Investigation of monotone link between confirmation measures and rule support and confidence

Izabela Brzezińska[1], Salvatore Greco[2], Roman Słowiński[1,3]

[1] *Institute of Computing Science, Poznań University of Technology, Piotrowo 2, 60-965 Poznań, Poland*
e-mail: {izabela.brzezinska, roman.slowinski}@cs.put.poznan.pl
[2] *Faculty of Economics, University of Catania, Corso Italia, 55, 95129 Catania, Italy*
e-mail: salgreco@unicit.it
[3] *Institute for Systems Research, Polish Academy of Sciences, 01-447 Warsaw, Poland*

Abstract:
In knowledge discovery and data mining many measures of interestingness have been proposed in order to reveal different characteristics of the discovered knowledge patterns. Among these measures, an important role is played by Bayesian confirmation measures, which express in what degree a piece of evidence (premise) confirms a hypothesis (conclusion). This paper focuses on knowledge patterns in a form of "*if…, then…*" rules with a fixed conclusion. We are investigating the question of monotone relationship between a particular Bayesian confirmation measure on one side, and rule support and confidence, on the other side. We also provide general conclusions for the monotone link between any confirmation measure enjoying some desirable properties and rule support and confidence.

Keywords: Knowledge discovery; Association rules; Decision rule; Bayesian confirmation measures; Monotonicity in rule support, confidence and confirmation; Pareto-optimal rules

1. Introduction

In data mining and knowledge discovery, the discovered knowledge patterns are often expressed in a form of "*if…, then…*" rules. They are consequence relations representing correlation, association, causation etc. between independent and dependent attributes. Typically, the number of rules generated from massive datasets is very large, but only a few of them are likely to be useful for the domain expert analysing the data. In order to measure the relevance and utility of selected rules, quantitative measures, also known as attractiveness or interestingness measures (metrics), have been proposed and studied. Among them there are: confidence and support [1], gain [10], conviction [3], and many others.

Moreover, there exists an important group of interestingness measures called Bayesian confirmation measures. In general, Bayesian confirmation measures quantify the degree to which a piece of evidence built of the independent attributes provides "evidence for or against" or "support for or against" the hypothesis built of the dependent attributes [9]. Due to some valuable properties confirmation measure denoted in [9] and other studies by *f*, and confirmation measure *s* proposed by [6] became the subject of our particular analysis.

Bayardo and Agrawal in [2] have proved that for a class of rules with a fixed conclusion, the set of non-dominated, Pareto-optimal rules with respect to both rule support and confidence (i.e. the upper support-confidence Pareto-optimal border) includes optimal rules according to several different interestingness measures, such as gain, Laplace [7], lift [13], conviction, an unnamed measure proposed by Piatetsky-Shapiro [16]. This practically useful result allows to identify, the most interesting rules according to several interestingness measures by solving an optimised rule mining problem with respect to rule support and confidence only.

Due to the high utility of the scale of confirmation measures, an analytical verification of the monotone link between particular Bayesian confirmation measures on one side and rule support and confidence on the other becomes a problem worth investigating.

The paper is organized as follows. In the next section, there are preliminaries on rules and their quantitative description. In section 3, we briefly remind our previous analysis of monotonicity of confirmation measure *f* in rule support and confidence. Section 4 is devoted to verification whether confidence depends monotonically on rule support when the value of confirmation *f* is held fixed, and on confirmation measure *f* for a constant value of rule support. Section 5 concentrates on the analysis of monotone link between confirmation measure *s* and rule support and confidence. In section 6, we generalize the approach from sections 3 and 4 to a broader class of confirmation measures. The paper ends with conclusions.

2. Preliminaries

As discovering rules from data is the domain of inductive reasoning, its starting point is a sample of larger reality often given in a form of a data table. Formally, a *data table* is a pair $S = (U, A)$, where U is a nonempty finite set of objects called *universe*, and A is a nonempty finite set of *attributes* such that $a : U \to V_a$ for every $a \in A$. The set V_a is a domain of a. Let us associate a formal language L of logical formulas with every subset of attributes. Formulas for a subset $B \subseteq A$ are built up from attribute-value pairs (a,v), where $a \in B$ and $v \in V_a$, using logical connectives $\neg$ (not), $\wedge$ (and), $\vee$ (or). A *rule* induced from S and expressed in L is denoted by $\phi \to \psi$ (read as "*if* ϕ, *then* ψ"). It consists of antecedent ϕ and consequent ψ formulas in L, called *premise* and *conclusion*, respectively, and therefore it can be seen as a consequence relation (see critical discussion about interpretation of rules as logical implications in [12]) between premise and conclusion. The rules mined from data may be either *decision* rules or *association* rules, depending on whether the division of A into condition and decision attributes has been fixed or not.

2.1. Monotonicity

A function $g(x)$ is understood to be *monotone* (resp. *anti-monotone*) in x, if $x_1 < x_2$ implies that $g(x_1) \leq g(x_2)$ (resp. $g(x_1) > g(x_2)$).

2.2. Support and confidence measures of rules

With every rule induced from information table S measures called *support* and *confidence* are often associated. The *support* of condition ϕ, denoted as $sup(\phi)$, is equal to the number of objects in U having property ϕ. The support of rule $\phi \to \psi$, denoted as $sup(\phi \to \psi)$, is equal to the number of objects in U having both property ϕ and ψ; for those objects, both conditions ϕ and ψ evaluate to true.

The *confidence* of a rule (also called *certainty*), denoted as $conf(\phi \to \psi)$, is defined as follows:

$$conf\,(\phi \to \psi) = \frac{sup(\phi \to \psi)}{sup(\phi)}\,.$$

Of course, through out the whole paper it is assumed that the set of objects having property ϕ is not empty.

2.3. Bayesian confirmation measures

In general, Bayesian measures of confirmation quantify the strength of confirmation that premise ϕ gives to conclusion ψ. A confirmation measure denoted as $c(\phi, \psi)$ is required to satisfy the following definition:

$$c(\phi,\psi) = \begin{cases} > 0 \;\; if \;\; \Pr(\psi \mid \phi) > \Pr(\psi), \\ = 0 \;\; if \;\; \Pr(\psi \mid \phi) = \Pr(\psi), \\ < 0 \;\; if \;\; \Pr(\psi \mid \phi) < \Pr(\psi). \end{cases}$$

Many authors have considered desirable properties of confirmation measures. Fitelson and Eells in [8] have analysed properties of symmetry introduced by Carnap in [5] and concluded that the *hypothesis symmetry* is a desirable property whereas as *evidence*, *commutativity* and *total symmetries* are not.

Another profitable property of confirmation measures called property of monotonicity (M) has been proposed in [12]:

$$\text{(M)}\;\; c(\phi \to \psi) =$$
$$F[sup(\phi \to \psi), sup(\neg\phi \to \psi),$$
$$sup(\phi \to \neg\psi), sup(\neg\phi \to \neg\psi)]$$

is a function non-decreasing with respect to $sup(\phi \to \psi)$ and $sup(\neg\phi \to \neg\psi)$, and non-increasing with respect to $sup(\neg\phi \to \psi)$ and $sup(\phi \to \neg\psi)$.

2.3.1. Bayesian confirmation measure f and s

This paper concentrates on two well-known and widely studied confirmation measures denoted by f and s and defined as follows:

$$f\,(\phi \to \psi) = \frac{\Pr(\phi \mid \psi) - \Pr(\phi \mid \neg\psi)}{\Pr(\phi \mid \psi) + \Pr(\phi \mid \neg\psi)}\,,$$

$$s(\phi \to \psi) = \Pr(\psi \mid \phi) - \Pr(\psi \mid \neg\phi)\,.$$

Taking into account that conditional probability $\Pr(\mathbf{o} \mid *) = conf\,(\mathbf{o} \to *)$, confirmation measures f and s can be expressed as:

$$f\,(\phi \to \psi) = \frac{conf\,(\psi \to \phi) - conf\,(\neg\psi \to \phi)}{conf\,(\psi \to \phi) + conf\,(\neg\psi \to \phi)}\,,$$

$$s(\phi \to \psi) = conf\,(\phi \to \psi) - conf\,(\neg\phi \to \psi)\,.$$

Among authors advocating for confirmation measure f are Good [11], Pearl [15], Fitelson [9]. Measure s has been proposed by Christensen [6] and Joyce [14].

Both of those confirmation measures have been positively verified in [12] to have the property of monotonicity (M). Moreover, as shown in [8] they satisfy the property of hypothesis symmetry. The confirmation measures s, however, also satisfies the undesirable properties of evidence and total symmetry.

2.4. Partial preorder on rules in terms of two interestingness measures

Let us denote by $\mathbf{p}_{AB}$ a partial preorder on rules in terms of any two different interestingness measures A and B. Recall that a *partial preorder* on a set X is any binary relation R on X that is reflexive (i.e. for all $x \in X$, xRx) and transitive (i.e. for all $x,y,z \in X$, xRy and yRz imply xRz). In simple words, if the semantics of xRy is "x is at most as good as y", then a complete preorder permits to order the elements of X from the best to the worst, with possible ex-aequo (i.e. cases of $x,y \in X$ such that xRy and yRx) and with possible incomparability (i.e. cases of $x,y \in X$ such that *not* xRy and *not* yRx). The partial order $\mathbf{p}_{AB}$ can be decomposed into its asymmetric part $\mathbf{p}_{AB}$ and its symmetric part $\sim_{AB}$ in the following manner:

given two rules r_1 and r_2, $r_1 \ \mathbf{p}_{AB} \ r_2$ if and only if

$A(r_1) \leq A(r_2) \wedge B(r_1) < B(r_2)$, or

$A(r_1) < A(r_2) \wedge B(r_1) \leq B(r_2)$,

moreover $r_1 \sim_{AB} r_2$ if and only if

$A(r_1) = A(r_2) \wedge B(r_1) = B(r_2)$.

2.5. Implication of a total order $\underline{\mathbf{p}}_t$ by partial preorder $\mathbf{p}_{AB}$

Application of measures such as gain, Laplace, lift, conviction, measure proposed by Piatetsky-Shapiro, or confirmation measures *f* and *s* etc. results in a total order on the set of rules, ordering them according to their interestingness value.

A total order $\underline{\mathbf{p}}_t$ *is implied by* a partial preorder $\underline{\mathbf{p}}_{AB}$ if:

$r_1 \ \mathbf{p}_{AB} \ r_2 \Rightarrow r_1 \ \mathbf{p}_t \ r_2$, *and*

$r_1 \sim_{AB} r_2 \Rightarrow r_1 \sim_t r_2$.

It has been proved by Bayardo and Agrawal in [2] that if a total order $\underline{\mathbf{p}}_t$ is implied by a particular support-confidence partial order $\underline{\mathbf{p}}_{sc}$, then the optimal rules with respect to $\underline{\mathbf{p}}_t$ can be found in the set of non-dominated rules with respect to rule support and confidence. Thus, having proved that a total order defined over a new interestingness measure is implied by $\underline{\mathbf{p}}_{sc}$, one can concentrate on discovering non-dominated rules with respect to rule support and confidence. Moreover, Bayardo and Agrawal have shown in [2] that the following conditions are sufficient for proving that a total order $\underline{\mathbf{p}}_t$ defined over a rule value function *g*(r) is implied by partial order $\underline{\mathbf{p}}_{AB}$:

- *g*(r) is monotone in *A* over rules with the same value of *B*, and
- *g*(r) is monotone in *B* over rules with the same value of *A*.

3. Analysis of the monotonicity of confirmation measure *f* in rule support and confidence

Due to the semantic importance of confirmation measure *f*, a verification of existence of a monotone link between *f* and, rule support and confidence has been conducted.

Theorem 1. [4] Confirmation measure *f* is independent of rule support, and, therefore, monotone in rule support, when the value of confidence is held fixed.

Proof. Let us consider the confirmation measure *f* transformed such that, for given *U* and ψ, it only depends on confidence of rule $\phi \to \psi$ and support of ψ:

$$f(\phi \to \psi) = \frac{|U| conf(\phi \to \psi) - sup(\psi)}{(|U| - 2sup(\psi)) conf(\phi \to \psi) + sup(\psi)}.$$

As we consider rules with a fixed conclusion ψ, the values of |*U*| and *sup*(ψ) are constant. Thus, for a fixed confidence, we have a constant value of the confirmation measure *f*, no matter what the rule support is. Hence, confirmation measure *f* is monotone in rule support when the confidence is held constant. □

Theorem 2. [4] Confirmation measure *f* is increasing in confidence, and, therefore, monotone in confidence.

Proof. Again, let us consider confirmation measure *f* given in the same form as above:

$$f(\phi \to \psi) = \frac{|U| conf(\phi \to \psi) - sup(\psi)}{(|U| - 2sup(\psi)) conf(\phi \to \psi) + sup(\psi)}$$

For the clarity of the presentation, let us express the above formula as a function of confidence, still regarding |*U*| and *sup*(ψ) as constant values greater than 0:

$$y = \frac{kx - m}{nx + m},$$

where $y = f(\phi \to \psi)$, $x = conf(\phi \to \psi)$, $k = |U|$, $m = sup(\psi)$, $n = |U| - 2sup(\psi)$.

It is easy to observe that:

- $k = |U| > 0$, and
- $0 < m \leq |U|$.

In order to verify the monotonicity of *f* in confidence, let us differentiate *y* with respect to *x*. We obtain:

$$\frac{\partial y}{\partial x} = \frac{m(k+n)}{(nx+m)^2}.$$

As $m > 0$, and $k + n = |U| + |U| - 2sup(\psi) = 2|U| - 2sup(\psi) \geq 0$ for $|U| \geq sup(\psi)$, the whole derivative is always not smaller than 0. Therefore, confirmation measure *f* is monotone in confidence. □

As proved in [2], mining the support-confidence border identifies optimal rules according to several different interestingness metrics. The above results show that even confirmation measure *f* is among those interestingness metrics and, therefore, all rules maximizing confirmation measure *f* can be found on the Pareto-optimal support-confidence border (concerning rules with a fixed conclusion).

4. Analysis of the monotonicity of confidence in rule support and confirmation measure *f*

It has been proved in [4] that rules maximizing confirmation measure *f* can be found on the Pareto-optimal support-confidence border. However, the utility of confirmation measure *f* outranks the utility of confidence, which has no means to show, that the rule is useless when its premise disconfirms the conclusion. Such situation is expressed by a negative value of any confirmation measure, thus useless rules can be filtered out simply by observing the confirmation measure's sign. Therefore, there arises a need to analyse whether a Pareto-optimal border with respect to rule support and confirmation measure *f* would contain rules optimal in confidence. In order to verify that an analysis of the monotonicity of confidence:

- in rule support for a fixed value of confirmation *f*, as well as
- in confirmation *f* for a fixed value of support

has been performed.

Theorem 3. Confidence is independent of rule support, and thus monotone in rule support, for a fixed value of confirmation measure *f*.

Proof.

The definition of the confirmation measure *f* can be transformed to outline how confidence depends on confirmation measure *f*:

$$conf(\phi \to \psi) = \frac{f(\phi \to \psi)sup(\psi) + sup(\psi)}{|U| - f(\phi \to \psi)(|U| - 2sup(\psi))} \qquad (1)$$

As we consider a set of rules with a fixed conclusion ψ, the values of $|U|$ and $sup(\psi)$ are constant. Thus, for a fixed confirmation *f*, we have a constant value of confidence, no matter what the rule support is. Hence, we can conclude that confidence is monotone in rule support when the confirmation measure *f* is held constant. □

Now, let us verify whether confidence is monotone in confirmation measure *f* for a constant value of rule support.

Theorem 4. Confidence is monotone in confirmation measure *f*.

Proof.

For the simplicity of the presentation, let us express formula (1) in the following manner:

$$y = \frac{mx + m}{lx + k}$$

where $y = conf(\phi \to \psi)$, $x = f(\phi \to \psi)$, $k = |U|$, $l = -(|U| - 2sup(\psi))$, $m = sup(\psi)$.

As we consider a set of rules with a fixed conclusion ψ, the values of $|U|$ and $sup(\psi)$ are constant, and therefore *m*, *l* and *k* are also constant.

In order to verify the monotonicity of *y* in *x*, let us now derive *y* with respect to *x*. We obtain:

$$\frac{\partial y}{\partial x} = \frac{m(k-l)}{(lx+k)^2}.$$

Since

$$m(k-l) = sup(\psi)(|U| + |U| - 2sup(\psi)) =$$
$$= 2sup(\psi)(|U| - sup(\psi))$$

and knowing that $0 < |U| \geq sup(\psi) > 0$

the derivative is always not smaller than 0, thus *y* (i.e. $conf(\phi \to \psi)$) is monotone in *x* (i.e. $f(\phi \to \psi)$). □

Thus, both of Bayardo and Agrawal's sufficient conditions for proving that a total order defined over confidence is implied by support-confirmation measure *f* partial order are held. This means that, for a class of rules with a fixed conclusion, rules optimal according to confidence will be found in the set of rules that are best with respect to both rule support and confirmation measure *f*.

Moreover *Theorem 1*, *Theorem 2*, *Theorem 3* and *Theorem 4* show that the set of rules located on the support-confidence Pareto-optimal border is exactly the same as the set of rules located on the support-confirmation measure *f* Pareto-optimal border. It is straightforward to observe that interestingness measures that are monotone in confidence, must also be monotone in confirmation measure *f*, due to the monotone link between confidence and confirmation measure *f*. Hence, all the interestingness measures that were found on the support-confidence Pareto-optimal border shall also reside on the Pareto-optimal border with respect to rule support and confirmation measure *f*.

However, any non-dominated rule with a negative value of confirmation measure *f* must be discarded from further analysis as its premise only disconfirms the conclusion. In particular, if the highest value of confirmation measure *f* in the Pareto set is negative, then the whole set should be excluded as it does not contain any interesting rules. A final set of rules representing patterns discovered from the whole dataset shall be a union of all the non-negative-in-*f* rules from all the Pareto-optimal borders (all possible conclusions) with respect to rule support and confirmation measure *f*.

Concluding, due to the fact that the utility of scale of confirmation measure *f* outranks the scale of confidence, we are strongly in favour of mining the Pareto-optimal border with respect to rule support and confirmation *f* and not rule support and confidence as it was proposed in [2].

5. Analysis of the monotonicity of confirmation measure *s* in rule support and confidence

In the group of Bayesian confirmation measures, an important role is also played by confirmation measure *s*. Its monotone link (within a set of rules with the same conclusion) with confidence and rule support came into the scope of our analysis.

For the clarity of further presentation, let us use the following notation:

$$a = sup(\phi \to \psi),$$
$$b = sup(\neg\phi \to \psi),$$
$$c = sup(\phi \to \neg\psi),$$
$$d = sup(\neg\phi \to \neg\psi).$$

and make an ongoing assumption that *a*, *b*, *c* and *d* are positive numbers.

The confirmation measure *s* is then defined as:

$$s(\phi \to \psi) = \frac{a}{a+c} - \frac{b}{b+d}.$$

The analysis considers only a set of rules with the same conclusion, thus the values of $|U| = a+b+c+d$ and $sup(\psi) = a+b$ are constant.

An analysis of the monotonicity of confirmation measure *s*:

- in confidence for a fixed value of rule support, and
- in rule support for a fixed value of confidence

has been performed.

Theorem 5. When the rule support value is held fixed, then confirmation measure *s* is increasing with respect to confidence (i.e. confirmation measure *s* is monotone in confidence).

Proof.

Let us consider the confirmation measure *s*:

$$s(\phi \to \psi) = \frac{a}{a+c} - \frac{b}{b+d}.$$

For the hypothesis, $sup(\phi \to \psi) = a$ is constant. Therefore, it is clear that $conf(\phi \to \psi) = \frac{a}{a+c}$ can only increase with the decrease of *c*. Hence, let us consider $c' = c - \Delta$, where $\Delta > 0$. Now, operating on c' the only way to guarantee that $|U|$ and $sup(\psi)$ still remain constant is to increase *d* such that $d' = d + \Delta$. The values of *a* and *b* cannot change: $a' = a$ and $b' = b$. Now, the new value of confirmation measure *s* takes the following form:

$$s'(\phi \to \psi) = \frac{a'}{a'+c'} - \frac{b'}{b'+d'} = \frac{a}{a+c-\Delta} - \frac{b}{b+d+\Delta}.$$

Since $\Delta > 0$, it is clear that $s'(\phi \to \psi) > s(\phi \to \psi)$. This means that for a fixed value of rule support, increasing confidence results in an increase of the value of confirmation measure *s* and therefore confirmation measure *s* is monotone with respect to confidence.□

Theorem 6. When the confidence value is held fixed, then:

a) confirmation measure *s* is increasing in rule support (i.e. monotone) if and only if *s*>0,

b) confirmation measure *s* is constant in rule support (i.e. monotone) if and only if *s*=0,

c) confirmation measure *s* is decreasing in rule support (i.e. anti-monotone) if and only if *s*<0.

Let us present the proof only of *Theorem 6.a* as the other points are analogous.

Proof.

Let us consider the confirmation measure *s*:

$$s(\phi \to \psi) = \frac{a}{a+c} - \frac{b}{b+d}.$$

Let us consider an increase of $sup(\phi \to \psi) = a$ expressed in the form of $a' = a + \Delta$, where $\Delta > 0$.

Since $conf(\phi \to \psi) = \frac{a}{a+c}$ is to be constant, thus *c* should change into $c' = c + \varepsilon$ in such a way that:

$$conf(\phi \to \psi) = \frac{a}{a+c} = \frac{a'}{a'+c'} = \frac{a+\Delta}{a+\Delta+c+\varepsilon}.$$

Simple mathematical transformation lead to the conclusion that:

$$\frac{a}{a+c} = \frac{a+\Delta}{a+\Delta+c+\varepsilon} \Leftrightarrow \frac{\Delta}{\Delta+\varepsilon} = \frac{a}{a+c} \tag{2}$$

Let us observe that (2) implies that if $c = 0$ then $\varepsilon = 0$ and moreover if $c > 0$ then $\varepsilon > 0$. Since $|U|$ and $sup(\psi)$ must be kept constant, *b* and *d* need to decrease in such a way that $b' = b - \Delta$ and $d' = d - \varepsilon$. In this situation, the new confirmation measure *s* will be:

$$s'(\phi \to \psi) = \frac{a'}{a'+c'} - \frac{b'}{b'+d'} =$$
$$= \frac{a+\Delta}{a+\Delta+c+\varepsilon} - \frac{b-\Delta}{b-\Delta+d-\varepsilon}$$

Remembering that $conf(\phi \to \psi) = \frac{a}{a+c}$ is constant, let us observe that:

$$s'(\phi \to \psi) > s(\phi \to \psi) \Leftrightarrow \frac{b}{b+d} > \frac{b-\Delta}{b-\Delta+d-\varepsilon} \Leftrightarrow$$
$$\Leftrightarrow d\Delta > b\varepsilon \Leftrightarrow d\Delta + b\Delta > b\varepsilon + b\Delta \Leftrightarrow$$
$$\Leftrightarrow \frac{\Delta}{\Delta+\varepsilon} > \frac{b}{b+d} \tag{3}$$

Considering (2) and (3) it can be concluded that:

$$s'(\phi \to \psi) > s(\phi \to \psi) \Leftrightarrow \frac{a}{a+c} > \frac{b}{b+d} \Leftrightarrow$$
$$\Leftrightarrow s(\phi \to \psi) > 0.$$

Thus, it proves that, for a fixed value of confidence, confirmation measure *s* is increasing with respect to rule support if and only if $s(\phi \to \psi) > 0$ and therefore in its positive range confirmation measure *s* is monotone in rule support.□

As rules with negative values of confirmation measure *s* should always be discarded from consideration, the result from *Theorem 6* states the monotone relationship just in the interesting subset of rules.

Since confirmation measure *s* is monotone in confidence when the value of rule support is held fixed, and monotone in rules support when the value of confidence remains unchanged, we propose to generate interesting rules by searching for rules maximizing confirmation measure *s* and support, i.e. substituting the confidence in the support-confidence Pareto-optimal border with the confirmation measure *s* and obtaining in this way a support-confirmation-*s* Pareto-optimal border. This approach differs from the idea of finding the Pareto-optimal border according to rule support and confirmation measure *f*, because support-confirmation-*f* Pareto-optimal border contains the same rules as the support-confidence Pareto-optimal border, while in general support-confirmation-*s* Pareto-optimal border does not.

6. Analysis of the monotonicity of any confirmation measure having the property of monotonicity (M) in rule support and confidence

The investigation of monotone link with confidence and rule support has also been extended to a more general class of all the confirmation measures that have the property of monotonicity (M). For a set of rules with a fixed conclusion a general analysis has been conducted verifying under what conditions a confirmation measure with the property of monotonicity (M):

- is monotone in confidence when the value of rule support is kept unchanged,
- is monotone in rule support when the value of confidence is held fixed.

Again for the simplicity of presentation let us use the following notation:

$$a = sup(\phi \to \psi),$$
$$b = sup(\neg\phi \to \psi),$$
$$c = sup(\phi \to \neg\psi),$$
$$d = sup(\neg\phi \to \neg\psi).$$

Let us consider a Bayesian confirmation measure $F(a, b, c, d)$ being differentiable and having the property of monotonicity (M). The analysis concerns only a set of rules with the same conclusion, thus the values of $|U| = a + b + c + d$ and $sup(\psi) = a + b$ are constant.

One can observe that *a*, *b*, *c*, and *d* can be transformed in the following way:

$$a = sup(\phi \to \psi),$$
$$b = sup(\psi) - sup(\phi \to \psi),$$
$$c = \frac{1}{conf(\phi \to \psi)} sup(\phi \to \psi) - sup(\phi \to \psi),$$
$$d = |U| - sup(\psi) - \frac{1}{conf(\phi \to \psi)} sup(\phi \to \psi) + sup(\phi \to \psi).$$

Then, a Bayesian confirmation measure *F* can be expressed as:

$$F(a,b,c,d) = F(sup(\phi \to \psi),\ sup(\psi) - sup(\phi \to \psi),\ \frac{1}{conf(\phi \to \psi)} sup(\phi \to \psi) - sup(\phi \to \psi),\ |U| - sup(\psi) - \frac{1}{conf(\phi \to \psi)} sup(\phi \to \psi) + sup(\phi \to \psi)).$$

Theorem 7. When the value of rule support is held fixed, then the confirmation measure $F(a, b, c, d)$ is monotone in confidence.

Proof.
Let us presume that $sup(\phi \to \psi)$ is constant. Now, let us derive $F(a, b, c, d)$ with respect to $\frac{1}{conf(\phi \to \psi)}$ and obtain:

$$\frac{\partial F}{\partial\left(\frac{1}{conf(\phi \to \psi)}\right)} = \frac{\partial F}{\partial c} sup(\phi \to \psi) - \frac{\partial F}{\partial d} sup(\phi \to \psi) = sup(\phi \to \psi)\left(\frac{\partial F}{\partial c} - \frac{\partial F}{\partial d}\right)$$

Since *F* is supposed to satisfy the property of monotonicity (M), it must be non-increasing with respect to *c* and non-decreasing with respect to *d*, such that $\frac{\partial F}{\partial c} \le 0$ and $\frac{\partial F}{\partial d} \ge 0$. Thus, $\frac{\partial F}{\partial c} - \frac{\partial F}{\partial d} \le 0$. Remembering, moreover, that $sup(\phi \to \psi) > 0$, we obtain:

$$\frac{\partial F}{\partial\left(\frac{1}{conf(\phi \to \psi)}\right)} \le 0$$

which brings the conclusion that $\frac{1}{conf(\phi \to \psi)}$ is non-increasing with respect to $conf(\phi \to \psi)$.

Thus, $\frac{\partial F}{\partial conf(\phi \to \psi)} \ge 0$, which means that *F* is non-decreasing with respect to $conf(\phi \to \psi)$. □

Theorem 8. When the value of confidence is held fixed, then the confirmation measure $F(a, b, c, d)$ is monotone in rule support if:

$$\frac{\partial F}{\partial c} = \frac{\partial F}{\partial d} = 0 \text{ or } \frac{\frac{\partial F}{\partial a} - \frac{\partial F}{\partial b}}{\frac{\partial F}{\partial d} - \frac{\partial F}{\partial c}} \ge \frac{1}{conf(\phi \to \psi)} - 1 .$$

Proof.
Let us presume that $conf(\phi \to \psi)$ is constant. Let us derive $F(a, b, c, d)$ with respect to $sup(\phi \to \psi)$ Then, we obtain:

$$\frac{\partial F}{\partial sup(\phi \to \psi)} =$$

$$= \frac{\partial F}{\partial a} - \frac{\partial F}{\partial b} + \left(\frac{\partial F}{\partial c} - \frac{\partial F}{\partial d}\right)\left(\frac{1}{conf(\phi \to \psi)} - 1\right)$$

Since F is supposed to satisfy the property of monotonicity (M), it must be non-increasing with respect to b, c and non-decreasing with respect to a, d, such that: $\frac{\partial F}{\partial b} \le 0$, $\frac{\partial F}{\partial c} \le 0$ and $\frac{\partial F}{\partial a} \ge 0$, $\frac{\partial F}{\partial d} \ge 0$.

Hence, if $\frac{\partial F}{\partial c} = \frac{\partial F}{\partial d}$, then $\frac{\partial F}{\partial sup(\phi \to \psi)} \ge 0$.

It is clear, that due to the property of monotonicity (M) of F, $\frac{\partial F}{\partial c} = \frac{\partial F}{\partial d}$ if and only if $\frac{\partial F}{\partial c} = \frac{\partial F}{\partial d} = 0$.

Let us observe moreover that, if $\frac{\partial F}{\partial c} \ne \frac{\partial F}{\partial d}$, then

$$\frac{\partial F}{\partial sup(\phi \to \psi)} \ge 0 \Leftrightarrow \frac{\frac{\partial F}{\partial a} - \frac{\partial F}{\partial b}}{\frac{\partial F}{\partial d} - \frac{\partial F}{\partial c}} \ge \frac{1}{conf(\phi \to \psi)} - 1 . \square$$

Theorem 7 states that for a set of rules with the same conclusion, any Bayesian confirmation measure satisfying the property of monotonicity (M) is always non-decreasing with respect to confidence when the value of rule support is kept fixed. Moreover, due to *Theorem 8*, all those confirmation measures that are independent of $sup(\phi \to \neg\psi)$ and $sup(\neg\phi \to \neg\psi)$ are always found monotone in rule support when the value of confidence remains unchanged. However, for a constant value of confidence, Bayesian confirmation measures which do depend on the value of $sup(\phi \to \neg\psi)$ and $sup(\neg\phi \to \neg\psi)$ are also non-decreasing with respect to rule support if and only if they satisfy the following condition:

$$\frac{\frac{\partial F}{\partial a} - \frac{\partial F}{\partial b}}{\frac{\partial F}{\partial d} - \frac{\partial F}{\partial c}} \ge \frac{1}{conf(\phi \to \psi)} - 1 .$$

The general analysis in *Theorem 7* and *Theorem 8* outlines an easy method of verification whether there exists a monotone link between any Bayesian confirmation measure with the property of monotonicity (M), and confidence and rule support, respectively.

7. Conclusions

Bayardo and Agrawal have proved in [2] that total orders of many interestingness measures such as gain, Laplace, lift, conviction, the one proposed by Piatetsky-Shapiro, etc. are implied by the rule support-confidence partial order. This practically useful result showed that the most-interesting rules according to any of the above measures, are included in the set of Pareto-optimal rules with respect to both rule support and confidence.

In this paper, for a class of rules with the same conclusion, we have analysed the monotone relationship between confidence and confirmation measure f and rule support. Moreover, we have verified the monotonicity of two Bayesian confirmation measures: f and s in rule support when the value of confidence is held fixed, and in confidence when the value of rule support remained unchanged. Those particular measures came into the scope of our interest for their valuable properties and the utility of scale of Bayesian confirmation measures in general.

The analysis has also been extended to a more general class of all the confirmation measures that have the property of monotonicity (M). As the result, precise conditions in which such confirmation measures are monotone in rule support and confidence were presented.

The overall results constitute the analytical bases for further research devoted to proposing new approaches to looking for interesting rules. Consequently, our future work will concentrate on developing new ways of mining attractive rules and on preparing algorithms realizing them.

References

[1] Agrawal, R., Imielinski, T. and Swami, A.: Mining Associations between Sets of Items in Massive Databases. [In]: *Proc. of the 1993 ACM-SIGMOD Int'l Conf. on Management of Data*, 207-216 (1993).

[2] Bayardo, R.J., Agrawal, R.: Mining the Most Interesting Rules. [In]: *Proc. of the Fifth ACM-SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*, 145-154 (1999).

[3] Brin, S., Motwani, R., Ullman, J. and Tsur, S.: Dynamic Itemset Counting and Implication Rules for Market Basket Data. [In]: *Proc. of the 1997 ACM-SIGMOD Int'l Conf. on the Management of Data*, 255-264 (1997).

[4] Brzezińska, I., Słowiński, R.: Monotonicity of a Bayesian confirmation measure in rule support and confidence. [In]: T. Burczyński, W. Cholewa, W. Moczulski (Eds.) *Recent Developments in Artificial Intelligence Methods*, AI-METH Series, Gliwice, 39-42 (2005).

[5] Carnap, R.: *Logical Foundations of Probability*, 2nd Edition. University of Chicago Press, Chicago (1962).

[6] Christensen, D.: Measuring confirmation. *Journal of Philosophy XCVI*, 437-461 (1999).

[7] Clark, P. and Boswell, P.: Rule Induction with CN2: Some Recent Improvements. [In]: *Machine Learning: Proc. of the Fifth European Conference*, 151-163 (1991).

[8] Eells, E., Fitelson, B.: Symmetries and asymmetries in evidential support. *Philosophical Studies*, 107 (2): 129-142 (2002).

[9] Fitelson, B.: *Studies in Bayesian Confirmation Theory*. Ph.D. Thesis, University of Wisconsin, Madison (2001).

[10] Fukuda, T., Morimoto, Y., Morishita, S. and Tokuyama, T.: Data Mining using Two-Dimensional

Optimized Association Rules: Scheme, Algorithms, and Visualization. [In]: *Proc. of the 1996 ACM-SIGMOD Int'l Conf. on the Management of Data*, 13-23 (1996).

[11] Good, I.J.: The best explicatum for weight of evidence. *Journal of Statistical Computation and Simulation*, 19: 294-299 (1984).

[12] Greco, S., Pawlak, Z. and Słowiński, R.: Can Bayesian Confirmation measures be useful for rough set decision rules? *Engineering Applications of Artificial Intelligence*, 17: 345-361 (2004).

[13] International Business Machines: *IBM Intelligent Miner User's Guide*, Version 1, Release 1 (1996).

[14] Joyce, J.: The Foundations of Causal Decision Theory. Cambridge University Press, Cambridge (1999).

[15] Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, San Francisco (1988).

[16] Piatetsky-Shapiro, G.: Discovery, Analysis, and Presentation of Strong Rules. Chapter 12 [In]: *Knowledge Discovery in Databases*, AAAI/MIT Press (1991).