*International Journal of Intelligence Science*, **2013, 3, 24-33**
doi:10.4236/ijis.2013.31A004 Published Online March 2013 (http://www.scirp.org/journal/ijis)

# Personalized Action Rules for Side Effects Object Grouping

**Hakim Touati[1], Jieyan Kuang[1], Ayman Hajja[1], Zbigniew W. Ras[1,2]**
[1]College of Computing and Informatics, University of North Carolina, Charlotte, USA
[2]Institute of Computer Science, Warsaw University of Technology, Warsaw, Poland
Email: htouati@uncc.edu, jkuang1@uncc.edu, ahajja@uncc.edu, ras@uncc.edu

Received November 8, 2012; revised January 2, 2013; accepted January 14, 2013

## ABSTRACT

There have been multiple techniques to discover action-rules, but the problem of triggering those rules was left exclusively to domain knowledge and domain experts. When meta-actions are applied on objects to trigger a specific rule, they might as well trigger transitions outside of the target action rule scope. Those additional transitions are called side effects, which could be positive or negative. Negative side effects could be devastating in some domains such as healthcare. In this paper, we strive to reduce those negative side effects by extracting personalized action rules. We proposed three object-grouping schemes with regards to same negative side effects to extract personalized action rules for each object group. We also studied the tinnitus handicap inventory data to apply and compare the three grouping schemes.

**Keywords:** Meta-Actions; Action-Rules; Personalization; Optimization; Grouping

## 1. Introduction

Action rules observe patterns, recorded on an information system, of domain experts applying their domain knowledge and expertise in real world situations. They provide efficient solutions to help naïve system users solve real world problems. There has been an increasing interest on action rule discovery algorithms since their creation by Ras and Wieczorkowska in [1]. Action rules have been used in healthcare to understand experts' practices and improve patients' care [2-5]. They are also used in distribution and customer loyalty systems, and maybe used in a wide range of industries such as education, and banking.

Action rules, that were first introduced in [1] and then investigated in [6-12], represent changing some of the objects' properties that will make the overall objects state change. They model the correlation between some specific classification features values and the decision feature values.

Meta-actions are used to provoke changes in objects' state and trigger action rules to solve specific problem. They are mainly defined by domain knowledge and domain experts and used by system users. When used to trigger actions rule, meta-actions trigger changes in objects state within the system to execute the action rule. In other words, they provoke changes in objects' features that will not only trigger the actions rule targeted but also side effects that could be negative. The negative side effects can damage the objects' features outside of the executed action rule scope. Naïve system users might not know about those negative side effects; thus, they will not be taken into consideration even though they might be harmful.

In this paper, we study closely the side effects of applying meta-action. We acknowledge that those negative side effects are not avoidable in most situations; therefore, we strive to personalize the action rules and their respective meta-actions applied to objects based on their reactions to meta-actions. We strongly believe that action rules should be extracted from data sets describing objects that have the same negative side effects, and we present three objects' grouping techniques based on negative side effects.

This paper was motivated by the tinnitus handicap dataset that was exploited in previous research [5] to extract action rules. In this dataset, patients are treated using four different treatments (in this context treatments represent meta-actions); however, some of the patient's features were changed to negative values (worse property values). Those negative changes affect objects' properties that are outside of the action rules scope and are therefore omitted by system users.

By analyzing the patient's negative side effects, and grouping patients based on their reactions to treatment, we can extract personalized action rules. The main contributions of this paper are:

1) Defining side effects resulting from applying meta-actions;

2) Presenting three object grouping methods for personalized action rules based on negative side effects;

3) Implementing the three grouping techniques and experimenting them on the tinnitus handicap dataset.

There are a number of software packages available for discovering action rules. For instance, Action4ft-Miner module of the Lisp-Miner project developed by Jan Rauch's group discovers action rules under different constraints which can be placed for the antecedent part of the rule [13].

## 2. Problem Definition

An action rule provides a set of atomic actions on its antecedent side, which will trigger the atomic action on its right side, if executed. Meta-actions are the triggers for those atomic actions to happen. However, the current solution does not provide a personalized procedure for specific objects or group of objects to whom applying certain meta-actions may result in negative side effects. For example, given a bank customer that is 24 years old, has medium salary, medium monthly expanses, high savings, low interest rate, and average loan profitability, if we apply meta-actions to increase the interest rate that triggers an action rule increasing the loan profitability, we may as well trigger a decrease in customer's savings, thus affecting negatively the saving account profitability. This scenario may not be suitable for the bank decision maker, and may not respect the strategy of the bank.

We strive to extract personalized action rules with regards to objects negative reactions to meta-actions. To achieve this goal, we group objects based on their negative reactions and extract personalized action rules on each object group concerned by specific negative side effects.

## 3. Side Effects Based Personalization

In this section we explore the different techniques of grouping objects and extracting action rules from a decision system, introduced by *Z.* Pawlak [14], that are personalized for each group of objects.

By a decision system we mean $S=\left(X, F\cup\{d\}, V\right)$, where:

(1) $X$ is a set of objects;

(2) $F$ is a set of classification features, $f: X\to V_f$ is a function for any $f\in F$, where $V_f$ is called the domain of $f$;

(3) $d$ is a decision feature, $d: X\to V_d$ is a function, where $V_d$ is called the domain of $d$;

(4) $V=V_F\cup V_d$, where $V_F=\bigcup\left\{V_f : f\in F\right\}$.

Also, for each $x\in X$ and $f\in F$, we assume that value $f(x)\in V_f$ is classified either as positive (normal) or negative (abnormal). To be more precise, we assume that $F(x)$ denotes the set $\left\{f(x): f\in F\right\}$, and that $F(x)=E_n(x)\bigcup E_p(x)$, where $E_p(x)$ is a set of positive values and $E_n(x)$ is a set of negative values for $x\in X$. If $f(x)\in E_n(x)$, then the value $F(x)$ is interpreted as abnormal (for instance: high temperature, cough, headache, $\cdots$). If $f(x)\in E_p(x)$, then value $F(x)$ is interpreted as normal.

### 3.1. Action-Rules

Action rules are rules that provide a set of actions to follow to drive the objects population from a certain state to a more profitable state. In addition, action rules are composed of features that are divided into two sets: stable features $F_{st}$, and flexible features $F_{fl}$ such that $F=F_{fl}\cup F_{st}$.

Stable features are object properties that we do no have control over in the context of our information system. For example, age and gender are stable features. Flexible features are object properties that can transition from an initial value to another value triggering a change in the object state. For instance, salary and benefits are flexible features since they can change values.

An atomic action term is an expression that defines a change of state for a distinct feature. For example, $\left(f, v_1\to v_2\right)$ is an atomic action term which defines a change of value of the attribute $f$ from $v_1$ to $v_2$, where $v_1, v_2\in V_f$. In the case when there is no change, we omit the right arrow sign, so for example, $\left(f, v_1\right)$ means that the value of attribute $f$ remains $v_1$, where $v_1\in V_f$.

Action terms are defined as the smallest collection of expressions such that:

- If $t$ is an atomic action term, then $t$ is an action term;
- If $t_1, t_2$ are action terms and $\wedge$ is a 2-argument functor called composition, then $t_1\wedge t_2$ is a candidate action term;
- If $t$ is a candidate action term and for any two atomic action terms $\left(f, v_1\to v_2\right), \left(g, w_1\to w_2\right)$ contained in $t$ we have $f\neq g$, then $t$ is an action term.

The domain $\mathrm{Dom}(t)$ of an action term $t$ is the set of features listed in the atomic action terms contained in $t$. For example, $t=\left[\left(f, v_1\to v_2\right)\wedge\left(g, w_1\right)\right]$ is an action term that consists of two atomic action terms, namely $\left(f, v_1\to v_2\right)$ and $\left(g, w_1\right)$. Therefore, $\mathrm{Dom}(t)=\left\{f, g\right\}$.

Action rules are expressions that take the following form: $r=\left[t_1\Rightarrow t_2\right]$, where $t_1, t_2$ are action terms. The interpretation of the action rule $r$ is that by triggering the action term $t_1$, we would get, as a result, the changes of states in action term $t_2$. We also assume that $\mathrm{Dom}(t_1)\cup\mathrm{Dom}(t_2)\subseteq F$, and $\mathrm{Dom}(t_1)\cap\mathrm{Dom}(t_2)=\varnothing$.

For example $r=\left[\left[\left(f, v_1\to v_2\right)\wedge\left(g, w_2\right)\right]\Rightarrow\left(d, d_1\to d_2\right)\right]$

means that by changing the state of feature $f$ from $v_1$ to $v_2$, and by keeping the state of feature $g$ as $w_2$, we would observe a change in attribute $d$ from the state $d_1$ to $d_2$, where $d$ is commonly referred to as the decision attribute.

In [9], it was observed that each action rule can be seen as a composition of two classification rules. For instance, the rule $r = \left[\left[\left(f, v_1 \to v_2\right) \wedge \left(g, w_2\right)\right] \Rightarrow \left(d, d_1 \to d_2\right)\right]$ is a composition of $r_1 = \left[\left(f, v_1\right) \wedge \left(g, w_2\right)\right] \to \left(d, d_1\right)$ and $r_2 = \left[\left(f, v_2\right) \wedge \left(g, w_2\right)\right] \Rightarrow \left(d, d_2\right)$. This fact can be recorded by the equation $r = r\left(r_1, r_2\right)$. Also, the definition of support $(\text{Sup})$ and confidence $(\text{Conf})$ of an action rule is based on support and confidence of classification rules (see below).

Assume that action rule $r$ is a composition of two classification rules $r_1$ and $r_2$. Then [9]:

$$\text{Sup}(r) = \min\left\{\text{card}\left(\sup\left(r_1\right)\right), \text{card}\left(\sup\left(r_2\right)\right)\right\},$$

$$\text{Conf}(r) = \text{conf}\left(r_1\right) \times \text{conf}\left(r_2\right).$$

By the support of a classification rule

$$r = \left[\left[\left(f_1, f_{11}\right) \wedge \left(f_2, f_{21}\right) \wedge \left(f_3, f_{31}\right) \wedge \cdots \wedge \left(f_k, f_{k1}\right)\right] \to \left(d, d_1\right)\right]$$

in a decision system $S = \left(X, F \cup \{d\}, V\right)$, where $\left(\forall i \leq k\right)\left(f_i \in F \;\&\; f_{i1} \in V\right), d_1 \in V_d$, we mean the set $\sup(r) = \left\{x \in X : \left(\forall i \leq k\right)\left[f_i(x) = f_{i1}\right] \& \, d(x) = d_1\right\}$.

## 3.2. Meta-Actions

By meta-actions associated with decision system $S$ we mean higher concepts used to model certain generalizations of actions rules [11]. Meta-actions, when executed, trigger changes in values of some flexible features in $S$ as described by influence matrix [11] and atomic action terms.

To give an example, let us assume that classification features in $S$ describe teaching evaluations at some school and the decision feature represents their overall score. Explain difficult concepts effectively, Speaks English fluently, Stimulate student interest in the course, Provide sufficient feedback are examples of classification features. Then, examples of meta-actions associated with $S$ will be: Change the content of the course, Change the textbook of the course, post all material on the Web. Clearly, any of these three meta-actions will not influence the feature Speaks English fluently and the same its values will remain unchanged [11]. Let us take Hepatitis as the application domain. Then *increase blood cell plague* and *decrease level of alkaline phosphatase* are examples of an atomic action term. Drugs like *Hepatil* or *Hepargen* are seen as meta-actions triggering changes described by these two atomic action terms [4, 15].

It should be noted that *Hepatil* is also used to get rid of obstruction, eructation, and bleeding. However, *Hepargen* is not used to get rid of obstruction but it is used to get rid of eructation and bleeding.

Also, it should be mentioned here that an expert knowledge concerning meta-actions involves only classification features. Now, if some of these features are correlated with the decision feature, then the change of their values will cascade to the decision through the correlation. The goal of action rule discovery is to identify possibly all such correlations.

Consider several meta-actions, denoted $M_1, M_2, \cdots, M_n$. Each one can invoke changes within values of some classification features in $F = \left\{f_1, f_2, \cdots, f_m\right\}$. The expected changes of values of classification features on objects from $S$ triggered by these meta-actions are described by the influence matrix $\left\{E_{ij} : 1 \leq i \leq n \,\&\, 1 \leq j \leq m\right\}$. **Table 1** gives an example of an influence matrix associated with 6 meta-actions and three features: $a$, $b$, and $c$.

For instance, let us take meta-action $M_2$. It says that by executing $M_2$ on objects in $S$, two atomic action terms are triggered. They are: $\left(a, a_2 \to a_1\right)$ and $\left(b, b_2 \to b_2\right)$. It means that objects in $S$ satisfying the description $\left(a, a_2\right) \wedge \left(b, b_2\right)$ are expected to change their description to $\left(a, a_1\right) \wedge \left(b, b_2\right)$.

Let us define $\boldsymbol{M}(S)$ as a set of meta-actions associated with a decision system $S$. Let $f \in F, x \in X$, *and* $M \subset \boldsymbol{M}(S)$, then, applying the meta-actions in the set $M$ on an object $x$ will result in $M\left(f(x)\right) = f(y)$, where object $x$ is converted to object $y$ by applying all meta-actions in $M$ to $x$. Similarly, $M\left(F(x)\right) = F(y)$, where $x$ is converted to $y$ by applying all meta-actions in $M$ to $x$ for all $f \in F$. Also, by $F(Y)$, where $Y \subseteq X$, we mean $\left\{F(x) : x \in Y\right\}$.

## 3.3. Side Effects

The main goal of meta-actions is to trigger action rules.

**Table 1. An example of an influence matrix associated with 6 meta-actions and three features: *a*, *b*, and *c*.**

| | $a$ | $b$ | $c$ |
|---|---|---|---|
| $M_1$ | - | $b_1$ | $c_2 \to c_1$ |
| $M_2$ | $a_2 \to a_1$ | $b_2$ | - |
| $M_3$ | $a_1 \to a_2$ | - | $c_2 \to c_1$ |
| $M_4$ | - | $b_1$ | $c_1 \to c_2$ |
| $M_5$ | - | - | $c_1 \to c_2$ |
| $M_6$ | $a_1 \to a_2$ | - | $c_1 \to c_2$ |

However, it is often the case that when applying meta-actions for the purpose of executing a specific action rule, a set of unrelated additional and potentially harmful atomic action terms are triggered. The additional action terms resulting from the meta-action application are called side effects.

Meta-actions might move some objects' features values from negative to positive values $f(x) \in E_n(x)$ and $f(y) \in E_p(y)$ (desirable positive side effects), and some object's features values from positive to negative values $f(x) \in E_p(x)$ and $f(y) \in E_n(y)$ (undesirable negative side effects).

Even though the features transitioning from positive to negative values might result in catastrophic situations, they were not fully investigated in previous work involving action rules discovery.

## 4. Personalized Object Grouping

Unfortunately, the negative side effects resulting from applying meta-actions are unavoidable in most situations. However, we can still lower the negative side effects resulting from executing the meta-actions by personalizing the action rules applied to objects. Action rules dictate the sets of meta-actions to apply to be triggered. There are multiple subsets of meta-actions that could be applied to different action rules and result in multiple subsets of negative side effects.

We aim to minimize the negative side effects for a large number of objects by discovering personalized action rules while keeping their utility and increasing their support and confidence.

In the following, we define three techniques to group objects based on their side effects resulting from applying meta-actions for a personalized action rules discovery system.

### 4.1. Side Effects Based Grouping

In this technique, as the title suggests, we group objects based on side effects. In most situations, objects have known side effects such as patients having allergies. However, more side effects can be determined based on the possible meta-actions applied.

Let us define the set of meta actions $\boldsymbol{M}(S)$ on the system $S = (X, F \cup \{d\}, V)$, such that

$F(x) = \{f(x) : f \in F\} = E_n(x) \cup E_p(x)$ for $x \in X$.

We aim at grouping objects $x \in X$ that have same negative side effects for any meta-action in $M = \{M_k\}_{k \in K} \subseteq \boldsymbol{M}(S)$ *where* $K = \{1, \cdots, |M|\}$. This grouping will result in a partition of $X$ defined by the equivalence relation given in the following:

$$x_i \approx_M x_j \text{ iff } (\forall k \in K)\left[E_{n,k}(y_i) = E_{n,k}(y_j)\right]$$

where $M_k(F(x_i)) = F(y_{i,k}), M_k(F(x_j)) = F(y_{j,k})$, and $x_i, x_j \in X$.

Also, we assume here that $x_i$ is converted to $y_{i,k}$ and $x_j$ is converted to $y_{j,k}$ by $M_k \in M$.

In a real setting, as explained earlier, we need to follow a set of steps in order to group objects in the most optimal way minimizing the negative side effects to extract the right action rules. In fact, meta-actions result in different negative side effects from one object to another. We defined the following steps for personalizing action rules based on common negative side effects:

- We first need to extract the negative side effects from applying the meta-actions for each object $x \in X$ if they are not defined yet. This process is performed by analyzing our decision system $S$ and extracting all negative side effects $E_n(x)$ for all $x \in X$ for each transaction (observation) that happens directly after applying the meta-actions.
- We then group the objects that have the same side effects using the previously defined equivalence relation that will result in a partition $G$. Those groups encompass the objects' observations in our decision system for each object in the group.
- We extract the personalized action rules for each group $g \in G$ using the objects' observation in each group and the algorithms presented in [3,9].
- Finally, we select the personalized action rules with the best support and confidence pair to apply in each group.

The personalized action rules extracted for each group of objects will result in the same negative side effects when applying the related meta-actions to trigger the action rules. However, grouping the objects by negative side effects first will decrease the object population for discovering the action rules for each group. This might result in decreasing the support of the rules or even not discovering all possible rules. To remediate to this problem, we propose the action rule based grouping that is described in the following.

### 4.2. Action Rules Based Grouping

Another way of grouping objects in $X$ is by grouping action rules with respect to their common support.

Let us assume that $AR_S$ is a set of action rules extracted from a decision system $S = (X, F \cup \{d\}, V)$, and $r_1 = r_1(r_{11}, r_{12}), r_2 = r_2(r_{21}, r_{22}) \in AR_S$. The binary relation $\equiv_S \subseteq AR_S \times AR_S$ is defined as follows:

$$r_1 \equiv_S r_2 \text{ iff } \left[\sup(r_{11}) = \sup(r_{21})\right].$$

Clearly, $\equiv_S$ is an equivalence relation which partitions the action rules in $AR_S$ into classes such that any two action rules in the same class have the same set of supporting objects in $X$. For instance, let us assume that

$r_1 = r_1(r_{11}, r_{12}) \in A \in R_S, r_{11} = [t_{11} \to t_{12}]$, and $\sup(r_{11}) = Y_1$. Also, we assume here that $r_2 = r_2(r_{21}, r_{22}) \in AR_S, r_{21} = [t_{21} \to t_{22}], \sup(r_{21}) = Y_2$, and $r_1 \equiv_S r_2$. Then, $\mathrm{Dom}(t_{11}) = \mathrm{Dom}(t_{21})$ and $F(Y_1)/\mathrm{Dom}(t_{11}) = F(Y_2)/\mathrm{Dom}(t_{21})$, where $F(Y_1)/\mathrm{Dom}(t_{11})$ is the set of values of attributes listed in $t_{11}$. Now, if $t_{11} = (f_1, v_{11}) \wedge (f_2, v_{21}) \wedge \cdots \wedge (f_k, v_{k1})$, then

$$F(Y_1)/\{f_1, f_2, \cdots, f_k\} = \{(f_1, v_{11}), (f_2, v_{21}), \cdots, (f_k, v_{k1})\}$$

displays all properties which objects in $X$ have to satisfy in order to be affected by $r_1$ or $r_2$.

So, the relation $\equiv_S \subseteq AR_S \times AR_S$ partitions the action rules in $AR_S$ into equivalence classes in such a way that each class $[r_1]_{\equiv S}$, where $r_1 = r_1(r_{11}, r_{12})$ and $r_{11} = [t_{11} \to t_{12}]$, has a unique set of attribute values $I([r_1]_{\equiv S})$ for $\mathrm{Dom}(t_{11})$ which is used as its identifier. In the section above, $I([r_1]_{\equiv S}) = \{(f_1, v_{11}), (f_2, v_{21}), \cdots, (f_k, v_{k1})\}$.Each identifier defines a subset of all objects in $X$ satisfying all properties listed in it. This way, we do not only group the objects in $X$ but also identify the largest subset of objects in $X$ which can be affected by minimum one action rule.

We still need to partition further the obtained groups of objects by taking into consideration personalized action rules based on their negative side effects. We use grouping mechanism similar to the side effects based grouping, presented in the previous section.

We define the following steps for personalizing action rules based grouping with respect to the negative side effects:

- We first extract the set of action rules $AR_S$ from $S$ and next we group the objects in $X$ based on the equivalence relation $\equiv_S \subseteq AR_S \times AR_S$. Let's assume that $G = \{G_k\}_{k \in K}$ represents that grouping.
- We extract the negative side effect $E_n(x)$ for all $x \in X$ resulting from applying the meta-actions to trigger the action rules in $AR_S(x)$ for each object $x \in X$.
- Then, we split each group $G_k \in G$ in such a way that objects having the same side effects with respect to action rules associated with $G_k$ are placed in the same sub-groups. For that purpose, we use the equivalence relation $\approx_M$ defined in the previous section. It will result in new sub-groupings $G_{ki}$ of $X$.
- We merge the sub-groups $G_{ki}$ if their respective action rules trigger the same negative side effects.
- Finally, we select the personalized action rules with the best support and confidence pair to apply in each group.

This grouping will not generate smaller groups than the previous technique since the merging step insures that groups with the same negative side effects are merged together.

### 4.3. Meta-Actions Based Grouping

The previous grouping technique based on action rules does not take the meta-actions into consideration. This may result in groups of objects with their respective action rules having different meta-actions applied to trigger their rules.

Another grouping technique groups objects first with respect to meta-actions applied to them. Of course meta-actions are not applied randomly to objects. They are either applied based on an action rule needs, or applied by an administrator making a decision based on his/her expertise.

Let us have a decision system $S = (X, F \cup \{d\}, V)$ and a set of meta-actions $\boldsymbol{M}(S)$ associated with $S$ which can be applied to objects in $X$. We first group objects by meta-actions and then we split the obtained groups further with respect to the negative side effects which are triggered by them. In order to group objects in $X$ with respect to a meta-action $M_p = \{E_{pk} : k = 1, 2, \cdots, m\} \in \boldsymbol{M}(S)$, where $E_{pk}$ are atomic actions triggered by $M_p$, we assume that $\mathrm{Dom}(M_p) = \{\mathrm{Dom}(E_{pk}) : k = 1, 2, \cdots, m\}$ and define the following relation for any $x_i, x_j \in X$:

$$x_i \approx_{Mp} x_j \text{iff}$$

$$\left[F(x_i) = F(x_j) \,\&\, M_p(F(x_i)) = M_p(F(x_j))\right]$$

In a similar way we can group objects in $X$ with respect to any set of meta-actions and in particular with respect to a minimal set of meta-actions $\{M_p\}_{p \in P}$ triggering a given action rule $r(r_1, r_2)$. Let us assume that $G = \{G_k\}_{k \in K}$ represents that grouping. Next, we split each group $G_k \in G$ in such a way that objects having the same side effects with respect to meta-actions associated with $G_k$ are placed in the same sub-groups. This strategy is similar to the one presented in Section 4.1.

## 5. Experiments

In this paper, we used R 2.15 to process the data and built in software to discover action rules.

### 5.1. Data-Set Description

The dataset used in those experiments is from the Tinnitus Handicap Inventory and represents physician's observations on patients. The data contains three categories of observations on patient's properties that are affected by tinnitus, which are: functioning (*F*), emotions (*E*), and how catastrophic it is (*C*). Each category consists of several related questions describing the patient's state. There are 25 multiple choice questions altogether, and the answers to all of them can be mapped to numeric scores: "Yes" is 4, "Sometimes" is 2 and "No" is 0. To evaluate the overall status for each patient, physicians observed three features "Sc *F*", "Sc *E*", and "Sc *C*", which are the total score of functioning category, the total score of emotions category and the total score of catastrophic category, respectively. Those three scores represent the sum of all the answer's scores for each category. Then feature "Sc *T*" (total score) is generated by adding results of "Sc *F*", "Sc *E*", and "Sc *C*" together to measure the tinnitus severity. The Tinnitus Handicap Inventory is completed during each patient visit and stored with patients' ID, visit date and number, and patient's gender (g). Another aspect of the data was the treatment performed on the patients at each visit. The treatment performed on the patients at each visit were divided into four treatments that are: Hearing Aid (HA), Sound Generator (SG), (CO) Combination of HA and SG, and a regular consultation (RC).

To be able to use the data in our experiments we had to perform a cleaning step along with a discretization step. In fact, the total number of patients' visits is of 2591 visit instances; however, there are multiple missing values and incomplete visit instances that had to be removed in order to be able to complete our experiments with the cleanest data possible. After cleaning the data, we ended up having only 517 visit instances.

We distinguished our data set classification features from the side effects. We assumed that the classification features are the functioning (*F*), emotions (*E*), and the catastrophic (*C*) features, and we kept their score values as they were already discretized. We further assumed that the side effects were the three scores of each category of the features Sc *F*, Sc *E*, and Sc *C*. We discretized the side effects based on the improvements on the category score (score = 1 positive side effects if the score decreases) and the declining of the category score (score = 0 negative side effects if the score increases). The decision feature is the total score, and the main goal of the treatments was to decrease the total score. We also discretized the decision feature, the total score, based on its improvement and its declining (score = 1 if improvement *i.e*. score decreases) and (score = 0 if declining *i.e*. score increases).

## 5.2. Side Effects Based Grouping

In this experiment we used the steps described in our proposed approach. However, we first cleaned the data and organized it by negative side effects.

Since we already know the three side effects in our dataset, we grouped patients that have the same negative side effects by codifying the different combinations of side effects values. Since we have three side effects and two possible values for each side effect (0 for negative and 1 for positive), we will have eight $\left(2^3 = 8\right)$ possible groups of patients $\left(000, 001, 010, \cdots, 111\right)$ in our partitioning.

Grouping patients with regards to same side effects resulted in the following number of patients in each group (see **Table 2**):

After grouping the patients by side effects, we extracted the action rules AR (2, 85%) using our action rules discovery software. We constrained the action rules confidence to 85% and support to a minimum of 2 in each one of the groups. However, since different groups might have different support, we increased the support sequentially to have the highest minimum support that returns actions rules. Furthermore, we fixed the decision transition from no-improvement to improvement $\left(\text{score}:\left(0 \to 1\right)\right)$.

The results of each group's action rules discovery after the partitioning are presented in the following **Table 3**.

Note that the groups of only negative side effects $g_{000}$ and the group of only positive side effects $g_{111}$ do not have any action rule. This is due to the decision feature being the same for all the group. In addition, note that the extracted rules do not have strong support. This is due to shrinking the patients populations in the groups.

An example of an action rule extracted from group $g_{110}$ is described in the following:

$$\left(F.24, 4 \to 4\right) \wedge \left(C.8, 4 \to 4\right) \wedge \left(F.7, 4 \to 2\right)$$
$$\wedge \left(C.19, 4 \to 4\right) \wedge \left(E.10, 4 \to 2\right)$$
$$\Rightarrow \left(ScT, 0 \to 1\right), \text{sup} = 3, \text{conf} = 100\%$$

**Table 2. Number of patients by side effects groups.**

| Group | $g_{000}$ | $g_{001}$ | $g_{010}$ | $g_{011}$ | $g_{100}$ | $g_{101}$ | $g_{110}$ | $g_{111}$ |
|---|---|---|---|---|---|---|---|---|
| #Patients | 46 | 38 | 17 | 54 | 18 | 44 | 32 | 268 |

**Table 3. Action rules by side effects groups.**

| Groups | Support | Confidence | #Rules |
|---|---|---|---|
| $g_{000}$ | - | - | 0 |
| $g_{001}$ | 2 | 100 | 251 |
| $g_{010}$ | 2 | 100 | 8128 |
| $g_{011}$ | 5 | 100 | 3 |
| | 5 | 86 | 1 |
| $g_{100}$ | 2 | 100 | 152213 |
| $g_{101}$ | 3 | 100 | 720 |
| $g_{110}$ | 3 | 100 | 5 |
| $g_{111}$ | - | - | 0 |

 

### 5.3. Action Rules Based Grouping

After the data cleaning step, we extracted all possible action rules from the entire decision system. We used our action rules extraction software setting up the minimal confidence to 85%, and the starting support at a minimum of 20. This support was then decreased sequentially until we reached a minimum support that resulted in discovering at least one action rule.

First, we extracted two action rules with minimum support 20 and minimum confidence 85%. Then, we decreased the minimum support to 19 to extract more action rules while keeping the confidence at least at 85%. This way 10 additional action rules has been found.

Next, we grouped them into sets of action rules that have the same antecedent side. Grouping the action rule with same antecedent side resulted in two groups for the first partition P1 for action rules with support 20. For the second set of action rules with support 19, we ended up having a partition P2 of 10 groups. Each one of those groups is summarized in **Table 4** and contains one action rule with a specific confidence and support:

For each action rule antecedent side set, we grouped patients that have the same preconditions as the antecedent part of action rules set together in the same group. This type of partitioning is natural since each patient $x \in X$ is associated with a set of possible action rules $AR_S(x)$. This step also insures that patients that do not have possible action rules $AR_S(x) = \varnothing$ are not part of the partition grouping.

Each group of action rules led to a number of patients having the same preconditions as the antecedent side of the rules. This experiment returned the same number of patients in each group as the support of the respective rules, which are 20 patients for the two groups of P1 and 19 patients for each group of P2.

**Table 4. Action rules by antecedent side grouping.**

| Partitions | Groups | Support | Confidence | #Rules |
|---|---|---|---|---|
| P1 | G1 | 20 | 86.95 | 1 |
| | G2 | 20 | 87.41 | 1 |
| P2 | G1 | 19 | 86.36 | 1 |
| | G2 | 19 | 85.58 | 1 |
| | G3 | 19 | 87.73 | 1 |
| | G4 | 19 | 87.55 | 1 |
| | G5 | 19 | 86.36 | 1 |
| | G6 | 19 | 86.70 | 1 |
| | G7 | 19 | 86.85 | 1 |
| | G8 | 19 | 86.36 | 1 |
| | G9 | 19 | 86.99 | 1 |
| | G10 | 19 | 86.85 | 1 |

Note that we do not get the same group size as the support of the corresponding rules in all situations since we are using the minimum support method to compute the rule support.

We followed the next step in our described approach where we further partitioned each group of patients presented in the previous **Table 4** to subgroups having the same negative side effects.

Each action rules based group was partitioned into a number of subgroups with regards to the same negative side effects. This partitioning is represented in **Table 5**, where you can note that the total number of patients for all negative side effects sub-group (row sum) is larger than the total number of patients in each corresponding parent group in the action rules grouping. This is due to an overlap between the groups for patients having different applicable action rules.

The previous table also represents the merging step, where the total number of patients in each sub-group is represented by the partition with respect to side effects. Note that this number is small due to the patients overlap described earlier in the table for the action rules groups. However there is no overlap between the different side effects groupings.

### 5.4. Meta-Action Based Grouping

This experiment requires grouping the objects based on meta-actions. Since the physician already applied treatments (meta-actions) to patients and those treatments are recorded in the dataset, we just need to place any two patients in the same group if the same treatments have been applied to them.

**Table 5. Number of patients by side effects and action rules grouping.**

| Partition | Group | $g_{000}$ | $g_{001}$ | $g_{010}$ | $g_{011}$ | $g_{100}$ | $g_{101}$ | $g_{110}$ | $g_{111}$ |
|---|---|---|---|---|---|---|---|---|---|
| P1 | G1 | 8 | 2 | 2 | 4 | 1 | 1 | 1 | 1 |
| | G2 | 8 | 6 | 3 | 2 | 0 | 1 | 0 | 0 |
| Total P1 | | 10 | 8 | 4 | 4 | 1 | 2 | 1 | 1 |
| P2 | G1 | 7 | 2 | 1 | 4 | 1 | 1 | 2 | 1 |
| | G2 | 7 | 2 | 1 | 4 | 1 | 1 | 2 | 1 |
| | G3 | 8 | 5 | 2 | 3 | 0 | 1 | 0 | 0 |
| | G4 | 7 | 5 | 2 | 3 | 0 | 1 | 1 | 0 |
| | G5 | 8 | 5 | 2 | 3 | 1 | 0 | 0 | 0 |
| | G6 | 7 | 5 | 3 | 3 | 0 | 1 | 0 | 0 |
| | G7 | 8 | 5 | 2 | 2 | 0 | 1 | 1 | 0 |
| | G8 | 9 | 5 | 1 | 3 | 0 | 0 | 1 | 0 |
| | G9 | 8 | 5 | 2 | 3 | 0 | 1 | 0 | 0 |
| | G10 | 8 | 5 | 2 | 3 | 0 | 1 | 0 | 0 |
| Total P2 | | 9 | 6 | 3 | 6 | 2 | 2 | 2 | 1 |

After cleaning the dataset and removing the missing and incomplete values, we ended up having three possible treatments that are Hearing Aid (HA), Sound Generators (SG), and Regular Consultation (RC). Thus, we grouped the patients into those three different groups. The results of the grouping based on same meta-actions are summarized in **Table 6**.

We then generated action rules from each group by setting up the minimum confidence to 85% and minimum support to 3; however, the support varies from one group to another depending on the group's action rules strength. We also fixed the decision feature transition from 0 to 1 (no improvement to improvement in the score $(ScT, 0 \rightarrow 1)$**)**. **Table 7** summarizes the results of action rules discovery:

We can note that the minimum support strength of the discovered action rules in each group is positively correlated to the number of patients in each group.

Here is the example of an action rule discovered in the RC group:

$$(g,3 \rightarrow 3) \wedge (C.8,4 \rightarrow 0) \wedge (F.7,4 \rightarrow 0) \wedge (C.19,4 \rightarrow 0)$$
$$\wedge (E.10,4 \rightarrow 0) \wedge (E.17,4 \rightarrow 0) \wedge (E.16,4 \rightarrow 0)$$
$$\Rightarrow (0 \rightarrow 1), \sup = 15, \text{conf} = 100\%$$

Once we have the groups of patients based on the meta-actions, we further partition each group to subgroups with respect to the same negative side effects. The results are summarized in **Table 8**.

## 5.5. Grouping Schemes Comparison

All three grouping schemes have their advantages and disadvantages with regards to action rules personalization.

**Table 6. Number of patients by meta-action.**

| Group | HA | SG | RC |
|---|---|---|---|
| #Patients | 16 | 87 | 414 |

**Table 7. Action rules by meta-actions groups.**

| Group | Support | AVG Confidence | #Rules |
|---|---|---|---|
| HA | 3 | 100 | 16 |
| SG | 4 | 100 | 16 |
| RC | 21 | 87.5 | 1 |
| | 19 | 87.56 | 7 |

**Table 8. Number of patients by meta-actions and side effects grouping.**

| Groups | $g_{000}$ | $g_{001}$ | $g_{010}$ | $g_{011}$ | $g_{100}$ | $g_{101}$ | $g_{110}$ | $g_{111}$ |
|---|---|---|---|---|---|---|---|---|
| HA | 2 | 0 | 0 | 2 | 1 | 2 | 1 | 8 |
| SG | 9 | 6 | 0 | 10 | 4 | 7 | 1 | 50 |
| RC | 35 | 32 | 17 | 42 | 13 | 35 | 30 | 210 |

The side effects-based grouping is a patients-centric scheme with regards to their negative side effects. It allows the extraction of more personalized action rules since each group dataset used in the action rules discovery process is exclusive to patients with exactly the same negative side effects. As we can note from **Tables 2** and **3** and **Figure 1**, the average group size is decent and the number of action rules generated is high. In addition, **Figure 2** denotes the highest confidence for this scheme. However the support of the extracted rules is rather small in comparison with the other schemes as seen in **Figure 3**. In fact, using the partitioned datasets to extract the rules limits the number of observations; thus, limits the strength of the action rules.

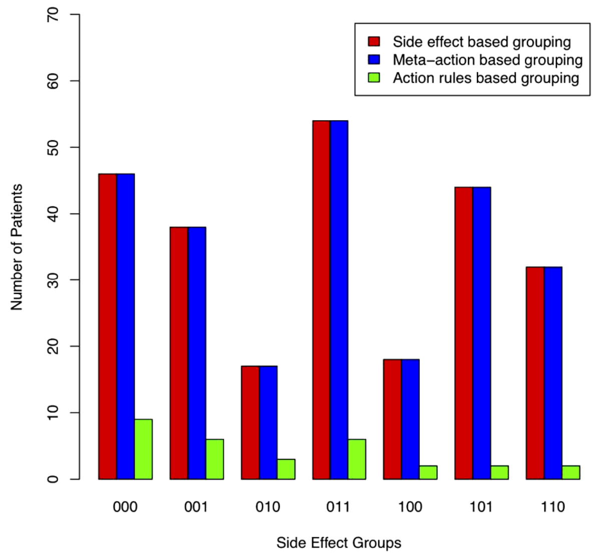


**Figure 1. Patient population by side effect groups (the last group $g_{111}$ was omitted for readability of the figure).**

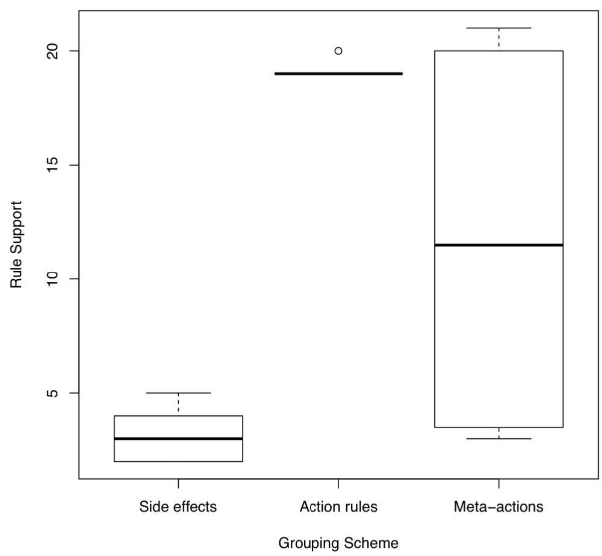


**Figure 2. Support of each grouping scheme.**

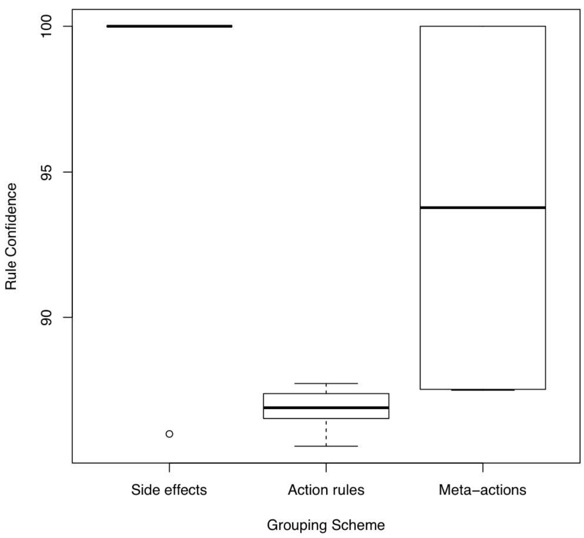


**Figure 3. Confidence of each grouping scheme.**

In this sense, we might argue that the second scheme, action rules based grouping, is more efficient than the previous scheme. It uses the whole dataset to generate the action rules first, and then groups patients by their side effects. It allows extracting action rules with stronger support as seen in **Figure 3** and **Table 4**. This partitioning is the most fine-grained since it allows distinguishing the patients not only by their side effects but also by their personalized rules. This is confirmed in our experiment where the number of patients in the fine-grained side effects groups is very small as seen in **Figure 1** and **Table 5**. This is due to shrinking the datasets or groups to only the action rules domains (patients with extracted applicable action rules $AR_S(x)$ ). The total number of patients will get bigger as we decrease the support, since we will extract more rules. However, the rules confidence is the smallest in this scheme as seen in **Figure 2**. In addition, in order to cover all patients in the dataset, we will eventually have to decrease the support; thus, the strength of the rules.

The third grouping scheme, meta-action based grouping, is the most efficient for our dataset. As we can note from **Table 7** and **Figure 3**, even though we used only subsets of the overall dataset in the meta-actions based grouping, we extracted a higher number of rules with the highest support for the RC group in the later scheme. Furthermore, extracted rules have a higher average confidence as seen in **Figure 2** than the second scheme. In addition, the average number of patients in each side effects based sub-groups is higher than the previous scheme and encompasses all patients for our dataset as seen in **Figure 1**. Grouping by meta-actions filtered all the noise, such as, patients that have the same initial state but different decision feature values then the ones applied in the extracted rules.

The reason for its efficiency is that meta-actions are the source of side effects, and are directly correlated to patients' negative side effects when applied to them. In fact, in our dataset, expert physicians apply treatments (meta-actions) to patient, based on their pathological state, knowing their negative side effects on them.

## 6. Related Work

There have been several research efforts on action rules since their introduction by [1]. The first effort to mine action rules from scratch was done in [7]. However multiple action rules discovery techniques were presented in [9,10].

Actio rules can also be seen as a composition of two classification rules as described in [9], where the authors described how to compute the support and confidence of an action rule based on it's two composing classification rules.

Meta-actions were first introduced in [11] as a higher level concepts used in modeling certain generalizations of actions rules. They were described either by an influence matrix or an antology as a set of value transitions in flexible attributes. Meta-action were formalized in [16], and used in a pruning process with tree classifiers to discover action rules. In [17], the authors show the cascadeing effect of meta-actions leading to a desired effects when generating association action rules and action paths. However, the previous work on meta-action neither studied the side effects of meta-actions nor the action rules personalization problem. In this paper we presented three grouping schemes aiming at discovering personalized action rules.

## 7. Conclusion

Actions rules are very important in modeling expert and domain knowledge. They were used in several domains including healthcare and music information retrieval. Action rules were augmented by the introduction of meta-actions that experts use to trigger them. In this paper, we studied the tinnitus handicap disease, and noted the importance of filtering negative side effects in applying treatments to patients. We then defined formally the negative side effects, and proposed to extract personalized action rules based on these side effects. We believe that expert physicians partition patients with the same pathological state based on their side effects response to treatments. Therefore, we proposed three grouping schemes with regards to negative side effects to extract patient's personalized action rules. We also implemented the three grouping schemes and tested them on the Tinnitus handicap dataset. We further compared the three grouping schemes and discussed their advantages and disadvantages, and justified our choice to apply the meta-action based grouping scheme. We trust that personalization is a very important aspect in filtering noise that skilled experts face when making decisions.

## REFERENCES


[1] Z. W. Ras and A. Wieczorkowska, "Action-Rules: How to Increase Profit of a Company," In: D. A. Zighed, J. Komorowski and J. Zytkow, Eds., *Principles of Data Mining and Knowledge Discovery*, Springer, Lyon, Vol. 1910, 2000, pp. 587-592.

[2] A. Hajja, A. Wieczorkowska, Z. W. Ras and R. Gubrynowicz, "Object-Driven Action Rules and Their Application to Hypernasality Treatment," *Proceedings of ECML-PKDD Workshop on New Frontiers in Mining Complex Patterns*, Bristol, 24-28 September 2012, pp. 104-115.

[3] Z. W. Ras, A. Dardzinska, L. S. Tsay and H. Wasyluk, "Association Action Rules," *IEEE International Conference on Data Mining Workshops*, Pisa, 15-19 December 2008, pp. 283-290.

[4] H. Wasyluk, Z. W. Ras and E. Wyrzykowska, "Application of Action Rules to HEPAR Clinical Decision Support System," *Journal of Experimental and Clinical Hepatology*, Vol. 4, No. 2, 2008, pp. 46-48.

[5] X. Zhang, Z. W. Ras, P. J. Jastreboff and P. L. Thompson, "From Tinnitus Data to Action Rules and Tinnitus Treatment," *Proceedings of* 2010 *IEEE Conference on Granular Computing*, IEEE Computer Society, Silicon Valley, 2010, pp. 620-625.

[6] S. Greco, B. Matarazzo, N. Pappalardo and R. Slowinski, "Measuring Expected Effects of Interventions Based on Decision Rules," *Journal of Experimental and Theoretical Artificial Intelligence*, Vol. 17, No. 1-2, 2005, pp. 103-118. doi:10.1080/09528130512331315864

[7] Z. He, X. Xu, S. Deng and R. Ma, "Mining Action Rules from Scratch," *Expert Systems with Applications*, Vol. 29, No. 3, 2005, pp. 691-699. doi:10.1016/j.eswa.2005.04.031

[8] Y. Qiao, K. Zhong, H.-A. Wang and X. Li., "Developing Event-Condition-Action Rules in Real-Time Active Database," *Proceedings of the* 2007 *ACM Symposium on Applied Computing*, Seoul, 11-15 March 2007, pp. 511-516.

[9] Z. Ras and A. Dardzinska, "From Data to Classification Rules and Actions," *International Journal of Intelligent Systems*, Vol. 26, No. 6, 2011, pp. 572-590.

[10] J. Rauch and M. Simunek, "Action Rules and the GUHA Method: Preliminary Considerations and Results," *Foundations of Intelligent Systems*, Vol. 5722, 2009, pp. 76-87.

[11] K. Wang, Y. Jiang and A. Tuzhilin, "Mining actionable Patterns by Role Models," *Proceedings of* 22*nd International Conference on Data Engineering*, Atlanta, 3-7 April 2006, pp. 16-26.

[12] H. Zhang, Y. Zhao, L. Cao and C. Zhang, "Combined Association Rule Mining," In: T. Washio, E. Suzuki, K. M. Ting and A. Inokuchi, Eds., *Proceedings of the* 12*th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining*, LNCS, Springer, Vol. 5012, 2008, pp. 1069-1074.

[13] J. Rauch, "Action4ft-Miner Module," Lisp-Miner Project, 2012. http://lispminer.vse.cz

[14] Z. Pawlak, "Information Systems—Theoretical Foundations," *Information Systems Journal*, Vol. 6, No. 3, 1981, pp. 205-218. doi:10.1016/0306-4379(81)90023-5

[15] Z. W. Ras, E. Wyrzykowska and H. Wasyluk, "ARAS: Action Rules Discovery Based on Agglomerative Strategy," *Mining Complex Data*, Vol. 4944, 2008, pp. 196-208.

[16] Z. W. Ras and A. Dardzinska, "Action Rules Discovery Based on Tree Classifiers and Meta-Actions," *Foundations of Intelligent Systems*, Vol. 5722, 2009, pp. 66-75. doi:10.1007/978-3-642-04125-9_10

[17] A. Tzacheva and Z. W. Ras, "Association Action Rules and Action Paths Triggered by Meta-Actions," *IEEE Conference on Granular Computing*, San Jose, 14-16 August 2010, pp. 772-776.


 