. J. Oper. Res. **136**(1), 181–189 (2002)
12. Wagner, C.G.: Modus tollens probabilized. British J. Philos. Sci. **55**(4), 747–753 (2004)
13. Pfeifer, N., Kleiter, G.D.: Inference in conditional probability logic. Kybernetika **42**(4), 391–404 (2006)
14. Ślęzak, D., Ziarko, W.: The investigation of the bayesian rough set model. Int. J. Approximate Reason. **40**(40), 81–91 (2005)

15. Sobel, J.H.: Modus Ponens and Modus Tollens for conditional probabilities, and updating on uncertain evidence. Theor. Decis. **66**(2), 103–148 (2009)
16. Yao, Y.Y.: The superiority of three-way decisions in probabilistic rough sets models. Inf. Sci. **181**(6), 1080–1096 (2011)
17. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Syst. Man Cybern. **3**, 28–44 (1973)
18. Zhang, H.Y., Zhou, J., Miao, D.Q., Gao, C.: Bayesian rough set model: a further investigation. Int. J. Approximate Reason. **53**(4), 541–557 (2012)
19. Zhang, X.Y., Miao, D.Q., Liu, C.H., Le, M.L.: Constructive methods of rough approximation operators and multigranulation rough sets. Knowl. Based Syst. **91**, 114–125 (2016)