Analysis of monotonicity properties of some rule interestingness measures

Salvatore Greco[1], **Roman Słowiński**[2,3], **Izabela Szczęch**[2]

Abstract: One of the crucial problems in the field of knowledge discovery is development of good interestingness measures for evaluation of the discovered patterns. In this paper, we consider quantitative, objective interestingness measures for "*if..., then...*" association rules. We focus on three popular interestingness measures being *rule interest function* of Piatetsky-Shapiro, *gain measure* of Fukuda et al., and *dependency factor* of Pawlak. We verify whether they satisfy valuable property (M) of monotonic dependency on the number of objects satisfying or not the premise or the conclusion of a rule, and property of hypothesis symmetry (HS). Moreover, analytically and through experiments we show an interesting relationship between those measures and two other commonly used measures of rule support and anti-support.

Keywords: Association rules, Piatetsky-Shapiro's Rule Interest Function, Gain Measure, Dependency Factor, Support, Anti-support, Pareto-optimal border

1. Introduction

It has been recognized early on in the knowledge discovery literature, that the number of knowledge patterns, often expressed in a form of "*if..., then...*" rules, discovered in databases can be quite large, and that only a small portion of them is actually useful for the user. To address the problem of evaluation of attractiveness of the mined rules, various quantitative measures of interestingness have been defined and studied (e.g. support, confidence, anti-support, gain, rule interest function, lift). They all reflect different characteristics of rules. However, the issue of studying and analyzing relationships between various measures, has not yet been fully investigated. Moreover, there is a need for verification whether particular interestingness measures satisfy some valuable features, which reflect the users' expectations towards the behavior of the measures in particular situations. For example, one may expect that the measure he uses will increase its value for a certain rule (or at least will not decrease) when the number of objects in the dataset that support the rule increases. It is, of course, quite intuitively understood property of the measure, however it draws well our attention to the properties of the applied measures. Studies verifying whether popular interestingness measures possesses

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

valuable features would widen our understanding of those measures and of their applicability. Moreover, through such property analysis one can also learn about relationships between different measures.

In this paper, we focus on three well-known measures: rule interest function proposed by Piatetsky-Shapiro (1991), gain measure of Fukuda *et al.* (1996) and dependency factor introduced by Pawlak (2004). We investigate whether they possess a useful feature called the property (M) introduced by Greco *et al.* (2004), and hypothesis symmetry (HS) advocated by Eells *et al.* (2002) and Fitelson (2001). Moreover, on the basis of satisfying the property (M), we draw some conclusions about very particular relationship between rule interest and gain, and two other simple but meaningful measures being rule support and anti-support.

In order to achieve the above objectives, the rest of the paper is organized as follows. In section 2, there are preliminaries on rules and their quantitative description. In section 3, we verify analytically whether rule interest function, gain measure and dependency factor have the analyzed property (M). In section 4, we investigate the relationship between the first two measures and the Pareto-optimal border with respect to support and anti-support. Illustration of the results on a real life dataset is presented to support the theoretical considerations with experimental results. Next, in section 5, we analyze if rule interest function, gain measure and dependency factor satisfy the hypothesis symmetry. The paper ends with conclusions.

2. Preliminaries

Let us consider information table $S = (U, A)$, where U and A are finite, non-empty sets called *universe* and *set of attributes*, respectively. One can associate a formal language L of logical formulas with every subset of attributes. A rule induced from S is denoted by $\phi \rightarrow \psi$ (read as "*if* ϕ, *then* ψ") and consists of condition and decision formulas, called premise and conclusion, respectively.

2.1. Support and Anti-support Measures of Rules

One of the most popular measures used to identify frequently occurring association rules in sets of items from information table S is *support*. 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 $\phi \rightarrow \psi$ (also simply referred to as support), denoted as $sup(\phi \rightarrow \psi)$, is the number of objects in U having property ϕ and ψ. Thus, it corresponds to statistical significance (Hilderman *et al.*, 2001).

Anti–support of a rule $\phi \rightarrow \psi$ (also simply referred to as anti-support), denoted as $anti{-}sup(\phi \rightarrow \psi)$, is equal to the number of objects in U having the property ϕ but not having the property ψ. Thus, anti-support is the number of counter-examples, i.e. objects for which the premise ϕ evaluates to true but which fall into a class different than ψ. Note that anti-support can also be regarded as $sup(\phi \rightarrow \neg\psi)$.

2.2. Piatetsky-Shapiro's Rule Interest Function, Gain and Dependency Factor

The *rule interest function RI* introduced by Piatetsky-Shapiro (1991) is used to quantify the correlation between the premise and conclusion. It is given by the following formula:

$$RI(\phi \to \psi) = sup(\phi \to \psi) - \frac{sup(\psi)sup(\phi)}{|U|} \tag{1}$$

For rule $\phi \to \psi$, when $RI = 0$, then ϕ and ψ are statistically independent and thus, such rule should be considered as uninteresting. When $RI > 0$ ($RI < 0$), then there is a positive (negative) correlation between ϕ and ψ (Hilderman *et al.*, 2001).

The *gain* function of Fukuda *et al.* (1996) is defined in the following manner:

$$gain(\phi \to \psi) = sup(\phi \to \psi) - \Theta sup(\phi) \tag{2}$$

where Θ is a fraction constant between 0 and 1. Note that, for a fixed value of $\Theta = sup(\psi)/|U|$, the gain measure becomes identical to the above rule interest function RI.

The *dependency factor* of Pawlak (2004) is defined in the following manner:

$$\eta(\phi \to \psi) = \frac{\frac{sup(\phi \to \psi)}{sup(\phi)} - \frac{sup(\psi)}{|U|}}{\frac{sup(\phi \to \psi)}{sup(\phi)} + \frac{sup(\psi)}{|U|}} \tag{3}$$

The dependency factor expresses the degree of dependency, and can be seen as a counterpart of correlation coefficient used in statistics. When ϕ and ψ are independent on each other, then $\eta(\phi \to \psi) = 0$. If $-1 < \eta(\phi \to \psi)$, then ϕ and ψ are negatively dependent, and if $0 < \eta(\phi \to \psi) < 1$, then ϕ and ψ are positively dependent on each other.

2.3. Property of monotonicity (M) and Hypothesis Symmetry (HS)

Greco *et al.* (2004) have considered Bayesian confirmation measures from the viewpoint of their usefulness for measuring interestingness of decision rules. They claim that confirmation measures should enjoy a valuable *property (M)* describing monotonic dependency on the number of objects satisfying or not the premise or the conclusion of the rule. The property was introduced in Greco *et al.* (2004) where it was defined as follows:

$$F = [sup(\phi \to \psi), sup(\neg\phi \to \psi), sup(\phi \to \neg\psi), sup(\neg\phi \to \neg\psi)] \tag{4}$$

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 $sup(\phi \to \neg\psi)$.

The property (M) with respect to $sup(\phi \to \psi)$ (or, analogously, with respect to $sup(\neg\phi \to \neg\psi)$) means that any evidence in which ϕ and ψ (or, analogously, neither ϕ nor ψ) hold together increases (or at least does not decrease) the credibility of the rule $\phi \to \psi$. On the other hand, the property (M) with respect to $sup(\neg\phi \to \psi)$ (or, analogously, with respect to $sup(\phi \to \neg\psi)$) means that any evidence in which ϕ does not hold and ψ holds (or, analogously, ϕ holds and ψ does not hold) decreases (or at least does not increase) the credibility of the rule $\phi \to \psi$.

Eells *et al.* (2002) have analysed some confirmation measures from the viewpoint of four properties of symmetry introduced by Carnap (1962). Considering an interestingness measure $c(\phi \to \psi)$, the considered symmetries were defined as follows:

- evidence symmetry (ES): $c(\phi \to \psi) = -c(\neg\phi \to \psi)$
- commutativity symmetry (CS): $c(\phi \to \psi) = c(\psi \to \phi)$
- hypothesis symmetry (HS): $c(\phi \to \psi) = -c(\phi \to \neg\psi)$
- total symmetry (TS): $c(\phi \to \psi) = c(\neg\phi \to \neg\psi)$

It has been concluded in Eells *et al.* (2002) that, in fact, only (HS) is a desirable property, while (ES), (CS) and (TS) are not. The meaning behind the hypothesis symmetry is that the significance of the premise with respect to the conclusion part of a rule should be of the same strength, but of the opposite sign, as the significance of the premise with respect to a negated conclusion.

Both, the property (M) as well as the hypothesis symmetry were introduced in the perspective of confirmation measures, however there are no adversities for applying them to any other interestingness measure, in particular RI, gain or dependency factor. In fact, we believe that these properties are valuable and it is worth verifying which commonly used interestingness measures really do have them.

2.4. Partial Preorder on Rules in terms of Rule Support and Anti-support

Let us denote by $\preceq_{s\neg a}$ a partial preorder given by the dominance relation on a set X of rules in terms of two interestingness measures: support and anti-support, i.e. given a set of rules X and two rules $r_1, r_2 \in X, r_1 \prec_{s\neg a} r_2$ if and only if

$$sup(r_1) \leq sup(r_2) \ \wedge \ anti-sup(r_1) \geq anti-sup(r_2) \tag{5}$$

Recall that a partial preorder on a set X is a binary relation R on X that is reflexive and transitive. The partial preorder $\preceq_{s\neg a}$ can be decomposed into its asymmetric part $\prec_{s\neg a}$ and its symmetric part $\sim_{s\neg a}$ in the following manner: given a set of rules X and two rules $r_1, r_2 \in X, r_1 \prec_{s\neg a} r_2$ if and only if

$$\begin{array}{c} sup(r_1) \leq sup(r_2) \ \wedge \ anti-sup(r_1) > anti-sup(r_2), \text{ or} \\ sup(r_1) < sup(r_2) \ \wedge \ anti-sup(r_1) \geq anti-sup(r_2) \end{array} \tag{6}$$

moreover, $r_1 \sim_{s\neg a} r_2$ if and only if

$$sup(r_1) = sup(r_2) \ \wedge \ anti - sup(r_1) = anti - sup(r_2) \tag{7}$$

If for a rule $r \in X$ there does not exist any rule $r' \in X$, such that $r \prec_{s\neg a} r'$ then r is said to be non–dominated (i.e. Pareto–optimal) with respect to support and anti-support. A set of all non-dominated rules forms a *Pareto-optimal border* of the set of rules in the evaluation space. A set of all non-dominated rules with respect to support and anti-support will be called a *support-anti-support Pareto-optimal border*. In other words, it is the set of rules such that there is no other rule having greater support and smaller anti-support.

The approach to evaluation of the set of rules in terms of two interestingness measures being rule support and anti-support was proposed and presented in detail in Brzezińska *et al.* (2007). The idea of combining those two dimensions came as a result of looking for a set of rules that would include all rules optimal with respect to any confirmation measure with the desirable property (M) (Greco *et al.*, 2004). It was proved by Brzezińska *et al.* (2007) that the best rule according to any of confirmation measures with (M) must reside on the support-anti-support Pareto-optimal border. Though the theorem was particularly intended for confirmation measures, it concerns, in fact, any measure that 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 $sup(\phi \to \neg\psi)$. Therefore, we can consider satisfying of the property of monotonicity (M) by a measure as a sufficient condition for stating that rules optimal with respect to this measure will be found on the support-anti-support Pareto-optimal border. It is a valuable result as it unveils some relationships between different interestingness measures. Moreover, it allows to identify a set of rules containing most interesting (optimal) rules according to any interestingness measure with the property (M) simply by solving an optimized rule mining problem with respect to rule support and anti-support.

3. Analysis of Property (M)

For the clarity of presentation, the following notation shall be used throughout the next sections: $a = sup(\phi \to \psi)$, $b = sup(\neg\phi \to \psi)$, $c = sup(\phi \to \neg\psi)$, $d = sup(\neg\phi \to \neg\psi)$, $a + c = sup(\phi)$, $a + b = sup(\psi)$, $b + d = sup(\neg\phi)$, $c + d = sup(\neg\psi)$, $a + b + c + d = |U|$. We also assume that set U is not empty, so that at least one of a, b, c or d is strictly positive. In order to prove that a measure has the property (M) we need to show that it is non-decreasing with respect to a and d, and non-increasing with respect to b and c.

THEOREM 1. *Measure RI has the property (M).*

Proof. Let us observe that measure RI can be rewritten as:

$$RI(\phi \to \psi) = a - \frac{(a+b)(a+c)}{a+b+c+d}. \tag{8}$$

After some simple algebraic transformation, we obtain

$$RI(\phi \to \psi) = \frac{ad - bc}{a + b + c + d}. \tag{9}$$

Taking into account equation (9), to prove the monotonicity of RI with respect to a we have to show that if a increases by $\Delta > 0$, then RI does not decrease, i.e.

$$\frac{(a + \Delta)d - bc}{a + b + c + d + \Delta} - \frac{ad - bc}{a + b + c + d} \geq 0. \tag{10}$$

After few simple algebraic passages, and remembering that a, b, c and d are non-negative, we get

$$\begin{aligned} &\tfrac{(a+\Delta)d-bc}{a+b+c+d+\Delta} - \tfrac{ad-bc}{a+b+c+d} = \\ &= \tfrac{b(b+c+d)\Delta+bc\Delta}{(a+b+c+d)(a+b+c+d+\Delta)} > 0 \geq 0 \end{aligned} \tag{11}$$

such that we can conclude that RI is non-decreasing (more precisely, strictly increasing) with respect to a. Analogous proofs hold for the monotonicity of RI with respect to b, c and d. ■

THEOREM 2. *Measure gain has the property (M).*

Proof. Let us consider measure gain expressed as follows:

$$gain(\phi \to \psi) = a - \Theta(a + c) \tag{12}$$

where Θ is a fractional constant between 0 and 1. As $gain(\phi \to \psi)$ does not depend on b nor d, it is clear that the change of b or d does not result in any change of $gain(\phi \to \psi)$. Thus, we only need to verify if :

- (i) the increase of a results in non-decrease of $gain(\phi \to \psi)$,
- (ii) the increase of c results in non-increase of $gain(\phi \to \psi)$.

Ad.(i). Let us assume that $\Delta > 0$ is the number by which a increases. Condition (i) will be satisfied if and only if

$$gain(\phi \to \psi) = a - \Theta(a + c) \leq gain'(\phi \to \psi) = (a + \Delta) - \Theta(a + \Delta + c) \tag{13}$$

Let us observe that

$$\begin{aligned} &a - \Theta(a + c) \leq (a + \Delta) - \Theta(a + \Delta + c) \Leftrightarrow \\ &\Leftrightarrow a - a\Theta - c\Theta \leq a + \Delta - a\Theta - c\Theta - \Theta\Delta \Leftrightarrow \\ &\Leftrightarrow \Delta - \Theta\Delta \geq 0 \Leftrightarrow \Delta(1 - \Theta) \geq 0 \end{aligned} \tag{14}$$

The last inequality is always satisfied as $\Delta > 0$ and $(1 - \Theta) \geq 0$ because Θ is a fractional constant between 0 and 1. Thus, condition (i) is satisfied.

Ad.(ii). Let us assume that $\Delta > 0$ is the number by which c increases. Condition (ii) will be satisfied if and only if

$$gain(\phi \rightarrow \psi) = a - \Theta(a+c) \geq gain'(\phi \rightarrow \psi) = a - \Theta(a+\Delta+c) \tag{15}$$

Let us observe that

$$\begin{gathered} a - \Theta(a+c) \geq a - \Theta(a+\Delta+c) \Leftrightarrow \\ \Leftrightarrow a - a\Theta - c\Theta \geq a - a\Theta - c\Theta - \Theta\Delta \Leftrightarrow \\ \Leftrightarrow 0 \geq -\Theta\Delta \Leftrightarrow \Delta\Theta \geq 0 \end{gathered} \tag{16}$$

The last inequality is always satisfied as $\Delta > 0$ and $\Theta \geq 0$. Thus, condition (ii) is satisfied. Since all four conditions are satisfied, the hypothesis that gain measure has the property (M) is true. ■

Having determined that both of the analyzed measures do satisfy the desired property (M), we can draw conclusion that rules optimal according to them will be found on the support-anti-support Pareto-optimal border.

Now, let us prove by counterexample that the dependency factor $\eta(\phi \rightarrow \psi)$ does not have the property (M).

THEOREM 3. *Dependency factor $\eta(\phi \rightarrow \psi)$ does not have the property (M).*

Proof. Let us consider the dependency factor rewritten as follows:

$$\eta(\phi \rightarrow \psi) = \frac{\frac{a}{a+c} - \frac{a+b}{a+b+c+d}}{\frac{a}{a+c} + \frac{a+b}{a+b+c+d}} \tag{17}$$

It will be shown by the following counterexample that $\eta(\phi \rightarrow \psi)$ does not satisfy the condition that the increase of a results in non-decrease of $\eta(\phi \rightarrow \psi)$, thus this measure does not have the property (M). Let us consider case α, in which a=7, b=2, c=3, d=3, and case α', in which a increases to 8 and b, c, d remain unchanged. The dependency factor does not have the property (M) as such increase of a results in the decrease of the measure:

$$\eta(\phi \rightarrow \psi) = 0.0769 > 0.0756 = \eta'(\phi \rightarrow \psi). \tag{18}$$

■

4. Experimental illustration of the result

It was proved by Brzezińska *et al.* (2007) that rules optimal with respect to any interestingness measure that has the property (M) will reside on the support-anti-support Pareto-optimal border. Since, the above analysis shows that both RI and gain enjoy the property (M), we can concude that rules optimal with respect to them will be found in the set of rules non-dominated according to support and anti-support. Several computational experiments analyzing rules optimal with respect to RI and gain in the perspective of rule support and anti-support have been conducted in order to illustrate the theoretical results concerning their possession of the property (M) and thus, their occurrence on the support-anti-support Pareto-optimal border.

Below, in Fig. 1, there is an exemplary diagram from those experiments. For a real life dataset containing information about technical state of buses, a set of all possible rules was generated. A set of 85 rules with the same conclusion was then isolated and rules non-dominated with respect to support and anti-support were found. The support-anti-support Pareto-optimal border is indicated in Fig. 1 by circles connected by a line. Four points marked as r_1, r_2, r_3, r_4 form the Pareto-optimal border. Each of those points respresents rules characterized by particular values of support and anti-support (i.e., r_1 represents rules with $sup(\phi \to \psi) = 50$ and $anti - sup(\phi \to \psi) = 4$, r_2 rules with $sup(\phi \to \psi) = 49$ and $anti - sup(\phi \to \psi) = 2$, r_3 rules with $sup(\phi \to \psi) = 48$ and $anti - sup(\phi \to \psi) = 1$, and r_4 rules with $sup(\phi \to \psi) = 45$ and $anti - sup(\phi \to \psi) = 0$). In the generated set of 85 rules, we have distinguished rules optimal according to RI (marked by r_3), and gain for different values of Θ. For $\Theta = 0.33$ the rules with maximal gain are marked as r_1; when $\Theta = 0.5$ these are rules marked as r_2 or r_3; finally when $\Theta = 0.66$ these are rules marked as r_3. The diagram shows that, indeed, rules optimal with respect to those measures lie on the support-anti-support Pareto-optimal border.

During this experiment we have also calculated the optimal value of the dependency factor. This measure does not have the property (M) so we could not conclude right away that rules optimal according to it will be on the support-anti-support Pareto-optimal border. However, since possession of the property (M) is only a sufficient condition for laying on that border, we cannot exclude a situation in which rules optimal with respect to the dependency factor will be found on the support-anti-support Pareto-optimal border. For this dataset we have such a case. Rules marked as r_4 are optimal according to dependency factor and they also form the support-anti-support Pareto-optimal border.

5. Analysis of Hypothesis Symmetry (HS)

The verification of the property of hypothesis symmetry was done for all three considered measures separately, by checking if their values for rules $\phi \to \psi$ and $\phi \to \neg\psi$ are the same but of opposite sign.

THEOREM 4. *Measure RI has the property of hypothesis symmetry.*

Proof. Let us consider RI expressed as follows:

$$RI(\phi \to \psi) = a - \frac{(a+c)(a+b)}{a+b+c+d}. \tag{19}$$

For a negated conclusion RI is defined as:

$$RI(\phi \to \neg\psi) = c - \frac{(a+c)(c+d)}{a+b+c+d}. \tag{20}$$

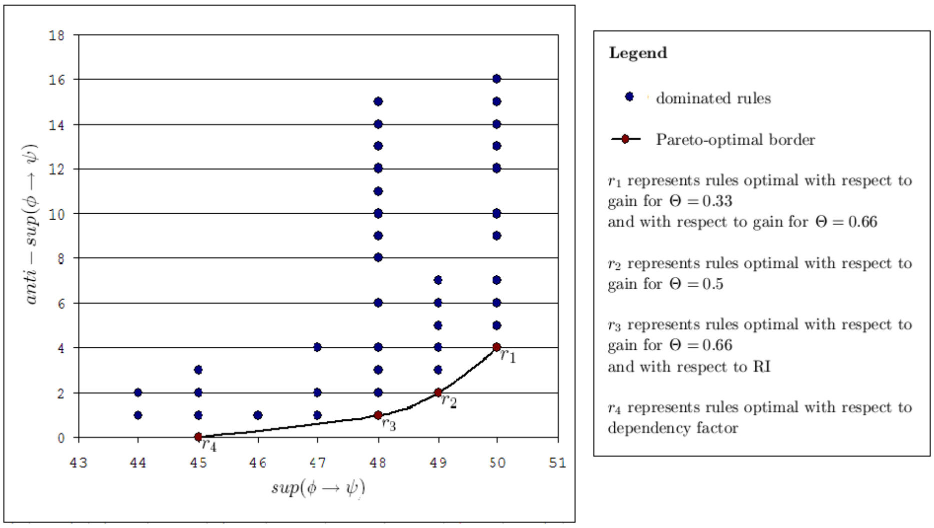

Figure 1. Pareto-optimal border with respect to rule support and anti-support includes rules being optimal in RI, and gain

The hypothesis symmetry will be satisfied by RI if:

$$a - \frac{(a+c)(a+b)}{a+b+c+d} = -[c - \frac{(a+c)(c+d)}{a+b+c+d}]. \tag{21}$$

Through simple mathematical transformation we obtain that:

$$a - \frac{(a+c)(a+b)}{a+b+c+d} = \frac{ad-bc}{a+b+c+d} \tag{22}$$

and

$$-c + \frac{(a+c)(c+d)}{a+b+c+d} = \frac{ad-bc}{a+b+c+d} \tag{23}$$

and thus, we can conclude that RI has the property of hypothesis symmetry.

■

THEOREM 5. *Measure gain has the property of hypothesis symmetry iff* $\Theta = 1/2$.

Proof. Let us consider gain expressed as follows:

$$gain(\phi \to \psi) = a - \Theta(a+c). \tag{24}$$

For a negated conclusion gain is defined as:

$$gain(\phi \to \neg\psi) = c - \Theta(a + c). \tag{25}$$

The hypothesis symmetry will be satisfied by gain if:

$$a - \Theta(a + c) = -[c - \Theta(a + c)]. \tag{26}$$

Through simple mathematical transformation we obtain that the above equality is satisfied only when $\Theta = 1/2$.

■

THEOREM 6. *The dependency factor η does not have the property of hypothesis symmetry.*

Proof. Let us consider dependency factor expressed as follows:

$$\eta(\phi \to \psi) = \frac{\frac{a}{a+c} - \frac{a+b}{a+b+c+d}}{\frac{a}{a+c} + \frac{a+b}{a+b+c+d}} \tag{27}$$

For a negated conclusion it is defined as:

$$\eta(\phi \to \neg\psi) = \frac{\frac{c}{a+c} - \frac{c+d}{a+b+c+d}}{\frac{c}{a+c} + \frac{c+d}{a+b+c+d}} \tag{28}$$

To prove that the dependency factor does not satisfy the hypothesis symmetry let us set $a = b = c = 10$ and $d = 20$. We can easily verify that

$$\eta(\phi \to \psi) = 0.11 \neq 0.09 = \eta(\phi \to \neg\psi). \tag{29}$$

■

6. Conclusions

As an active research area in data mining, rule evaluation has been considered by many authors from different perspectives. This paper concentrated on measuring the relevance and utility of induced rules according to three popular interestingness measures: rule interest function of Piatetsky-Shapiro, gain measure of Fukuda et al., and dependency factor of Pawlak. A theoretical analysis has been conducted for verifying which of those measures satisfy valuable properties (M) and hypothesis symmetry (HS). It has been proved that the rule interest function and gain measure are characterized by both of those properties, while the dependency factor does not satisfy any of them. Such analysis of properties of interestingness measures was carried out in order to widen our knowledge and understanding of them, as well as their applicability. Moreover, the possession of the property (M) unveils an interesting relationship between rule interest function and gain on one hand, and two other interestingness measures: rule support and anti-support, on the other hand. It has been shown that rules maximizing rule interest function or gain will surely be found on the rule support-anti-support Pareto-optimal border. It has also been illustrated on a real life dataset.

The obtained results are useful for practical applications because they show which interestingness measures are relevant for meaningful rule evaluation. Using the measures which enjoy the desirable properties one can avoid analyzing unimportant rules.

References

BAYARDO, R.J. and AGRAWAL, R. (1999) Mining the Most Interesting Rules. *Proc. of Fifth ACM-SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining.* **1**45–154.

BRZEZIŃSKA, I. and GRECO, S. and SŁOWIŃSKI, R. (2007) Mining Pareto-optimal rules with respect to support and anti-support. *Engineering Applications of Artificial Intelligence* Vol. 20(5) **5**87–600.

CARNAP, R. (1962) *Logical Foundations of Probability, 2nd ed.* University of Chicago Press, Chicago.

EELLS, E and FITELSON, B. (2002) Symmetries and assymmetries in evidential support. *Philosophical Studies* 107 (2) **1**29–142.

FITELSON, B. (2001) *Studies in Bayesian confirmation theory.* Ph.D. Thesis, University of Wisconsin, Madison.

FUKUDA, T. and MORIMOTO, Y. and MORISHITA, S. and TOKUYAMA, T. (1996) Data Mining using Two-Dimensional Optimized Association Rules: Schemes, Algorithms, and Visualization. *Proc. of the 1996 ACM SIGMOD Int'l Conference on Management of Data.* **1**3–23.

GRECO, S. and PAWLAK, Z. and SŁOWIŃSKI, R. (2004) Can Bayesian confirmation measures be useful for rough set decision rules? *Engineering Applications of Artificial Intelligence.***17**, 345–361.

HILDERMAN, R. and HAMILTON, H. (2001) Knowledge Discovery and Measures of Interest. *K*luwer Academic Publishers, Boston.

PAWLAK, Z. (2004) Some issues on Rough Sets. *Transactions on Rough Sets I* LNCS 3100 **,** 1–58.

PIATETSKY-SHAPIRO, G. (1991) Discovery, analysis and presentation of strong rules. *Knowledge Discovery in Databases.* AAAI/MIT Press. **2**29–248.