Bayesian Confirmation Measures within Rough Set Approach

Salvatore Greco[1], Zdzisław Pawlak[2,3], and Roman Słowiński[4,5]

[1] Faculty of Economics, University of Catania, Corso Italia, 55, 95129 Catania, Italy
salgreco@mbox.unict.it
[2] Institute of Theoretical and Applied Informatics, Polish Academy of Sciences
44-100 Gliwice, Poland
[3] Warsaw School of Information Technology, 01-447 Warsaw, Poland
zpw@ii.pw.edu.pl
[4] Institute of Computing Science, Poznań University of Technology, 60-965 Poznań, Poland
[5] Institute for Systems Research, Polish Academy of Sciences, 01-447 Warsaw, Poland
slowinsk@sol.put.poznan.pl

Abstract. Bayesian confirmation theory considers a variety of non-equivalent confirmation measures quantifying the degree to which a piece of evidence supports a hypothesis. In this paper, we apply some of the most relevant confirmation measures within the rough set approach. Moreover, we discuss interesting properties of these confirmation measures and we propose a new property of monotonicity that is particularly relevant within rough set approach. The main result of this paper states which one of the confirmation measures considered in the literature have the desirable properties from the viewpoint of the rough set approach.

1 Introduction

Reasoning from data is the domain of inductive reasoning. Contrary to deductive reasoning, where axioms expressing some universal truths constitute a starting point of reasoning, inductive reasoning uses data about a sample of a larger reality to start inference.

Rough set theory (Pawlak 1982, 1991) is a mathematical approach to data analysis. Rough-set-based data analysis starts from a data table, called *information table*. The information table contains data about objects of interest, characterized by a finite set of attributes. Among the attributes, *condition attributes C* and *decision attributes D* are distinguished, in order to analyze how values of attributes *C* associate with values of attributes *D*. An information table where condition attributes and decision attributes are distinguished is called *decision table*. From a decision table one can induce some relationships (patterns) in form of "*if … then …*" decision rules. More exactly, the decision rules say that if some condition attributes have given values, then some decision attributes have other given values. With every decision rule induced from a decision table, three coefficients are traditionally associated: the strength, the certainty and the coverage factors of the rule. They are useful to show that discovering patterns in data can be represented in terms of Bayes' theorem (Pawlak 2002; Greco,

S. Tsumoto et al. (Eds.): RSCTC 2004, LNAI 3066, pp. 264–273, 2004.

Pawlak, Słowiński 2002) in a different way from that offered by standard Bayesian inference techniques, without referring to prior and posterior probabilities, inherently associated with Bayesian inference methodology.

Within inductive reasoning, classical Bayesian theory considers a variety of non-equivalent confirmation measures (see (Fitelson 2001) for a survey) which quantify the degree to which a piece of evidence E provides, "evidence for or against" or "support for or against" a hypothesis H. In this paper, we take into account some of the most relevant of these confirmation measures and apply them within rough set approach to data analysis. Moreover, we discuss some interesting properties of these confirmation measures, which are particularly relevant within rough set approach.

Our research is strongly related to the rich discussion about the interestingness measure for decision rules in data mining (see, for example, (Hilderman and Hamilton 2002) and (Yao and Zhong 1999) for exhaustive reviews of the subject). Moreover, some confirmation measures considered in this paper may remember statistical independence tests of a contingency table. Indeed, some interestingness measures of decision rules, which are based on these statistical tests, have been proposed in the specialized literature (see, for example, (Flach and Lachiche 2001), (Tsumoto 2002), (Zembowicz and Zytkow 1996)). It is worth stressing that the confirmation measures take a different perspective than the statistical approach. First, observe that the independence (dependence) measures are symmetric while decision rules, for which these measures are conceived, are not symmetric. Even if some authors tried to generalize classical statistical analysis of a contingency table in order to handle typical asymmetries of rule induction (Flach and Lachiche 2001), our approach is different in nature because we are interested in some desirable properties of confirmation measures rather than in their statistical properties.

We think that our research, besides operational impact, can also be interesting for philosophical research about confirmation. In fact, quantitative confirmation theory is strongly based on probability functions, however, there is a great and well-known controversy relative to interpretation, origin and status of probability. Conclusively, in this paper, we use the theory of quantitative confirmation theory which, instead, is based on observed data, without any consideration of probability functions.

Let us also remark that the concept of confirmation we are interested in is related to the concept of independence of logical formulas (propositions), as presented by Łukasiewicz (1913). In brief, his definition of independence between two propositions Φ and Ψ amounts to say that the credibility of Ψ given Φ is the same as the credibility of Ψ given $\neg\Phi$. Thus, independence means that the credibility of Φ does not influence the credibility of Ψ. For this definition Łukasiewicz proved the law of multiplication which says that if propositions Φ and Ψ are independent, then the credibility of Ψ given Φ is equal to the product of the individual credibilities of Φ and Ψ. From this law, Pawlak (2003) derived a dependency factor for flows in decision networks and then he applied this formula to decision rules (Pawlak 2004). The dependency factors derived from the concept of Łukasiewicz and the measures of confirmation studied in this paper are based, however, on different desiderata.

The article is organized as follows. Section 2 introduces confirmation measures and recalls some desirable properties of symmetry and asymmetry proposed by Eells and Fitelson. Section 3 gives some basic notions concerning decision rules and decision algorithms within rough set approach. Section 4 introduces rough set confirmation measures. In section 5, we introduce a specific monotonicity property of rough set confirmation measures. Section 6 investigates which one among the considered rough set confirmation measures satisfies the monotonicity property. Final section draws conclusions and some directions of future research. Proofs of theorems and many interesting extensions can be found in (Greco, Pawlak, Słowiński 2004).

2 Confirmation Measures

According to Fitelson (2001), measures of confirmation quantify the degree to which a piece of evidence E provides, "evidence for or against" or "support for or against" a hypothesis H. Fitelson remarks, moreover, that measures of confirmation are supposed to capture the impact rather than the final result of the "absorption" of a piece of evidence.

Bayesian confirmation assume the existence of a probability Pr. In the following, given a proposition X, $Pr(X)$ is the probability of X. Given X and Y, $Pr(X|Y)$ represents the probability of X given Y, i.e.

$$Pr(X|Y) = Pr(X \wedge Y)/Pr(Y).$$

In this context, a measure of confirmation of a piece of evidence E with respect to a hypothesis H is denoted by $c(E,H)$. $c(E,H)$ is required to satisfy the following minimal property:

$$c(E,H) = \begin{cases} >0 & \text{if } Pr(H|E) > Pr(H) \\ =0 & \text{if } Pr(H|E) = Pr(H) \\ <0 & \text{if } Pr(H|E) < Pr(H) \end{cases}$$

The most well known confirmation measures proposed in the literature are the following:

$$d(E,H) = Pr(H|E) - Pr(H), \quad r(E,H) = log\left[Pr(H|E)/Pr(H)\right]$$

$$l(E,H) = log\left[Pr(E|H)/Pr(E|\neg H)\right], \quad f(E,H) = \frac{Pr(E|H) - Pr(E|\neg H)}{Pr(E|H) + Pr(E|\neg H)}$$

$$s(E,H) = Pr(H|E) - Pr(H|\neg E), \quad b(E,H) = Pr(H \wedge E) - Pr(H)\,Pr(E)$$

Measure $d(E,H)$ has been supported by Earman (1992), Eells (1982), Gillies (1986), Jeffrey (1992) and Rosenkrantz (1994). Measure $r(E,H)$ has been defended by Horwich (1982), Keynes (1921), Mackie (1969), Milne (1995, 1996), Schlesinger (1995) and Pollard (1999). Measure $l(E,H)$ and $f(E,H)$ have been supported by Kemeny and Oppenheim (1952), Good (1984), Heckerman (1988), Pearl (1988), Schumm (1994). Fitelson (2001) has advocated for measure $f(E,H)$. Measure $s(E,H)$ has been proposed by Christensen (1999) and Joyce (1999). Measure $b(E,H)$ has been introduced by Carnap (1962).

Many authors have considered, moreover, some more or less desirable properties of confirmation measures. Fitelson (2001) makes a comprehensive survey of these considerations. At the end of his retrospective, Fitelson concludes that the most convincing confirmation measures are $l(E,H)$ and $f(E,H)$. He also proves that $l(E,H)$ and $f(E,H)$ are ordinally equivalent, i.e. for all E,H and E', H',

$$l(E,H) \geq l(E',H') \text{ if and only if } f(E,H) \geq f(E',H').$$

Among the properties of confirmation measures reviewed by Fitelson (2001), there are properties of symmetry introduced by Carnap (1962) and investigated recently by Eells and Fitelson (2000). For all E and H, one can have:

- Evidence Symmetry (ES): $c(E,H) = -c(\neg E,H)$
- Commutativity Symmetry (CS): $c(E,H) = c(H,E)$
- Hypothesis Symmetry (HS): $c(E,H) = -c(E,\neg H)$
- Total Symmetry (TS): $c(E,H) = c(\neg E,\neg H)$

Eells and Fitelson (2000) remarked that, given (CS), (ES) and (HS) are equivalent, and that (TS) follows from the conjunction of (ES) and (HS). Moreover, they advocate in favor of (HS) and against (ES), (CS) and (TS). The reason in favor of (HS) is that the significance of E with respect to H should be of the same strength, but of opposite sign, as the significance of E with respect to $\neg H$.

Eells and Fitelson (2000) prove that

1) s and b satisfy (ES), while d,r,l and f do not satisfy (ES),
2) d,s,b,f and l satisfy (HS), while r does not satisfy (HS),
3) r and b satisfy (CS), while d,s,f and l do not satisfy (CS),
4) s and b satisfy (TS), while d,r,f and l do not satisfy (CS).

Thus, assuming that (HS) is a desirable property, while (ES), (CS) and (TS) are not, Eells and Fitelson (2000) conclude that with respect to the property of symmetry, d, f and l are satisfying confirmation measures while s, r and b are not satisfying confirmation measures.

3 Decision Rules and Decision Algorithm

Let $S = (U, A)$ be an information table, where U and A are finite, non-empty sets called the *universe* and the set of *attributes*, respectively. If in set A two disjoint subsets of *condition* and *decision attributes* are distinguished (C and D, respectively), then the system is called *decision table* and is denoted by $S = (U, C, D)$. With every subset of attributes, one can associate a formal language of logical formulas $\boldsymbol{L}$ defined in a standard way and called the *decision language*. Formulas for a subset $B \subseteq A$ are build up from attribute-value pairs (a, v), where $a \in B$ and $v \in V_a$ (set V_a is a domain of a), by means of logical connectives $\wedge$ (*and*), $\vee$ (*or*), $\neg$ (*not*). We assume that the set of all formula sets in $\boldsymbol{L}$ is partitioned into two classes, called *condition* and *decision formulas*, respectively.

A *decision rule* induced from S and expressed in $\boldsymbol{L}$ is presented as $\Phi \rightarrow \Psi$, read "*if* Φ*, then* Ψ", where Φ and Ψ are condition and decision formulas in $\boldsymbol{L}$, called *premise* and *conclusion*, respectively. A decision rule $\Phi \rightarrow \Psi$ is also seen as a binary relation between premise and conclusion, called *consequence relation* (see critical discussion about interpretation of decision rules as logical implications in (Greco, Pawlak, Słowiński 2004)).

Let $\|\Phi\|$ denote the set of all objects from universe U, having the property Φ in S.

If $\Phi \rightarrow \Psi$ is a decision rule, then $supp_S(\Phi,\Psi) = card(\|\Phi \wedge \Psi\|)$ will be called the *support* of the decision rule and $\sigma_S(\Phi,\Psi) = supp_S(\Phi,\Psi)/card(U)$ will be referred to as the *strength* of the decision rule.

With every decision rule $\Phi\to\Psi$ we associate a *certainty factor* $cer_S(\Phi,\Psi) = supp_S(\Phi,\Psi)/card(\|\Phi\|)$ and a *coverage factor* $cov_S(\Phi,\Psi) = supp_S(\Phi,\Psi)/card(\|\Psi\|)$.

If $cer_S(\Phi,\Psi)=1$, then the decision rule $\Phi\to\Psi$ is called *certain*, otherwise the decision rule is referred to as *uncertain*. A set of decision rules supported in total by the universe U creates a *decision algorithm* in S.

4 Confirmation Measures and Decision Algorithms

Given a decision rule $\Phi\to\Psi$, the confirmation measure we want to introduce should give the credibility of the proposition: **Ψ is satisfied more frequently when Φ is satisfied rather than when Φ is not satisfied**.

Differently from Bayesian confirmation, however, we start from a decision table rather than from a probability measure. In this context, the probability Pr of Φ is substituted by the relative frequency Fr in the considered data table S, i.e.

$$Fr_S(\Phi) = card(\|\Phi\|)/card(U).$$

Analogously, given Φ and Ψ, $Pr(\Psi|\Phi)$ – the probability of Ψ given Φ – is substituted by the certainty factor $cer_S(\Phi,\Psi)$ of the decision rule $\Phi\to\Psi$.

Therefore, a measure of confirmation of property Ψ by property Φ, denoted by $c(\Phi,\Psi)$, where Φ is a condition formula in $\boldsymbol{L}$ and Ψ is a decision formula in $\boldsymbol{L}$, is required to satisfy the following minimal property

$$c(\Phi,\Psi) = \begin{cases} >0 & if\ \ cer_S(\Phi,\Psi) > Fr_S(\Psi) \\ =0 & if\ \ cer_S(\Phi,\Psi) = Fr_S(\Psi) \\ <0 & if\ \ cer_S(\Phi,\Psi) < Fr_S(\Psi) \end{cases} \tag{i}$$

(i) can be interpreted as follow:

- $c(\Phi,\Psi)>0$ means that property Ψ is satisfied more frequently when Φ is satisfied (then, this frequency is $cer_S(\Phi,\Psi)$), rather than generically in the whole decision table (where this frequency is $Fr_S(\Psi)$),
- $c(\Phi,\Psi)=0$ means that property Ψ is satisfied with the same frequency when Φ is satisfied and generically in the whole decision table,
- $c(\Phi,\Psi)<0$ means that property Ψ is satisfied less frequently when Φ is satisfied, rather than generically in the whole decision table.

The specific confirmation measures recalled in section 2 can be rewritten in this context as follows:

$$d(\Phi,\Psi) = cer_S(\Phi,\Psi) - Fr_S(\Psi), \quad r(\Phi,\Psi) = log\left[\frac{cer_S(\Phi,\Psi)}{Fr_S(\Psi)}\right]$$

$$l(\Phi,\Psi) = log\left[\frac{cer_S(\Psi,\Phi)}{cer_S(\neg\Psi,\Phi)}\right], \quad f(\Phi,\Psi) = \frac{cer_S(\Psi,\Phi) - cer_S(\neg\Psi,\Phi)}{cer_S(\Psi,\Phi) + cer_S(\neg\Psi,\Phi)}$$

$$s(\Phi,\Psi) = cer_S(\Phi,\Psi) - cer_S(\neg\Phi,\Psi), \quad b(\Phi,\Psi) = cer_S(\Phi,\Psi) - Fr_S(\Phi)\,Fr_S(\Psi)$$

Clearly, all the results about confirmation measures obtained within Bayesian confirmation theory are valid for the confirmation measures defined in the context of decision algorithms considered within rough set theory. Therefore, according to Fitelson's conclusions reminded in section 2, we believe that $l(\Phi,\Psi)$ and $f(\Phi,\Psi)$ are the most convincing confirmation measures, which continue to be ordinally equivalent in this new context. Below, we call the confirmation measures presented in the above language of decision algorithms, the *rough set confirmation measures*.

5 Desirable Properties for Rough Set Confirmation Measures

Even if all the formal properties of the Bayesian confirmation measures hold also for the corresponding rough set confirmation measures, we think that there is a new property which would be desirable for the latter measures.

To introduce this new property, let us remark that for each formula Φ, Ψ in $\boldsymbol{L}$, one can express the rough set confirmation measures in terms of the following four values:

- $a=supp_S(\Phi,\Psi)$ - number of objects in U for which Φ and Ψ hold together,
- $b=supp_S(\neg\Phi,\Psi)$ - number of objects in U for which Φ doesn't hold while Ψ holds,
- $c=supp_S(\Phi,\neg\Psi)$ - number of objects in U for which Φ holds while Ψ doesn't hold,
- $d=supp_S(\neg\Phi,\neg\Psi)$ - number of objects in U for which both Φ and Ψ don't hold.

Therefore, the rough set confirmation measures can be expressed as follows:

$$d(\Phi,\Psi) = \frac{ad-bc}{(a+c)(a+b+c+d)},\quad r(\Phi,\Psi) = log\left[\left(\frac{a}{a+c}\right)\Big/\left(\frac{a+b}{a+b+c+d}\right)\right]$$

$$l(\Phi,\Psi) = log\left[\left(\frac{a}{a+b}\right)\Big/\left(\frac{c}{c+d}\right)\right],\quad f(\Phi,\Psi) = \frac{ad-bc}{ad+bc+2ac}$$

$$s(\Phi,\Psi)=\frac{ad-bc}{(a+c)(b+d)},\quad b(\Phi,\Psi)=\frac{ad-bc}{(a+b+c+d)^2}$$

In this context, we propose the following property of monotonicity:

(M) $c(\Phi,\Psi) = F[supp_S(\Phi,\Psi), supp_S(\neg\Phi,\Psi), supp_S(\Phi,\neg\Psi), supp_S(\neg\Phi,\neg\Psi)]$ is a function non-decreasing with respect to $supp_S(\Phi,\Psi)$ and $supp_S(\neg\Phi,\neg\Psi)$ and non-increasing with respect to $supp_S(\neg\Phi,\Psi)$ and $supp_S(\Phi,\neg\Psi)$.

The monotonicity property (M) has the following interpretation. Monotonicity of $c(\Phi,\Psi)$ with respect to $supp_S(\Phi,\Psi)$ means that any evidence in which Φ and Ψ hold together increases (or at least does not decrease) the credibility of the decision rule $\Phi\to\Psi$. Monotonicity of $c(\Phi,\Psi)$ with respect to $supp_S(\Phi,\neg\Psi)$ means that any evidence in which Φ holds and Ψ does not hold decreases (or at least does not increase) the credibility of the decision rule $\Phi\to\Psi$. Analogously, with respect to $supp_S(\neg\Phi,\Psi)$, any evidence in which Φ does not hold and Ψ holds decreases (or at least does not increase) the credibility of the decision rule $\Phi\to\Psi$, and with respect to $supp_S(\neg\Phi,\neg\Psi)$, any evidence in which both Φ and Ψ do not hold increases (or at least does not decrease) the credibility of the decision rule $\Phi\to\Psi$.

Let us remark that the monotonicity with respect to $supp_S(\Phi,\Psi)$ and $supp_S(\Phi,\neg\Psi)$ is rather intuitive. In the context of automated analysis of data, Hajek and Havranek

(1978) suggest the same monotonicity property M, however, not for a confirmation measure of rule $\Phi \rightarrow \Psi$ (for which they suggest the increasing monotonicity with respect to $supp_S(\Phi,\Psi)$ and the decreasing monotonicity with respect to $supp_S(\Phi,\neg\Psi)$ only), but for a more specific association measure.

We can explain the monotonicity with respect to $supp_S(\neg\Phi,\Psi)$ and $supp_S(\neg\Phi,\neg\Psi)$, considering our interpretation of property (i) from section 4: a positive value of a confirmation measure $c(\Phi,\Psi)$ means that property Ψ is satisfied more frequently when property Φ is satisfied rather than when Φ is not satisfied. From this viewpoint, an evidence in which Φ is not satisfied and Ψ is satisfied (objects $\|\neg\Phi\wedge\Psi\|$) increases the frequency of Ψ in the situations where Φ is not satisfied and thus it should decrease the value of the confirmation. Analogously, an evidence in which both Φ and Ψ are not satisfied (objects $\|\neg\Phi\wedge\neg\Psi\|$) decreases the frequency of Ψ in the situations where Φ is not satisfied and thus it should increase the value of the confirmation.

We want to give also a more formal justification to the monotonicty of confirmation measures with respect to $supp_S(\neg\Phi,\Psi)$ and $supp_S(\neg\Phi,\neg\Psi)$. Let us consider the following definition of confirmation: property Φ confirms property Ψ if

$$cer_S(\Phi,\Psi) > Fr_S(\Psi). \tag{iii}$$

Let us remark that definition (iii) corresponds to the definition of *incremental confirmation* introduced by Carnap (1962, new preface) under the name of "*confirmation as increase in firmness*" in the following form: evidence Φ confirms hypothesis Ψ if $Pr(\Psi|\Phi) > Pr(\Psi)$.

The confirmation measures $d(\Phi,\Psi)$, $r(\Phi,\Psi)$, $l(\Phi,\Psi)$, $f(\Phi,\Psi)$, $s(\Phi,\Psi)$ and $b(\Phi,\Psi)$ can be seen as quantitative generalizations of the qualitative incremental confirmation (Fitelson 2001).

Redefining (iii) in terms of $a=supp_S(\Phi,\Psi)$, $b=supp_S(\neg\Phi,\Psi)$, $c=supp_S(\Phi,\neg\Psi)$ and $d=supp_S(\neg\Phi,\neg\Psi)$, we get:

$$a/(a+c) > (a+b)/(a+b+c+d). \tag{iii′}$$

The following theorem is useful for justifying the property of monotonicity.

Theorem 1. Let us consider case α in which

$$a=supp_S(\Phi,\Psi),\quad b=supp_S(\neg\Phi,\Psi),\quad c=supp_S(\Phi,\neg\Psi),\quad d=supp_S(\neg\Phi,\neg\Psi),$$

and case α' in which

$$a'=supp_S(\Phi',\Psi'),\quad b'=supp_S(\neg\Phi',\Psi'),\quad c'=supp_S(\Phi',\neg\Psi'),\quad d'=supp_S(\neg\Phi',\neg\Psi').$$

Let us suppose, moreover, that

$$cer_S(\Phi,\Psi) < Fr_S(\Psi),\quad \text{while}\quad cer_S(\Phi',\Psi') > Fr_S(\Psi').$$

The following implications are satisfied:

1) if $a'=a+\Delta$, $b'=b$, $c'=c$ and $d'=d$, then $\Delta>0$,
2) if $a'=a$, $b'=b+\Delta$, $c'=c$ and $d'=d$, then $\Delta<0$,
3) if $a'=a$, $b'=b$, $c'=c+\Delta$ and $d'=d$, then $\Delta<0$,
4) if $a'=a$, $b'=b$, $c'=c$ and $d'=d+\Delta$, then $\Delta>0$.

Theorem 1 has the following interpretation. Passing from case α to case α', we pass from a situation in which property Φ does not confirm property Ψ, to a situation in which property Φ' confirms property Ψ'. Theorem 1 says that this passage from non-confirmation to confirmation is permitted by an increase of $supp_S(\Phi,\Psi)$ or $supp_S(\neg\Phi,\neg\Psi)$, or by a decrease of $supp_S(\neg\Phi,\Psi)$ or $supp_S(\Phi,\neg\Psi)$. Thus, the theorem supports the claim that confirmation given by property Φ to property Ψ is positively related to $supp_S(\Phi,\Psi)$ and $supp_S(\neg\Phi,\neg\Psi)$, and negatively related to $supp_S(\neg\Phi,\Psi)$ and $supp_S(\Phi,\neg\Psi)$.

In fact, Theorem 1 supports the monotonicity property (M) because, if the passage from a situation of non-confirmation to a situation of confirmation implies a specific sign of modifications of the four values $supp_S(\Phi,\Psi)$, $supp_S(\neg\Phi,\neg\Psi)$, $supp_S(\neg\Phi,\Psi)$ and $supp_S(\Phi,\neg\Psi)$, it is natural to expect that confirmation measures will react analogously to modifications of the above values.

6 Rough Set Confirmation Measures Satisfying Monotonicity

Theorem 2. $l(\Phi,\Psi)$, $f(\Phi,\Psi)$ and $s(\Phi,\Psi)$ satisfy monotonicity property (M), while $d(\Phi,\Psi)$, $r(\Phi,\Psi)$ and $b(\Phi,\Psi)$ do not satisfy (M).

The content of Theorem 2 is quite clear and immediate, however, a more detailed comment may be useful. From our viewpoint, the most important discovery coming from Theorem 2 is that the confirmation measure $d(\Phi,\Psi)$ does not satisfy the monotonicity property. The importance of this result is threefold:

1) $d(\Phi,\Psi)$ is a very simple rough set confirmation measure, coherent with the definition of incremental confirmation; it is rather counterintuitive that $d(\Phi,\Psi)$ does not satisfy monotonicity, while other confirmation measures having as complex formulation as $l(\Phi,\Psi)$ and $f(\Phi,\Psi)$ do;
2) $d(\Phi,\Psi)$ does not satisfy monotonicity with respect to $supp_S(\Phi,\Psi)$; in this case the monotonicity property is indeed an uncontestable principle;
3) $d(\Phi,\Psi)$ is not ruled out by the symmetry/asymmetry test performed by Eells and Fitelson (2000); this means that the contribution of monotonicity property (M) in reducing the field of "coherent" confirmation measures is very relevant; in fact, the only confirmation measures which satisfy both symmetry/asymmetry properties of Eells and Fitelson and monotonicity property (M) are the two ordinally equivalent confirmation measures $l(\Phi,\Psi)$ and $f(\Phi,\Psi)$.

7 Conclusions

The main result of this paper states that, from among the confirmation measures considered in the literature and recalled in section 2, the only confirmation measures satisfying the desirable properties of symmetry/asymmetry of Eells and Fitelson (2000), as well as our monotonicity property (M) are the two ordinally equivalent measures $l(\Phi,\Psi)$ and $f(\Phi,\Psi)$. In particular, our property (M) rules out $d(\Phi,\Psi)$. Let us remark that using the symmetry/asymmetry properties, it is not possible to discard $d(\Phi,\Psi)$, while using our monotonicity property, it is not possible to discard $s(\Phi,\Psi)$. This can be interpreted in the sense that the symmetry/asymmetry properties together with our monotonicity property (M) can be considered as complementary basic principles on which a sound theory of confirmation measures can be founded. A special attention merits, moreover, the violation of the monotonicity property by confirmation measure $b(\Phi, \Psi)$ which is the corroboration measure proposed by Carnap (1962). From our point of view, the violation of the monotonicity property by this confirmation measure is more troubling than its violation of the symmetry/asymmetry property proposed by Eells and Fitelson. In fact Carnap (1962) liked that his corroboration measure b satisfies all four symmetry properties ES, HS, CS and TS, because he was interested in representing quantitatively a completely symmetric relevance relation.

We think that the quite theoretical results presented in this paper can be the basis for important operational development within rough set theory and, in general, within data analysis. Only to give some idea of interesting issues for future researches consider the use of measures $l(\Phi, \Psi)$ and $f(\Phi, \Psi)$ for assessing the interest of decision rules induced from a data table, as well as classification with these rules. Considering the huge number of decision rules which can be induced from a data set, and the necessity of presenting only the most interesting rules to the users, this is a problem of primary importance for data analysis. We believe that the contribution of rough set confirmation measures to solving this problem is very important.

Acknowledgements

The research of the first author has been supported by the Italian Ministry of Education, University and Scientific Research. The third author wishes to acknowledge financial support from the State Committee for Scientific Research and from the Foundation for Polish Science.

References

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

Christensen, D. (1999), Measuring Confirmation. *Journal of Philosophy* XCVI, 437–61.

Earman, J. (1992), *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*. MIT Press, Cambridge.

Eells, E. (1982), *Rational Decision and Causality*. Cambridge University Press.

Eells, E., Fitelson, B. (2000), Measuring confirmation and evidence. *Journal of Philosophy* 97, 663–672.

Fitelson, B. (2001), *Studies in Bayesian Confirmation Theory*. Ph.D. thesis, University of Wisconsin – Madison.

Flach, P.A., Lachiche, N. (2001), Confirmation-guided discovery of first-order rules with tertius. *Machine Learning* 42, 61–95.

Gillies, D. (1986), In defense of the Popper-Miller argument. *Philosophy of Science* 53, 110–113.

Good, I. (1984), The best explicatum for weight of evidence. *Journal of Statistical Computation and Simulation* 19, 294–299.

Greco, S., Pawlak, Z., Słowiński, R. (2002), Generalized decision algorithms, rough inference rules, and flow graphs. [In]: J.J.Alpigini, J.F.Peters, A.Skowron, N.Zhong (eds.), *Rough Sets and Current Trends in Computing*. LNAI 2475, Springer, Berlin, pp. 93–104.

Greco, S., Pawlak, Z., Słowiński, R. (2004), Can Bayesian confirmation measures be useful for rough set decision rules? Submitted.

Hajek, P., Havranek, T. (1978), *Mechanising Hypothesis Formulation: Mathematical Foundation for a General Theory*. Springer-Verlag, Berlin.
Heckerman, D. (1988), An axiomatic framework for belief updates. [In]: L.Kanal & J. Lemmer (eds.), *Uncertainty in Artificial Intelligence* (2). Elsevier, New York, pp.11–22.
Hilderman, R.J., Hamilton, H.J. (2002), *Knowledge Discovery and Measures of Interest* , Kluwer, Boston.
Horwich, P.(1982), *Probability and Evidence*. Cambridge University Press.
Horvitz, E. and Heckerman, D. (1986), The inconsistent use of certainty measures in artificial intelligence research. In L. Kanal and J. Lemmer (eds.), *Uncertainty in Artificial Intelligence* (1). Elsevier, New York, pp. 137–151.
Jeffrey, R. (1992), *Probability and the Art of Judgment.* Cambridge University Press.
Joyce, J. (1999), *The Foundations of Causal Decision Theory*. Cambridge University Press.
Kyburg, H. (1983), Recent work in inductive logic. [In]: Lucey and Tabor (eds.), *Recent Work in Philosophy*. Rowman and Allanfield, Totowa, N.J., pp. 89–150.
Keynes, J.(1921), *A Treatise on Probability*. Macmillan, London.
Łukasiewicz, J. (1913), *Die Logische Grundlagen der Wahrscheinlichkeitsrechnung*, Cracow; English translation in: L. Borkowski (ed.), *Jan Łukasiewicz – Selected Works*, North Holland & Polish Scientific Publishers, Amsterdam-Warsaw, 1970.
Mackie, J. L. (1963), The paradox of confirmation. *The British Journal for the Philosophy of Science* 13, 265–277.
Mackie, J. (1969), The relevance criterion of confirmation. *The British Journal for the Philosophy of Science* 20, 27–40.
Milne, P. (1995), A Bayesian defence of Popperian science? *Analysis* 55, 213–215.
Milne, P. (1996), Log [P(h|e)/P(h|b)] is the one true measure of confirmation. *Philosophy of Science* 63, 21–26.
Pawlak, Z. (1982), Rough Sets. *Int. J. of Computer and Information Sciences* 11, 341–356.
Pawlak, Z. (1991), *Rough Sets – Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Boston.
Pawlak, Z. (2002), Rough Sets, decision algorithms and Bayes' Theorem. *European Journal of Operational Research* 136, 181–189.
Pawlak, Z. (2003), Flow graphs and decision algorithms. [In]: G. Wang, Q. Liu, Y.Y. Yao, A. Skowron (eds.), *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*. LNAI 22639, Springer, Berlin, pp. 1–11.
Pawlak, Z. (2004), Decision networks. [In]: J.W. Grzymala-Busse, J. Komorowski, R. Słowiński, S. Tsumoto (eds.), *Rough Sets and Current Trends in Computing*. LNAI, Springer-Verlag, Berlin, in this volume.
Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufman, San Francisco.
Pollard, S. (1999), Milne's measure of confirmation. *Analysis* 59, 335–337.
Schlesinger, G. (1995), Measuring degrees of confirmation. *Analysis* 55, 208–212.
Rosenkrantz, R. (1994), Bayesian confirmation: paradise regained. *The British Journal for the Philosophy of Science* 45, 467–476.
Schum, D. (1994), *The Evidential Foundations of Probabilistic Reasoning*. Wiley, New York.
Tsumoto, S. (2002), Statistical test for rough set approximation based on Fisher's exact test. [In]: J. J. Alpigini, J. F. Peters, A. Skowron, N. Zhong (eds.), *Rough Sets and Current Trends in Computing*. LNAI 2475, Springer, Berlin, 2002, pp. 381-388.
Yao, Y.Y., and Zhong, N. (1999), An analysis of quantitative measures associated with rules. [In]: N. Zhong and L. Zhou (eds.), *Methodologies for Knowledge Discovery and Data Mining*. LNAI 1574, Springer, Berlin, pp. 479-488.
Zembowicz, R., Zytkow, J. (1996), From contingency tables to various forms of knowledge in databases. [In]: U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy, (eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press & The MIT Press, Menlo Park and Cambridge, pp. 329-349.