



Representative Rule Templates for Association Rules Satisfying Multiple Canonical Evaluation Criteria

Marzena Kryszkiewicz^(✉)

Institute of Computer Science, Warsaw University of Technology,
Nowowiejska 15/19, 00-665 Warsaw, Poland
mkr@ii.pw.edu.pl

Abstract. Originally, strong association rules were defined as those that have sufficiently high values of two parameters: support and confidence. However, it has been shown in the literature that in general neither of these two measures is capable of determining (in)dependence between rules' constituents correctly. Thus, usage of other measures for rule evaluation is also under research. In this paper, we formulate a generic notion of a canonical measure and show important examples of canonical measures. For association rules satisfying any set of criteria based on canonical measures, we offer their concise lossless representation in the form of so called representative rule templates. Also, we derive a number of properties of this representation.

1 Introduction

Originally, strong association rules were defined in [1] as those that have sufficiently high values of two parameters: (relative) support (probability that an association rule is satisfied) and confidence (conditional probability of the rule consequent given its antecedent). However, it has been shown in the literature that in general neither of these two measures is capable of determining (in)dependence between rules' constituents correctly (see e.g. [12]). Thus, usage of other measures for rule evaluation is also under research [3, 5, 9, 10, 13–18]. In this paper, we formulate a generic notion of a *canonical measure*. For a number of example evaluation measures of association rules, we show that they are canonical. For association rules satisfying any set of criteria based on canonical measures, we offer their concise lossless representation in the form of so called *representative rule templates*. Also, we derive a number of properties of this representation.

Our paper has the following layout. In Sect. 2, we briefly recall and comment basic notions of itemsets, (association) rules, support, confidence, dependency of events, ACBC-measures, and a good interestingness measure. Our main contribution is presented in Sects. 3–5. In Sect. 3, we introduce and justify usefulness of the notion of a canonical measure and show that a number of example ACBC-measures are canonical. In Sect. 4, we introduce the notion of rule templates and show that representative rule templates are sufficient to represent all association rules that are strong with respect to any criteria expressible in terms of canonical measures. Also, we derive a number of

properties of (representative) rule templates. In particular, we derive a formula determining how many association rules are covered by a rule template. In Sect. 5, we briefly address related work on lossless representations of association rules and make additional claims about our new representation. Section 6 concludes our contribution.

2 Basic Notions and Properties

In this section, we provide definitions and properties related to *rules* and their specific type called *association rules* [1]. Let \mathcal{I} be a set of *items* (e.g. products). Any set $X \subseteq \mathcal{I}$ is called an *itemset*. A *transaction dataset* is denoted by \mathcal{D} and is defined as a set of itemsets. Each itemset in \mathcal{D} is called a *transaction*.

Let X and Y be any itemsets. An expression $X \rightarrow Y$ is called a *rule*. If additionally $X \cap Y = \emptyset$, then $X \rightarrow Y$ is called an *association rule*. The set of all association rules is denoted by AR . Itemsets X and Y , which occur in $X \rightarrow Y$, are called its *antecedent* and *consequent*, respectively. The itemset $Z = X \cup Y$ is called the *base* of $X \rightarrow Y$.

Itemsets and rules are typically characterized by *support* (or *relative support*) and *confidence* as follows. *Support* of an itemset X is denoted by $sup(X)$ and is defined as the number of transactions in \mathcal{D} that contain X . Alternatively, instead of a support, a notion of a relative support is used: *Relative support* of an itemset X is defined as $sup(X)/|\mathcal{D}|$. The relative support of itemset X can be regarded as the probability of the occurrence of X in a transaction. The relative support of X will be denoted by $P(X)$. Clearly, if $X \subseteq Y$, then $sup(X) \geq sup(Y)$ and $P(X) \geq P(Y)$.

Support of a rule $X \rightarrow Y$ is denoted by $sup(X \rightarrow Y)$ and is defined as the support of its base $X \cup Y$; that is, $sup(X \rightarrow Y) = sup(X \cup Y)$.

The *confidence* of a rule $X \rightarrow Y$ is denoted by $conf(X \rightarrow Y)$ and is defined as the conditional probability that Y occurs in a transaction provided X occurs in the transaction; that is: $conf(X \rightarrow Y) = sup(X \rightarrow Y)/sup(X) = P(XY)/P(X)$.

An association rule is defined in [1] as *strong* if its (relative) support and confidence are greater than or equal to user defined *minimum (relative) support* and *minimum confidence thresholds*, respectively. Nevertheless, antecedents and consequents of rules that are strong with respect to support and confidence are not guaranteed to be dependent. Statistically, X and Y are *dependent* if $P(XY) \neq P(X) \times P(Y)$. Otherwise, X and Y are *independent*. The *dependence* between X and Y is considered as *positive* if $P(XY) > P(X) \times P(Y)$, and *negative* if $P(XY) < P(X) \times P(Y)$. The example beneath shows the case when an antecedent and consequent of a rule are dependent negatively despite high values of (relative) support and confidence.

Example 1. Let \mathcal{D} be a transaction dataset from Table 1 and $\{z\} \rightarrow \{v\}$ be a rule of interest. The probabilities of the antecedent, consequent and base of rule $\{z\} \rightarrow \{v\}$ are as follows: $P(\{z\}) = 0.5$, $P(\{v\}) = 0.9$, $P(\{zv\}) = 0.4$ (see Table 2). Since $P(\{z\}) \times P(\{v\}) = 0.45$, which is greater than $P(\{zv\})$, z and v are dependent negatively. On the other hand, $conf(\{z\} \rightarrow \{v\}) = 0.4/0.5 = 0.8$, which is a high value. \square

Table 1. Example dataset \mathcal{D}

Transaction Id	Transaction
#1	{v}
#2	{v}
#3	{x v}
#4	{x v}
#5	{x y v}
#6	{x y z v}
#7	{x y z v}
#8	{y z v}
#9	{y z v}
#10	{y z}

Table 2. Example rules found in \mathcal{D} from Table 1

Rule	{x} → {y}	{z} → {y}	{z} → {v}
$P(\text{antecedent})$	0.5	0.5	0.5
$P(\text{consequent})$	0.6	0.6	0.9
$P(\text{base})$	0.3	0.5	0.4
$\text{conf}(\text{rule})$	1/5	1	4/5
$P(\text{antecedent}) \times P(\text{consequent})$	0.3	0.3	0.45
Are consequent and antecedent dependent?	no	yes - positively	yes - negatively

Please note that (in)dependence of antecedent X and consequent Y of rule $X \rightarrow Y$ is determined based on three probabilities: $P(X)$, $P(Y)$ and $P(XY)$. In fact, large number of rule evaluation measures can be defined in terms of at most these three probabilities. Such measures were called *ACBC-measures* (Antecedent-Consequent-Base-Constants-measures) in [11]. Some of their popular examples are provided in Table 3.

Table 3. Example ACBC-measures of (association) rules

Measure	Definition
$\text{relativeSup}(X \rightarrow Y)$	$P(XY)$
$\text{conf}(X \rightarrow Y)$	$P(XY) / P(X)$
$\text{novelty}(X \rightarrow Y)$	$P(XY) - P(X) \times P(Y)$
$\text{lift}(X \rightarrow Y)$	$P(XY) / (P(X) \times P(Y))$
$\text{cosine}(X \rightarrow Y)$	$P(XY) / \sqrt{P(X) \times P(Y)}$
$\text{Jaccard}(X \rightarrow Y)$	$P(XY) / (P(X) + P(Y) - P(XY))$
$\text{accuracy}(X \rightarrow Y)$	$P(XY) + P(\bar{X}\bar{Y}) = 1 + 2P(XY) - P(X) - P(Y)$
$F\text{-score}(X \rightarrow Y)$	$\frac{(P(XY)/P(X)) \times (P(XY)/P(Y))}{((P(XY)/P(X)) + (P(XY)/P(Y))) / 2} = \frac{2P(XY)}{P(X) + P(Y)}$

In [15], a rule evaluation measure μ was defined as a *good interestingness measure* if the following conditions were fulfilled for any rules $X \rightarrow Y, X' \rightarrow Y'$:

1. If $P(XY) = P(X) \times P(Y)$, then $\mu(X \rightarrow Y) = 0$.
2. If $P(X') = P(X)$, $P(Y') = P(Y)$ and $P(X'Y') > P(XY)$, then $\mu(X' \rightarrow Y') > \mu(X \rightarrow Y)$.
3. If $P(X') = P(X)$, $P(Y') < P(Y)$ and $P(X'Y') = P(XY)$, then $\mu(X' \rightarrow Y') > \mu(X \rightarrow Y)$.
4. If $P(X') < P(X)$, $P(Y') = P(Y)$ and $P(X'Y') = P(XY)$, then $\mu(X' \rightarrow Y') > \mu(X \rightarrow Y)$.

Nevertheless, a number of rule evaluation measures, including all measures listed in Table 3, do not satisfy at least one of the postulates. In particular, relative support does not satisfy postulates 1, 3, 4; confidence does not satisfy postulates 1, 3; novelty does

not satisfy postulate 3 (if $P(X) = 0$) and postulate 4 (if $P(Y) = 0$); lift, cosine, Jaccard, accuracy and F-score do not fulfill postulate 1.

In the next section, we offer a generic notion of a *canonical measure*, which embraces a large number of ACBC-measures (including all measures from Table 3).

3 Canonical Measures for Evaluating Rules

Let us consider two rules $X \rightarrow Y$ and $X' \rightarrow Y'$ such that X' occurs in \mathcal{D} no more frequently than X ($P(X') \leq P(X)$), Y' occurs in \mathcal{D} no more frequently than Y ($P(Y') \leq P(Y)$), but $X' \cup Y'$ occurs in \mathcal{D} no less frequently than $X \cup Y$ ($P(X'Y') \geq P(XY)$). In such a case, if the dependence between Y' and X' is positive, then it is natural to expect that Y and X are either dependent positively, but to degree that is not higher than for Y' and X' , or are independent, or are dependent negatively. A rule measure reflecting this expectation would evaluate rule $X' \rightarrow Y'$ not lower than rule $X \rightarrow Y$. A notion of a *canonical measure*, which we offer in this paper, is based on this observation.

Example 2. Let \mathcal{D} be a transaction dataset from Table 1 and $\{x\} \rightarrow \{y\}$, $\{z\} \rightarrow \{y\}$ and $\{z\} \rightarrow \{v\}$ be rules under investigation. In Table 2, we provide the values of probabilities of their antecedents, consequents and bases. We note that the probabilities of antecedents of $\{x\} \rightarrow \{y\}$ and $\{z\} \rightarrow \{y\}$ are the same and that the probabilities of their consequents are the same, but the probabilities of their bases are different; namely, $P(\{zy\}) > P(\{xy\})$, and the dependence between y and z is positive, while y and x are independent.

Let us now compare rule $\{z\} \rightarrow \{y\}$ with rule $\{z\} \rightarrow \{v\}$. The probabilities of their antecedents are the same, but the probability of the consequent of $\{z\} \rightarrow \{y\}$ is lower than the probability of the consequent of $\{z\} \rightarrow \{v\}$, while the probability of the co-occurrence of $\{zy\}$ is greater than the probability of the co-occurrence of $\{zv\}$. In this case, the dependence between v and z is negative in contrast to the positive dependence between y and z . \square

We define a *measure* $\mu : [0..1] \times [0..1] \times [0..1] \rightarrow R$ as *canonical* if for any values $p_a, p_c, p_b \in [0, 1]$ and $p_{a'}, p_{c'}, p_{b'} \in [0, 1]$ such that $(p_a, p_c \geq p_b)$, $(p_{a'}, p_{c'} \geq p_{b'})$, $(p_{a'} \leq p_a)$, $(p_{c'} \leq p_c)$ and $(p_{b'} \geq p_b)$ the following holds:

$$\mu(p_{a'}, p_{c'}, p_{b'}) \geq \mu(p_a, p_c, p_b).$$

Proposition 1. Let μ be a canonical measure, $X \rightarrow Y$ and $X' \rightarrow Y'$ be rules such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(XY)$. Then

$$\mu(P(X'), P(Y'), P(X'Y')) \geq \mu(P(X), P(Y), P(XY)).$$

Proof. Let μ be a canonical measure, $X \rightarrow Y$ and $X' \rightarrow Y'$ be rules such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(XY)$. Let $p_a = P(X)$, $p_c = P(Y)$, $p_b = P(XY)$, $p_{a'} = P(X')$, $p_{c'} = P(Y')$, $p_{b'} = P(X'Y')$. Then, $p_a, p_c, p_b \in [0, 1]$, $p_{a'}, p_{c'}, p_{b'} \in [0, 1]$, $(p_a, p_c \geq p_b)$, $(p_{a'}, p_{c'} \geq p_{b'})$, $(p_{a'} \leq p_a)$, $(p_{c'} \leq p_c)$ and $(p_{b'} \geq p_b)$.

Hence and by the fact that μ is a canonical measure, $\mu(p_{a'}, p_{c'}, p_{b'}) \geq \mu(p_a, p_c, p_b)$. Therefore, $\mu(P(X'), P(Y'), P(X'Y')) \geq \mu(P(X), P(Y), P(XY))$. \square

In the remainder of the paper, when using a canonical measure to evaluate a rule $X \rightarrow Y$, we will write interchangeably $\mu(P(X), P(Y), P(XY))$ and $\mu(X \rightarrow Y)$. Thus, Proposition 1 can be rewritten equivalently as Proposition 2:

Proposition 2. Let μ be a canonical measure, $X \rightarrow Y$ and $X' \rightarrow Y'$ be rules such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(XY)$. Then, $\mu(X' \rightarrow Y') \geq \mu(X \rightarrow Y)$.

It can be easily shown for many ACBC-measures that they are canonical. In Proposition 3, we show that in fact all ACBC-measures from Table 3 are canonical.

Proposition 3. Let $X \rightarrow Y$ and $X' \rightarrow Y'$ be rules such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(XY)$. Then:

- (a) $relativeSup(X' \rightarrow Y') \geq relativeSup(X \rightarrow Y)$
- (b) $conf(X' \rightarrow Y') \geq conf(X \rightarrow Y)$
- (c) $novelty(X' \rightarrow Y') \geq novelty(X \rightarrow Y)$
- (d) $lift(X' \rightarrow Y') \geq lift(X \rightarrow Y)$
- (e) $cosine(X' \rightarrow Y') \geq cosine(X \rightarrow Y)$
- (f) $Jaccard(X' \rightarrow Y') \geq Jaccard(X \rightarrow Y)$
- (g) $accuracy(X' \rightarrow Y') \geq accuracy(X \rightarrow Y)$
- (h) $F - score(X' \rightarrow Y') \geq F - score(X \rightarrow Y)$

Proof. Let $X \rightarrow Y$ and $X' \rightarrow Y'$ are rules such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(XY)$. Then:

Ad (a) $relativeSup(X' \rightarrow Y') = P(X'Y') \geq P(XY) = relativeSup(X \rightarrow Y)$.

Ad (b) $conf(X' \rightarrow Y') = \frac{P(X'Y')}{P(X')} \geq \frac{P(XY)}{P(X)} = conf(X \rightarrow Y)$.

Ad (c) $novelty(X' \rightarrow Y') = P(X'Y') - P(X') \times P(Y') \geq P(XY) - P(X) \times P(Y) = novelty(X \rightarrow Y)$.

Ad (d–h) Proof in these cases is analogous to the proof of case (c). \square

In the remainder of the paper, we assume that a set $\mathcal{M} = \{\mu_1, \dots, \mu_n\}$ of canonical measures is used to evaluate (association) rules with respect to their corresponding threshold values $E = \{\varepsilon_1, \dots, \varepsilon_n\}$. An expression $\mu_i(X \rightarrow Y) \geq \varepsilon_i$, where μ_i is a canonical measure and ε_i is a threshold value, will be called a *canonical evaluation criterion*.

An association rule $X \rightarrow Y$ will be called (\mathcal{M}, E) -strong if $\forall \mu_i \in \mathcal{M} (\mu_i(X \rightarrow Y) \geq \varepsilon_i)$. The set of all (\mathcal{M}, E) -strong association rules will be denoted by $AR_{(\mathcal{M}, E)}$.

4 Representative Rule Templates

In this section, we will first introduce a new notion of a *rule template* and examine its properties and relationship with association rules. Then, we will offer a lossless representation of (\mathcal{M}, E) -strong association rules based on so called (\mathcal{M}, E) -representative rule templates.

4.1 Rule Templates and Association Rules

A construct $X \xrightarrow{Z} Y$ will be called a *rule template* if $X \cap Y = \emptyset$ and $(X \cup Y) \subseteq Z$. Itemsets X , Y and Z , which occur in the rule template $X \xrightarrow{Z} Y$, are called its *antecedent*, *consequent* and *base*, respectively. The set of all rule templates is denoted by RT .

Please note that for $Z = X \cup Y$, rule template $X \xrightarrow{Z} Y$ coincides with association rule $X \rightarrow Y$. However, in general case, the base Z of rule template $X \xrightarrow{Z} Y$ is a (proper or improper) superset of the base $X \cup Y$ of association rule $X \rightarrow Y$. In spite of this difference, both the probability of base $X \cup Y$ of rule $X \rightarrow Y$ as well as the probability of base Z of rule template $X \xrightarrow{Z} Y$ exceeds neither $P(X)$ nor $P(Y)$.

Let μ be a canonical measure. We redefine μ for a rule template $X \xrightarrow{Z} Y$ as follows:

$$\mu(X \xrightarrow{Z} Y) = \mu(P(X), P(Y), P(Z)).$$

For example, $\text{novelty}(X \rightarrow Y) = P(XY) - P(X) \times P(Y)$, while $\text{novelty}(X \xrightarrow{Z} Y) = P(Z) - P(X) \times P(Y)$.

A rule template $X \xrightarrow{Z} Y$ is called (\mathcal{M}, E) -strong if $\forall \mu_i \in \mathcal{M} (\mu_i(X \xrightarrow{Z} Y) \geq \varepsilon_i)$. The set of all (\mathcal{M}, E) -strong rule templates will be denoted by $RT_{(\mathcal{M}, E)}$.

Proposition 4. Let $X \xrightarrow{Z} Y$ be a rule template and μ be a canonical measure. Then for each rule $X' \rightarrow Y'$ such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(Z)$, the following holds: $\mu(X' \rightarrow Y') \geq \mu(X \xrightarrow{Z} Y)$.

Proof. Let $X \xrightarrow{Z} Y$ be a rule template, μ be a canonical measure and $X' \rightarrow Y'$ be a rule such that $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(Z)$. Let $p_a = P(X)$, $p_c = P(Y)$, $p_b = P(XY)$, $p_{a'} = P(X')$, $p_{c'} = P(Y')$, $p_{b'} = P(X'Y')$. Then, $p_a, p_c, p_b \in [0, 1]$, $p_{a'}, p_{c'}, p_{b'} \in [0, 1]$, $(p_a, p_c \geq p_b)$, $(p_{a'}, p_{c'} \geq p_{b'})$, $(p_{a'} \leq p_a)$, $(p_{c'} \leq p_c)$ and $(p_{b'} \geq p_b)$. Hence and by the fact that μ is a canonical measure, $\mu(p_{a'}, p_{c'}, p_{b'}) \geq \mu(p_a, p_c, p_b)$. So, $\mu(X' \rightarrow Y') = \mu(P(X'), P(Y'), P(X'Y')) \geq \mu(P(X), P(Y), P(Z)) = \mu(X \xrightarrow{Z} Y)$. \square

Proposition 5. Let $X \xrightarrow{Z} Y$ be a rule template and μ be a canonical measure. Then for each rule $X' \rightarrow Y'$ such that $X' \supseteq X$, $Y' \supseteq Y$, $(X' \cup Y') \subseteq Z$, the following holds: $\mu(X' \rightarrow Y') \geq \mu(X \xrightarrow{Z} Y)$.

Proof. Let $X \xrightarrow{Z} Y$ be a rule template and μ be a canonical measure (*). Let $X' \rightarrow Y'$ be a rule such that $X' \supseteq X$, $Y' \supseteq Y$, $(X' \cup Y') \subseteq Z$. Then, $P(X') \leq P(X)$, $P(Y') \leq P(Y)$ and $P(X'Y') \geq P(Z)$. Hence, by (*) and Proposition 4, $\mu(X' \rightarrow Y') \geq \mu(X \xrightarrow{Z} Y)$. \square

Proposition 5 tells us that whenever a rule template $X \xrightarrow{Z} Y$ reaches some value for a canonical measure, then a number of association rules $X' \rightarrow Y'$ associated in a certain way with this rule template ($X' \supseteq X$, $Y' \supseteq Y$, $(X' \cup Y') \subseteq Z$) will also reach at least the same value of this measure. We will use this observation when defining a cover of a rule template.

The cover of rule template $X \xrightarrow{Z} Y$ is denoted by $TC(X \xrightarrow{Z} Y)$ and defined as follows:

$$TC(X \xrightarrow{Z} Y) = \{(X' \rightarrow Y') \in AR \mid X' \supseteq X, Y' \supseteq Y, (X' \cup Y') \subseteq Z\}.$$

Each association rule in $TC(X \xrightarrow{Z} Y)$ will be called *covered* by $X \xrightarrow{Z} Y$.

Proposition 6. Let $X \xrightarrow{Z} Y$ be an (\mathcal{M}, E) -strong rule template. Then, each association rule $(X' \rightarrow Y')$ in $TC(X \xrightarrow{Z} Y)$ is (\mathcal{M}, E) -strong.

Proof. Follows from definition of the cover of a rule template and Proposition 5. \square

Thus, all association rules covered by an (\mathcal{M}, E) -strong rule template are (\mathcal{M}, E) -strong, too. We will focus now on determining a set of association rules that are covered by a given rule template and on determining its cardinality.

Example 3. Let $\{ab\} \xrightarrow{\{abcde\}} \{e\}$ be a rule template. Then, $TC(\{ab\} \xrightarrow{\{abcde\}} \{e\}) = \{\{ab\} \rightarrow \{e\}, \{abc\} \rightarrow \{e\}, \{abd\} \rightarrow \{e\}, \{ab\} \rightarrow \{ce\}, \{ab\} \rightarrow \{de\}, \{abc\} \rightarrow \{de\}, \{abd\} \rightarrow \{ce\}, \{abcd\} \rightarrow \{e\}, \{ab\} \rightarrow \{cde\}\}.$ \square

We find that the number of association rules covered by a single rule template depends on the number of items in its base that occur neither in the antecedent of the rule template nor in its consequent (see Theorem 1).

Theorem 1. Let $X \xrightarrow{Z} Y$ be a rule template. Then:

$$|TC(X \xrightarrow{Z} Y)| = 3^m, \text{ where } m = |Z \setminus (X \cup Y)|.$$

Proof. By definition of TC , $X \rightarrow Y$ belongs to $TC(X \xrightarrow{Z} Y)$. In addition, $TC(X \xrightarrow{Z} Y)$ contains association rules having supersets of X as antecedents and supersets of Y as consequents provided additional items in antecedents and consequents come from $Z \setminus (X \cup Y)$. In fact, association rules belonging to $TC(X \xrightarrow{Z} Y)$ could be created from $X \rightarrow Y$ by extending its antecedent and consequent in the following way. Let V be a set of items from Z that occur neither in antecedent of $X \xrightarrow{Z} Y$ nor in its consequent; that is $V = Z \setminus (X \cup Y)$. A new rule to be included in $TC(X \xrightarrow{Z} Y)$ can be obtained from rule $X \rightarrow Y$ by performing one of the following three operations on each item v from V : (1) add v to the antecedent of $X \rightarrow Y$, (2) add v to the consequent of $X \rightarrow Y$, (3) do not use v when building a new rule. Thus, there are $3^{|V|} = 3^{|Z \setminus (X \cup Y)|}$ possible ways of building distinct association rules based on given rule template $X \xrightarrow{Z} Y$. \square

Example 4. Let $\{ab\} \xrightarrow{\{abcdefghijk\}} \{e\}$ be a rule template and $m = |\{abcdefghijk\} \setminus (\{ab\} \cup \{e\})| = 11 - 3 = 8$. So,

$|TC(\{ab\} \xrightarrow{\{abcdefghijk\}} \{e\})| = 3^8 = 6591$. In consequence, if $\{ab\} \xrightarrow{\{abcdefghijk\}} \{e\}$ is (\mathcal{M}, E) -strong, then 6591 association rules are (\mathcal{M}, E) -strong, too. \square

Theorem 2. Let $X \xrightarrow{Z} Y$ and $V \xrightarrow{U} W$ be distinct rule templates such that $V \supseteq X$, $W \supseteq Y$, and $U \subseteq Z$. Then, $TC(X \xrightarrow{Z} Y) \supset TC(V \xrightarrow{U} W)$.

Proof. Let $X \xrightarrow{Z} Y$ and $V \xrightarrow{U} W$ be distinct rule templates such that $V \supseteq X$, $W \supseteq Y$ and $U \subseteq Z$. $TC(X \xrightarrow{Z} Y) = \{(X' \rightarrow Y') \in AR \mid X' \supseteq X, Y' \supseteq Y, (X' \cup Y') \subseteq Z\} \supseteq \{(X' \rightarrow Y') \in AR \mid X' \supseteq V \supseteq X, Y' \supseteq W \supseteq Y, (X' \cup Y') \subseteq U \subseteq Z\} = TC(V \xrightarrow{U} W)$. Hence, we proved that $TC(X \xrightarrow{Z} Y) \supseteq TC(V \xrightarrow{U} W)$ (*). However, $X \xrightarrow{Z} Y$ and $V \xrightarrow{U} W$ are distinct rule templates, so $V \supset X$ or $W \supset Y$ or $U \supset Z$.

Case $V \supset X$: Then, association rule $X \rightarrow Y$ will belong to $TC(X \xrightarrow{Z} Y)$, but will not belong to $TC(V \xrightarrow{U} W)$ (since no rule in $TC(V \xrightarrow{U} W)$ will have antecedent X). Hence and by (*), $TC(X \xrightarrow{Z} Y) \supset TC(V \xrightarrow{U} W)$.

Case $W \supset Y$: Then, association rule $X \rightarrow Y$ will belong to $TC(X \xrightarrow{Z} Y)$, but will not belong to $TC(V \xrightarrow{U} W)$ (since no rule in $TC(V \xrightarrow{U} W)$ will have consequent Y). Hence and by (*), $TC(X \xrightarrow{Z} Y) \supset TC(V \xrightarrow{U} W)$.

Case $U \subseteq Z$: Then, association rule $X \rightarrow Z \setminus X$ will belong to $TC(X \xrightarrow{Z} Y)$, but will not belong to $TC(V \xrightarrow{U} W)$ (since no rule in $TC(V \xrightarrow{U} W)$ will have base Z). Hence and by (*), $TC(X \xrightarrow{Z} Y) \supset TC(V \xrightarrow{U} W)$. \square

4.2 Representative Rule Templates as a Lossless Representation of Association Rules

We denote (\mathcal{M}, E) -representative rule templates by $RRT_{(\mathcal{M}, E)}$ and define as follows:

$$RRT_{(\mathcal{M}, E)} = \{(X \xrightarrow{Z} Y) \in RT_{(\mathcal{M}, E)} \mid \neg \exists (V \xrightarrow{U} W) \in RT_{(\mathcal{M}, E)} \text{ such that } (X \xrightarrow{Z} Y) \neq (V \xrightarrow{U} W), V \subseteq X, W \subseteq Y, U \supseteq Z\}.$$

In Theorem 3, we claim that (\mathcal{M}, E) -representative rule templates cover all (\mathcal{M}, E) -strong association rules.

Theorem 3. $\bigcup_{(X \xrightarrow{Z} Y) \in RRT_{(\mathcal{M}, E)}} TC(X \xrightarrow{Z} Y) = AR_{(\mathcal{M}, E)}$.

Proof. We will prove Theorem 3 in two steps by showing that: (1) the union of covers of all (\mathcal{M}, E) -strong rule templates equals $AR_{(\mathcal{M}, E)}$; (2) the union of covers of all (\mathcal{M}, E) -representative rule templates equals the union of covers of all (\mathcal{M}, E) -strong rule templates, and by this, equals $AR_{(\mathcal{M}, E)}$.

- (1) We note that for each (\mathcal{M}, E) -strong association rule, say $X' \rightarrow Y'$, there is an (\mathcal{M}, E) -strong rule template; namely, $X' \overset{X' \cup Y'}{\rightarrow} Y'$ that covers $X' \rightarrow Y'$. On the other hand, for each (\mathcal{M}, E) -strong rule template $X \overset{Z}{\rightarrow} Y$, its cover $TC(X \overset{Z}{\rightarrow} Y) \subseteq AR_{(\mathcal{M}, E)}$ (by Proposition 6). Thus, $\bigcup_{(X \overset{Z}{\rightarrow} Y) \in RT_{(\mathcal{M}, E)}} TC(X \overset{Z}{\rightarrow} Y) = AR_{(\mathcal{M}, E)}$.
- (2) Let $RRT_{(\mathcal{M}, E)}$ do not contain an (\mathcal{M}, E) -strong rule template $X \overset{Z}{\rightarrow} Y$. By definition of $RRT_{(\mathcal{M}, E)}$, this implies that there is a non-empty set of rule templates $S = \{(V \overset{U}{\rightarrow} W) \in RT_{(\mathcal{M}, E)} \mid (V \overset{U}{\rightarrow} W) \neq (X \overset{Z}{\rightarrow} Y), V \subseteq X, W \subseteq Y, U \supseteq Z\}$. By Theorem 2, the cover of each rule template in S is a proper superset of $TC(X \overset{Z}{\rightarrow} Y)$. Let U_{\max} be the base of a rule template in S with the maximum number of items. Let S' contain all rule templates with base U_{\max} . Let V_{\min} be the antecedent of a rule template in S' with the minimum number of items. Let S'' contain all rule templates in S' with antecedent V_{\min} . Let W_{\min} be the consequent of a rule template in S'' with the minimum number of items. Let S''' contain all rule templates in S'' with consequent W_{\min} . By construction of S''' , each rule template in S''' is representative and its cover contains $TC(X \overset{Z}{\rightarrow} Y)$. In this way, for each non-representative rule template $(X \overset{Z}{\rightarrow} Y) \in RT_{(\mathcal{M}, E)}$, we will be able to identify at least one representative rule template $V \overset{U}{\rightarrow} W$ such that $TC(V \overset{U}{\rightarrow} W) \supset TC(X \overset{Z}{\rightarrow} Y)$. So, $\bigcup_{(X \overset{Z}{\rightarrow} Y) \in RRT_{(\mathcal{M}, E)}} TC(X \overset{Z}{\rightarrow} Y) = \bigcup_{(X \overset{Z}{\rightarrow} Y) \in RT_{(\mathcal{M}, E)}} TC(X \overset{Z}{\rightarrow} Y) = AR_{(\mathcal{M}, E)}$. \square

Theorem 4. Let $X' \rightarrow Y'$ be an association rule. There is an (\mathcal{M}, E) -representative rule template covering $X' \rightarrow Y'$ if and only if $X' \rightarrow Y'$ is an (\mathcal{M}, E) -strong association rule.

Proof. Let $X' \rightarrow Y'$ be an association rule.

(\Rightarrow) Follows from Proposition 6.

(\Leftarrow) By Theorem 3, each (\mathcal{M}, E) -strong association rule is covered by at least one (\mathcal{M}, E) -representative rule template. \square

5 Related Work on Lossless Representations of Association Rules

A number of concise lossless representations of strong association rules with respect to support and confidence have been proposed in the literature (see e.g. [7, 8] for an overview). Such representations typically consist of rules built from particular itemsets called *generators*; that is, itemsets the supports of which are less than the supports of all their proper subsets, and/or *closed itemsets*; that is, itemsets the supports of which are greater than the supports of all their proper supersets [2, 4, 6–8, 19, 20]. For example, in the case of representative association rules [6–8] and minimal non-redundant association rules [2], antecedents of rules are generators, whereas their consequents are set theoretical differences between closed itemsets and rules' antecedents. The set of representative rules is a subset of the set of minimal non-redundant association rules

[7, 8] and typically is by at least on order of magnitude less numerous. On the other hand, representative rules allow determination of pessimistic estimations of supports and confidences of strong association rules, while minimal non-redundant association rules allow exact determination of values of these measures. Yet, these and other developed representations of strong association rules with respect to support and confidence are not flexible enough to represent rules that are strong with respect to arbitrary ACBC/canonical measures.

In [11], we proposed the first representation of strong association rules with respect to any set of ACBC-measures. It consists of rule templates, where each rule template is a pair composed from a *lower rule* built from generators and an *upper rule* built from closed itemsets. For a rule template $(X \rightarrow Y, Z \rightarrow V)$, the following holds: $P(X) = P(Z)$, $P(Y) = P(V)$, $P(XY) = P(ZV)$. For each strong association rule $U \rightarrow W$, there is a rule template $(X \rightarrow Y, Z \rightarrow V)$ such that $X \subseteq U \subseteq Z, Y \subseteq W \subseteq V$, which implies that: $P(X) = P(U) = P(Z)$, $P(Y) = P(W) = P(V)$, $P(XY) = P(UW) = P(ZV)$ and that for any ACBC-measure μ : $\mu(X \rightarrow Y) = \mu(U \rightarrow W) = \mu(Z \rightarrow V)$.

In the current paper, we proposed a rule representation which does not require rule templates in the form of pairs of two rules. In addition, a representative rule templates we offered here, may cover strong association rules whose values of canonical measures may be the same or greater than values of respective canonical measures of covering rule templates. This should result in higher conciseness of the new representation. Beneath we formulate additional properties of representative rule templates (lack of space does not allow us to provide their proves).

Proposition 7.

- (a) Antecedents and consequents of representative rule templates are generators.
- (b) Bases of representative rule templates are closed itemsets.
- (c) The set of representative rule templates is never more numerous than the set of rule templates from [11].

6 Summary

In this paper, we offered the notion of a canonical measure and proved that a number of example measures definable (at most) in terms of the probabilities of rule antecedents, consequents and bases are, in fact, canonical. Then we proposed representative rule templates as a representation of association rules satisfying multiple evaluation criteria expressible in terms of canonical measures. We proved that this representation allows deriving all and only association rules strong with respect to a given set of criteria based on canonical measures. We proved a number of properties of rule templates. In particular, we derived the formula determining the number of association rules covered by a rule template. The found formula suggests that representative rule templates can be a very concise representation of strong association rules. By means of an example, we showed that a single rule template may represent thousands of not less strong association rules.

References

1. Agrawal, R., Imielinski, T., Swami, A.N.: Mining association rules between sets of items in large databases. In: ACM SIGMOD International Conference on Management of Data, pp. 207–216 (1993)
2. Bastide, Y., Pasquier, N., Taouil, R., Stumme, G., Lakhal, L.: Mining minimal non-redundant association rules using frequent closed itemsets. In: Lloyd, J., et al. (eds.) CL 2000. LNCS (LNAI), vol. 1861, pp. 972–986. Springer, Heidelberg (2000). https://doi.org/10.1007/3-540-44957-4_65
3. Brin, S., Motwani, R., Ullman, J.D., Tsur, S.: Dynamic itemset counting and implication rules for market basket data. In: ACM SIGMOD 1997 International Conference on Management of Data, pp. 255–264 (1997)
4. Hamrouni, T., Yahia, S.B., Nguifo, E.M.: Succinct minimal generators: theoretical foundations and applications. *Int. J. Found. Comput. Sci.* **19**(2), 271–296 (2008)
5. Hilderman, R.J., Hamilton, H.J.: Evaluation of interestingness measures for ranking discovered knowledge. In: Cheung, D., Williams, G.J., Li, Q. (eds.) PAKDD 2001. LNCS (LNAI), vol. 2035, pp. 247–259. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45357-1_28
6. Kryszkiewicz, M.: Closed set based discovery of representative association rules. In: Hoffmann, F., Hand, D.J., Adams, N., Fisher, D., Guimaraes, G. (eds.) IDA 2001. LNCS, vol. 2189, pp. 350–359. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44816-0_35
7. Kryszkiewicz, M.: Concise representations of frequent patterns and association rules. *Prace Naukowe Politechniki Warszawskiej. Elektronika* **142**, 5–207 (2002)
8. Kryszkiewicz, M.: Concise representations of association rules. In: Hand, D.J., Adams, N.M., Bolton, R.J. (eds.) Pattern Detection and Discovery. LNCS (LNAI), vol. 2447, pp. 92–109. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45728-3_8
9. Kryszkiewicz, M.: Dependence factor for association rules. In: Nguyen, N.T., Trawiński, B., Kosala, R. (eds.) ACIIDS 2015. LNCS (LNAI), vol. 9012, pp. 135–145. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-15705-4_14
10. Kryszkiewicz, M.: Dependence factor as a rule evaluation measure. In: Matwin, S., Mielniczuk, J. (eds.) Challenges in Computational Statistics and Data Mining. SCI, vol. 605, pp. 205–223. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-18781-5_12
11. Kryszkiewicz, M.: A lossless representation for association rules satisfying multiple evaluation criteria. In: Nguyen, N.T., Trawiński, B., Fujita, H., Hong, T.-P. (eds.) ACIIDS 2016. LNCS (LNAI), vol. 9622, pp. 147–158. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49390-8_14
12. Kryszkiewicz, M.: ACBC-adequate association and decision rules versus key generators and rough sets approximations. *Fundam. Inform.* **148**(1–2), 65–85 (2016)
13. Lavrač, N., Flach, P., Zupan, B.: Rule evaluation measures: a unifying view. In: Džeroski, S., Flach, P. (eds.) ILP 1999. LNCS (LNAI), vol. 1634, pp. 174–185. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48751-4_17
14. Lenca, P., Meyer, P., Vaillant, B., Lallich, S.: On selecting interestingness measures for association rules: user oriented description and multiple criteria decision aid. *Eur. J. Oper. Res.* **184**, 610–626 (2008). Elsevier
15. Piatetsky-Shapiro, G.: Discovery, analysis, and presentation of strong rules. In: Knowledge Discovery in Databases, pp. 229–248, AAAI/MIT Press (1991)
16. Sheikh, L.M., Tanveer, B., Hamdani, S.M.A.: Interesting measures for mining association rules. In: Proceedings of INMIC 2004. IEEE (2004)

17. Shortliffe, E., Buchanan, B.: A model of inexact reasoning in medicine. *Math. Biosci.* **23**, 351–379 (1975)
18. Suzuki, E.: Interestingness measures - limits, desiderata, and recent results. In: Lenca, P., Lallich, S. (eds.) *QIMIE/PAKDD 2009* (2009)
19. Stumme, G., Taouil, R., Bastide, Y., Pasquier, N., Lakhal, L.: Intelligent structuring and reducing of association rules with formal concept analysis. In: Baader, F., Brewka, G., Eiter, T. (eds.) *KI 2001. LNCS (LNAI)*, vol. 2174, pp. 335–350. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45422-5_24
20. Zaki, M.J.: Generating non-redundant association rules. In: 6th ACM SIGKDD (2000)