An Analysis of Quantitative Measures Associated with Rules

Y.Y. Yao[1] and Ning Zhong[2]

[1] Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada S4S 0A2
E-mail: yyao@cs.uregina.ca

[2] Department of Computer Science and Systems Engineering, Faculty of Engineering
Yamaguchi University, Tokiwa-Dai, 2557, Ube 755, Japan
E-mail: zhong@ai.csse.yamaguchi-u.ac.jp

Abstract. In this paper, we analyze quantitative measures associated with if-then type rules. Basic quantities are identified and many existing measures are examined using the basic quantities. The main objective is to provide a synthesis of existing results in a simple and unified framework. The quantitative measure is viewed as a multi-facet concept, representing the confidence, uncertainty, applicability, quality, accuracy, and interestingness of rules. Roughly, they may be classified as representing one-way and two-way supports.

1 Introduction

In machine learning and data mining, the discovered knowledge from a large data set is often expressed in terms of a set of if-then type rules [7, 21]. They represent relationships, such as correlation, association, and causation, among concepts. Typically, the number of potential rules observable in a large data set may be very large, and only a small portion of them is actually useful. In order to filter out useless rules, certain criteria must be established for rule selection. A common solution is the use of quantitative measures. One may select the rules which have the highest values. Alternatively, one may choose a threshold value and select rules whose measures are above the threshold value. The well known ID3 inductive learning algorithm [23] is an example of the former, and the approach for mining association rules in transaction databases [1] is an example of the latter. The use of quantitative measures also play a very important role in the interpretation of discovered rules, which provides guidelines for the proper uses of the rules.

Many quantitative measures have been proposed and studied, each of them captures different characteristics of rules. However, several important issues need more attention. Different names have been used for essentially the same measure, or a positive monotonic transformation of the same measure (called order preserving transformation [15]). Additional measures are being proposed, without realizing that the same measures have been studied in related fields such as expert systems, pattern recognition, information retrieval, and statistical data analysis. The relationships between various measures have not been fully investigated. There is clearly a need for a systematic study on the interpretation, classification, and axiomatization of quantitative measures associated with rules. Important initial studies have been reported by Piatetsky-Shapiro [25], and Major and Mangano [17] on the axiomatic characterization of quantitative measures, and by Klösgen [15] on the study of special classes of quantitative measures.

This paper may be viewed as a first step in the study of quantitative measures. A simple set-theoretic framework is suggested for interpreting if-then type rules. Basic quantities are identified and many existing measures are examined using the basic quantities. The results may lay down the groundwork for further systematic studies.

2 The Basic Framework and Basic Quantities

Consider an if-then type rule of the form:

$$\text{IF } E \text{ THEN } H \quad \text{with } \alpha_1, \ldots, \alpha_m, \tag{1}$$

which relates two concepts E and H. For clarity, we also simply write $E \longrightarrow H$. A rule does not necessarily represent a strict logical implication, with logical implication as the degenerate case. The values $\alpha_1, \ldots, \alpha_m$ quantifies different types of uncertainty and properties associated with the rule. In principle, one may connect any two concepts in the above rule form. The quantities $\alpha_1, \ldots, \alpha_m$ measures the degree or strength of relationships [34]. Examples of quantitative measures include confidence, uncertainty, applicability, quality, accuracy, and interestingness of rules.

We use the following set-theoretic interpretation of rules. It relates a rule to the data sets from which the rule is discovered. Let U denote a finite universe consisting of objects. Each object may be considered as one instance of a data set. If each object is described by a set of attribute-value pairs, the concepts E and H can be formally defined using certain languages, such as propositional and predicate languages [15]. We are not interested in the exact representation of the concepts. Instead, we focus on the set-theoretic interpretations of concepts [13, 18, 22]. For a concept E, let $m(E)$ denote the set of elements of U that satisfy the condition expressed by E. We also say that $m(E)$ is the set of elements satisfying E. Similarly, the set $m(H)$ consists of elements satisfying H. One may interpret m as a meaning function that associates each concept with a subset of U. The meaning function should obey the following conditions:

$$\begin{aligned} m(\neg E) &= U - m(E), \\ m(E \wedge H) &= m(E) \cap m(H), \\ m(E \vee H) &= m(E) \cup m(H), \end{aligned} \tag{2}$$

representing the sets of elements not satisfying E, satisfying both E and H, and satisfying at least one of E and H, respectively. With the meaning function m, a rule $E \longrightarrow H$ may be paraphrased as saying that "IF an element of the universe satisfies E, THEN the element satisfies H".

Using the cardinalities of sets, we obtain the following contingency table representing the quantitative information about the rule $E \longrightarrow H$:

	H	$\neg H$	Totals
E	$\|m(E) \cap m(H)\|$	$\|m(E) \cap m(\neg H)\|$	$\|m(E)\|$
$\neg E$	$\|m(\neg E) \cap m(H)\|$	$\|m(\neg E) \cap m(\neg H)\|$	$\|m(\neg E)\|$
Totals	$\|m(H)\|$	$\|m(\neg H)\|$	$\|U\|$

where $|\cdot|$ denotes the cardinality of a set. For clarity, we rewrite the table as follows:

	H	$\neg H$	Totals
E	a	b	$a+b$
$\neg E$	c	d	$c+d$
Totals	$a+c$	$b+d$	$a+b+c+d=n$

The values in the four cells are not independent. They are linked by the constraint $a+b+c+d=n$. The 2×2 contingency table has been used by many authors for representing information of rules [9, 11, 27, 29, 33]. From the contingency table, we can define some basic quantities.

The *generality* of E is defined by:

$$G(E) = \frac{|m(E)|}{|U|} = \frac{a+b}{n}, \tag{3}$$

which indicates the relative size of the concept E. A concept is more general if it covers more instances of the universe. If $G(E) = \alpha$, then $(100\alpha)\%$ of objects in U satisfy E. The quantity may be viewed as the probability of a randomly selected element satisfying E. Obviously, we have $0 \leq G(E) \leq 1$.

The *absolute support* of H provided by E is the quantity:

$$AS(H|E) = \frac{|m(H) \cap m(E)|}{|m(E)|} = \frac{a}{a+b}. \tag{4}$$

The quantity, $0 \leq AS(H|E) \leq 1$, shows the degree to which E implies H. If $AS(H|E) = \alpha$, then $(100\alpha)\%$ of objects satisfying E also satisfy H. It may be viewed as the conditional probability of a randomly selected element satisfying H given that the element satisfies E. In set-theoretic terms, it is the degree to which $m(E)$ is included in $m(H)$. Clearly, $AS(H|E) = 1$, if and only if $m(E) \subseteq m(H)$. The *change of support* of H provided by E is defined by:

$$CS(H|E) = AS(H|E) - G(H) = \frac{an - (a+b)(a+c)}{(a+b)n}. \tag{5}$$

Unlike the absolute support, the change of support varies from -1 to 1. One may consider $G(H)$ to be the prior probability of H and $AS(H|E)$ the posterior probability of H after knowing E. The difference of posterior and prior probabilities represents the change of our confidence regarding whether E actually causes H. For a positive value, one may say that E causes H; for a negative value, one may say that E does not cause H. The *mutual support* of H and E is defined by:

$$MS(E,H) = \frac{|m(E) \cap m(H)|}{|m(E) \cup m(H)|} = \frac{a}{a+b+c}. \tag{6}$$

One may interpret the mutual support, $0 \leq MS(E,H) \leq 1$, as a measure of the strength of the double implication $E \longleftrightarrow H$. It measures the degree to which E causes, and only causes, H. The mutual support can be reexpressed by:

$$MS(E,H) = 1 - \frac{|m(E) \varDelta m(H)|}{|m(E) \cup m(H)|}, \tag{7}$$

where $A \varDelta B = (A \cup B) - (A \cap B)$ is the symmetric difference between two sets. The measure $|A \varDelta B|/|A \cup B|$ is commonly known as the MZ metric for measuring distance between two sets [19]. Thus, MS may be viewed as a similarity measure of E and H.

The degree of *independence* of E and H is measured by:

$$IND(E,H) = \frac{G(E \wedge H)}{G(E)G(H)} = \frac{an}{(a+b)(a+c)}. \tag{8}$$

It is the ratio of the joint probability of $E \wedge H$ and the probability obtained if E and H are assumed to be independent. One may rewrite the measure of independence as [10]:

$$IND(E,H) = \frac{AS(H|E)}{G(H)}. \tag{9}$$

It shows the degree of the deviation of the probability of H in the subpopulation constrained by E from the probability of H in the entire data set [16, 31]. With this expression, the relationship to the change of support becomes clear. Instead of using the ratio, the latter is defined by the difference of $AS(H|E)$ and $G(H)$. When E and H are probabilistic independent, we have $CS(H|E) = 0$ and $IND(E,H) = 1$. Moreover, $CS(H|E) \geq 0$ if and only if $IND(E,H) \geq 1$, and $CS(H|E) \leq 0$ if and only if $IND(E,H) \leq 1$. This provides further support for use of CS as a measure of confidence that E causes H. However, CS is not a symmetric measure, while IND is symmetric. The difference of $G(H \wedge E)$ and $G(H)G(E)$:

$$D(H,E) = G(H \wedge E) - G(H)G(E), \tag{10}$$

is a symmetric measure. Compared with $D(H,E)$, the measure $CS(H|E)$ may be viewed as a relative difference.

The generality of a concept is related to the probability that a randomly selected element will be an instance of the concept. It is the basic quantity from which all other quantities can be expressed as follows:

$$
\begin{aligned}
AS(H|E) &= \frac{G(H \wedge E)}{G(E)}, \\
CS(H|E) &= \frac{G(H \wedge E) - G(H)G(E)}{G(E)}, \\
MS(E,H) &= \frac{G(E \wedge H)}{G(E \vee H)}, \\
IND(E,H) &= \frac{G(E \wedge H)}{G(E)G(H)}, \\
D(H,E) &= G(H \wedge E) - G(H)G(E).
\end{aligned} \tag{11}
$$

From the above definitions, we can establish the following relationships:

$$
\begin{aligned}
G(E) &= AS(E|U), \\
CS(H|E) &= (IND(E,H) - 1)G(H), \\
AS(H|E) &= \frac{G(H)}{G(E)} AS(E|H), \\
MS(E,H) &= \frac{1}{\frac{1}{AS(E|H)} + \frac{1}{AS(H|E)} - 1}, \\
D(H,E) &= CS(H|E)G(E).
\end{aligned} \tag{12}
$$

In summary, all measures introduced in this section have a probability related interpretation. They can be roughly divided into three classes:

generality:	G,
one-way association (single implication):	AS, CS,
two-way association (double implication):	MS, IND, D.

Each type of association measures can be further divided into absolute support and change of support. The measure of absolute one-way support is AS, and the measure of absolute two-way support is MS. The measures of change of support are CS for one-way, and IND and D for two-way. It is interesting to note that all measures of change of support are related to the *deviation* of joint probability of $E \wedge H$ from the probability obtained if E and H are assumed to be independent. In other words, a stronger association is presented if the joint probability is further away from the probability under independence. The association can be either positive or negative.

3 A Review of Existing Measures

This section is not intended to be an exhaustive survey of quantitative measures associated with rules. We will only review some of the measures that fit in the framework established in the last section.

3.1 Generality

The generality is one of the two standard measures used for mining association rules [1]. For rule $E \longrightarrow H$, the generality:

$$G(E \wedge H) = \frac{a}{n} \tag{13}$$

is commonly known as the *support* of the rule. It represents the percentage of positive instances of E that support the rule. On the other hand, the generality $G(E)$ is the percentage of instances to which the rule can be applied. Iglesia *et al.* [13] called the quantity $G(E)$ the *applicability* of the rule. Klösgen [15] referred to it as a measure of *coverage* of the concept E.

3.2 One-way support

The absolute support $AS(H|E)$ is the other standard measure used for mining association rules [1], called *confidence* of the rule $E \longrightarrow H$. Different names were given to this measure, including the *accuracy* [13, 29], *strength* [8, 15, 26], and *certainty factor* [15]. In the context of information retrieval, the same measure is referred to as the measure of *precision* [32]. Tsumoto and Tanaka [29] used the quantity $AS(E|H)$ for measuring the *coverage* or *true positive rate*. It is regarded as a measure of sensitivity by Klösgen [15]. The same measure was also used by Choubey *et al.* [5]. In the context of information retrieval, the measure is referred to as the measure of *recall* [32]. The use of change of support $CS(H|E)$ was discussed by some authors [4, 25].

Additional measures of one-way support can be obtained by combining basic quantities introduced in the last section. Yao and Liu [31] used the following quantity for measuring the significance of a rule $E \longrightarrow H$:

$$S_1(H|E) = AS(H|E) \log IND(E,H) = \frac{a}{a+b} \log \frac{an}{(a+b)(a+c)}. \tag{14}$$

The measure is a product of a measure of one-way support $AS(S|E)$ and the logarithm of a measure of two-way support $IND(E,H)$. Since logarithm is a monotonic increasing function, it reflects the properties of $IND(E,H)$. Gray and Orlowska [10] proposed a measure of one-way support, called measure of rule interestingness, by combining generality and absolute support:

$$i(H|E) = \left(IND(E,H)^l - 1\right) G(E \wedge H)^m = ((\frac{an}{(a+b)(a+c)})^l - 1)(\frac{a}{n})^m. \tag{15}$$

where l and m are parameters to weigh the relative importance of the two measures. Klösgen [15] studied another class of measures:

$$K(H|E) = G(E)^\alpha (AS(H|E) - G(H)). \tag{16}$$

It is a combination of generality and change of support. When $\alpha = 0$, the measure reduced to the change of support.

Following Duda, Gasching, and Hart [6], Kamber and Shinghal [14], Schlimmer and Granger [24] used the measure of *logical sufficiency*:

$$LS(H|E) = \frac{AS(E|H)}{AS(E|\neg H)} = \frac{a(b+d)}{b(a+c)}. \tag{17}$$

and the measure of *logical necessity*:

$$LN(H|E) = \frac{AS(\neg E|H)}{AS(\neg E|\neg H)} = \frac{c(b+d)}{d(a+c)}. \tag{18}$$

McLeish *et al.* [20] viewed LS as the weight of evidence if one treats E as a piece of evidence. A highly negative weight implies that there is significant reason to belief in $\neg H$, and a positive weight supports H. It should pointed out that weight of evidence plans an important rule in Bayesian inference. Ali, Manganaris and Srikant [2] defined the relative risk of a rule $E \longrightarrow H$ as follows:

$$r(H|E) = \frac{AS(H|E)}{AS(H|\neg E)} = \frac{a(c+d)}{c(a+b)}. \tag{19}$$

It is fact related to the measure LS, if one change the places of E and H.

Based on the probability related interpretation of $AS(H|E)$, Smyth and Goodman [28] defined the *information content* of rules. For $E \longrightarrow H$, we have:

$$\begin{aligned} J(H||E) &= G(E)\left(AS(H|E)\log\frac{AS(H|E)}{G(H)} + AS(\neg H|E)\log\frac{AS(\neg H|E)}{G(\neg H)}\right) \\ &= G(E)\left(AS(H|E)\log IND(E,H) + AS(\neg H|E)\log IND(\neg H,E)\right) \\ &= \frac{1}{n}\left(a\log\frac{an}{(a+b)(a+c)} + b\log\frac{bn}{(a+b)(b+d)}\right). \end{aligned} \tag{20}$$

This measure is closely related to the *divergence* measure proposed by Kullback and Leibler [12].

3.3 Two-way support

The measure of independence IND has been used by many authors. Silverstein *et al.* [27] referred to it as a measure of interest. Büchter and Wirth [3] regarded it as a measure of dependence. Gray and Orlowska [10] used the same measure, and provided the interpretation given by equation (9).

The measure of two-way support corresponding to equation (14) is given by Yao and Liu [31] as:

$$S_2(E,H) = G(E \wedge H)\log IND(E,H) = \frac{a}{n}\log\frac{an}{(a+b)(a+c)}. \tag{21}$$

By setting $l = m = 1$ in equation (15), we have:

$$i(H|E) = IND(E,H)D(E,H), \tag{22}$$

which is a multiplication of two basic measures of two-way support. By setting $\alpha = 1$ in equation (16), we immediately obtain the measure D.

The measure of two support corresponding to the measure of divergence (20) is given by the measure of mutual information. For rule $E \longrightarrow H$, we have:

$$\begin{aligned} M(E;H) = G(E \wedge H) \log \frac{G(E \wedge H)}{G(E)G(H)} + G(E \wedge \neg H) \log \frac{G(E \wedge \neg H)}{G(E)G(\neg H)} + \\ G(\neg E \wedge H) \log \frac{G(\neg E \wedge H)}{G(\neg E)G(H)} + G(\neg E \wedge \neg H) \log \frac{G(\neg E \wedge \neg H)}{G(\neg E)G(\neg H)} \\ = \frac{1}{n} \left(a \log \frac{an}{(a+c)(a+b)} + b \log \frac{bn}{(b+d)(a+b)} + \right. \\ \left. c \log \frac{cn}{(a+c)(c+d)} + d \log \frac{dn}{(b+d)(c+d)} \right) \end{aligned} \quad (23)$$

The relationship between J and M can be established as:

$$M(E;H) = J(H||E) + J(H||\neg E). \quad (24)$$

By extending the above relationship, in general one may obtain measures of two-way support by combining measures of one-way support. For example, both $AS(H|E) + AS(E|H)$ and $AS(H|E)AS(E|H)$ are measures of two-way support.

3.4 Axioms for quantitative measures of rules

Piatetsky-Shapiro [25] suggested that a quantitative measure of rule $E \longrightarrow H$ may be computed as a function of $G(E)$, $G(H)$, $G(E \wedge H)$, rule complexity, and possibly other parameters such as the mutual distribution of E and H or the domain size of E and H. For the evaluation of rules, Piatetsky-Shapiro [25] introduced three axioms. Major and Mangano [17] added a fourth axioms. Klösgen [15] studied a special class of measures that are characterized by two quantities, the absolute one-way support $AS(H|E)$ and the generality $G(E)$. The generality $G(H \wedge E)$ is obtained by $AS(H|E)G(E)$. Suppose $Q(E,H)$ is a measure associated with rule $E \longrightarrow H$. The version of the four axioms given by Klösgen [15] is:

(i). $Q(E,H) = 0$ if E and H are statistically independent,
(ii). $Q(E,H)$ monotonically increases in $AS(H|E)$ for fixed $G(E)$,
(iii). $Q(E,H)$ monotonically decreases in $G(E)$ for fixed $G(E \wedge H)$,
(iv). $Q(E,H)$ monotonically increases in $G(E)$ for fixed $AS(H|E) > G(H)$.

Axiom (i) implies that only measures of change of support are considered. Other axioms states that all measures must have the property of monotonicity. Many of the measures discussed in this paper fall into this class.

4 Conclusion

We have presented a simple and unified framework for the study of quantitative measures associated with rules. Some basic measures have been proposed and studied. Many existing measures have been investigated in terms of these basic measures.

This paper is a preliminary step towards a systematic study on quantitative measures associated with rules. Further investigations on the topic are planed. We will examine the semantics and implications of various measures, and study axioms for distinct types of measures.

References

1. Agrawal, R., Imielinski, T., and Swami, A. Mining association rules between sets of items in large databases, *Proceedings of the ACM SIGMOD International Conference on the Management of Data*, 207-216, 1993.
2. Ali, K., Manganaris, S., and Srikant, R. Partial classification using association rules, *Proceedings of KDD-97*, 115-118, 1997.
3. Büchter, O. and Wirth R. Discovery of association rules over ordinal data: a new and faster algorithm and its application to basket analysis, in [30], 36-47, 1998.
4. Chen, M., Han, J. and Yu, P.S. Data mining, an overview from a databases perspective, *IEEE Transactions on Knowledge and Data Engineering*, **8**, 866-883, 1996.
5. Choubey, S.K., Deogun, J.S., Raghavan, V.V., and Sever, H. Comparison of classification methods, *Proceedings of 1997 Joint Conference in Information Sciences*, 371-374, 1997.
6. Duda, R.O., Gasching, J., and Hart, P.E. Model design in the Prospector consultant system for mineral exploration, in: Webber, B.L. and Nilsson, N.J. (Eds.), *Readings in Artificial Intelligence*, Tioga, Palo Atlto, CA, 334-348, 1981.
7. Fayyad, U.M, Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (Eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press / MIT Press, California, 1996.
8. Fhar, V. and Tuzhilin, A. Abstract-driven pattern discovery in databases, *IEEE Transactions on Knowledge and Data Engineering*, **5**, 926-938, 1993.
9. Gaines, B.R. The trade-off between knowledge and data in knowledge acquisition, in: Piatetsky-Shapiro, G. and Frawley, W.J. (Eds.), *Knowledge Discovery in Databases*, AAAI/MIT Press, 491-505, 1991.
10. Gray, B. and Orlowska, M.E. CCAIIA: clustering categorical attributes into interesting association rules, in [30], 132-143, 1998.
11. Ho, K.M. and Scott, P.D. Zeta: a global method for discretization of continuous variables, *Proceedings of KDD-97*, 191-194, 1997.
12. Kullback, S. and Leibler, R.A. On information and sufficiency, *Annals of Mathematical Statistics*, **22**, 79-86, 1951.
13. Iglesia, B. Debuse, J.C.W. and Rayward-Smith V.J. Discovering knowledge in commercial databases using modern heuristic techniques, *Proceedings of KDD-96*, 44-49, 1996.
14. Kamber, M. and Shinghal, R. Evaluating the interestingness of characteristic rules, *Proceedings of KDD-96*, 263-266, 1996.

15. Klösgen, W. Explora: a multipattern and multistrategy discovery assistant, in [7], 249-271, 1996.
16. Liu, H., Lu, H., and Yao, J. Identifying relevant databases for multidatabase mining, in [30], 211-221, 1998.
17. Major, J. and Mangano, J. Selecting among rules induced from a hurricane database, *The Journal of Intelligent Information Systems*, **4**, 1995.
18. Mannila, H. and Toivonen, H. Multiple uses of frequent sets and condensed representations, *Proceedings of KDD-96*, 189-194, 1996.
19. Marczewski, E. and Steinhaus, H. On a certain distance of sets and the corresponding distance of functions, *Colloquium Mathemmaticum*, **6**, 319-327, 1958.
20. McLeish, M., Yao, P., Garg, M., and Stirtzinger, T. Discovery of medical diagnostic information: an overview of methods and results, in: Piatetsky-Shapiro, G. and Frawley, W.J. (Eds.), *Knowledge Discovery in Databases*, AAAI/MIT Press, 477-490, 1991.
21. Michalski, R.S., Carbonell, J.G., and Mitchell, T.M. (Eds.), *Machine Learning*, Tioga, 1983.
22. Pawlak, Z. *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Boston, 1991.
23. Quinlan, J.R. Induction of decision trees, *Machine Learning*, **1**, 81-106, 1986.
24. Schlimmer, J.C. and Granger, Jr. R.H. Incremental learning from noisy data, *Machine Learning*, **1**, 317-354, 1986.
25. Piatetsky-Shapiro, G. Discovery, analysis, and presentation of strong rules, in: Piatetsky-Shapiro, G. and Frawley, W.J. (Eds.), *Knowledge Discovery in Databases*, AAAI/MIT Press, 229-238, 1991.
26. Shen, W. and Leng, B. Metapattern generation for integrated data mining, *Proceedings of KDD-96*, 152-157, 1996.
27. Silverstein, C., Brin, S., and Motwani, R. Beyond market baskets: generalizing association rules to dependence rules, *Data Mining and Knowledge Discovery*, **2**, 39-68, 1998.
28. Smyth, P. and Goodman, R.M. Rule induction using information theory, in: Piatetsky-Shapiro, G. and Frawley, W.J. (Eds.), *Knowledge Discovery in Databases*, AAAI/MIT Press, 159-176, 1991.
29. Tsumoto, S. and Tanaka, H. Automated discovery of functional components of proteins from amino-acid sequences based on rough sets and change of representation, *Proceedings of KDD-95*, 318-324, 1995.
30. Wu, X., Kotagiri, R,. and Bork, K.B. (Eds.), *Research and Development in Knowledge Discovery and Data Mining*, Springer, Berlin, 1998.
31. Yao, J. and Liu, H. Searching multiple databases for interesting complexes, in: Lu, H., Motoda, H., and Liu, H. (Eds.), *KDD: Techniques and Applications*, World Scientific, Singapore, 1997.
32. Yao, Y.Y. Measuring retrieval performance based on user preference of documents, *Journal of the American Society for Information Science*, **46**, 133-145, 1995.
33. Zembowicz, R. and Żytkow, J.M. From contingency tables to various forms of knowledge in database, in [7], 39-81, 1996.
34. Zhong, N., Dong, J., Fujitsu, S., and Ohsuga, S. Soft techniques for rule discovery in data, *Transactions of Information Processing Society of Japan*, **39**, 2581-2592, 1998.

This article was processed using the LaTeX macro package with LLNCS style