Pair-based Object-driven Action Rules

Ayman Hajja[1], Alicja A. Wieczorkowska[2], Zbigniew W. Raś[1,3], and Ryszard Gubrynowicz[2]

[1] University of North Carolina, Dept. of Computer Science, 9201 University City Blvd., Charlotte, NC 28223, USA
[2] Polish-Japanese Institute of Information Technology, Koszykowa 86, 02-008 Warsaw, Poland
[3] Warsaw University of Technology, Institute of Computer Science, Nowowiejska 15/19, 00-665 Warsaw, Poland

ahajja@uncc.edu
alicja@poljap.edu.pl
ras@uncc.edu
rgubryn@pjwstk.edu.pl

Abstract. Action rules, as proposed by Raś and Wieczorkowska in [11], can be defined as actionable tasks that describe possible transitions of objects from one state to another with respect to a distinguished attribute. Recently, a new specialized case of action rules, namely object-driven action rules, has been introduced by Ayman et al. in [4]. Object-driven action rules are action rules that are extracted from information systems with temporal and object-based nature. By object-based nature, we refer to systems that contain multiple observations for each object. A typical example of an object-based system would be a system of patients recording multiple visits; each patient is considered a distinct object. In this paper, we will further investigate the concept of object-driven action rules by proposing a new pair-based way of examining object-driven systems, which we believe is more intuitive for temporal and object-driven systems. The focus of this paper will be on our proposed pair-based approach, along with the modifications required to extract action rules and calculate their properties.

Keywords: action rules, object-driven action rules, temporal data, hypernasality

1 Introduction and Background

Action Rules, as proposed by Raś and Wieczorkowska in [11], describe possible transitions of objects from one state to another with respect to a specific attribute, called the decision attribute. Action rules have been successfully applied in many domain areas including business [11], medical diagnosis and treatment [16], [17], and music automatic indexing and retrieval [6], [9].

System users are mainly interested in actionable tasks that trigger state transitions that move objects from a less desirable state to a more desirable state;

action rules specify the actions needed to be taken to reach that desired goal.

In this paper, we will introduce a novel approach for extracting action rules from object-driven and temporal systems. There has been considerable research on the varied methodologies for extracting action rules from information systems [5, 7, 13]. However, adapted action rules systems that are designed for datasets with particular nature is to some extent new. In [4], Ayman et al. proposed an adapted action rules extraction method for information systems of temporal and object-based nature. In this work, we will extend the approach presented in [4].

2 Object-driven Action Rules Revisited

The drive behind the introduction of object-driven action rules in [4] was to bring forth an adapted approach to extract action rules from systems of temporal and object-driven nature.

In [4], we proposed an object-independency assumption that suggests extracting patterns from subsystems defined by unique objects, and then aggregating similar patterns amongst all objects. The motivation behind this approach is based on the fact that same-object observations share similar features that are not shared with other objects, and these features are possibly not explicitly included in our dataset. Therefore, by individualizing objects prior to calculating action rules, variance is reduced, and over-fitting is potentially avoided. In addition to the object-independency assumption, temporal information is exploited by taking into account only the state transitions that occurred in the valid direction.

In this section, we will start with providing the necessary background concerning action rules. A complete description of the motivation and the concept of the pair-based approach will be presented next, along with the modifications required to extract object-driven action rules and calculate their properties.

2.1 Action Rules

The notion of action rules was first proposed by Z. W. Raś and A. Wieczorkowska in [11]. Action rules describe possible transition of objects from one state to another with respect to a specific attribute, called the decision attribute. The goal of action rules is to provide system users with actionable tasks that can be directly applied to objects listed in information systems to reach a desired goal.

Let $S = (X, A, V)$ denotes an information system [8], where:

1. X is a nonempty, finite set of instances (objects),
2. A is a nonempty, finite set of attributes;
   $a : X \to V_a$ is a function for any $a \in A$, where V_a is called the domain of a,

3. $V = \bigcup\{V_a : a \in A\}$.

By a decision table, we mean an information system that makes a clear explicit distinction between attributes in A, and will therefore label each attribute as either a *decision attibute*, or a non-decision attribute, called *condition attribute*. The decision attribute(s), which normally but not necessarily is a single attribute, is the attribute that we are interested in most. For system users, the eventual goal would be to change the decision attribute from less desirable to more desirable state. For example, a company would be interested in moving clients' states of loyalty from lower to higher.

All non-decision, or condition, attributes are further partitioned into two mutually exclusive sets; the first one is the *stable* attributes set, and the second one is the *flexible* attributes set. By stable attributes set we mean the set that contains attributes that we have no control over; their values cannot be changed by the users of our system. An example of a *stable* attribute is the place where the person was born. On the other hand, values of *flexible* attributes can be influenced and changed; an example of a *flexible* attribute is the patient's prescribed medications. In this paper, A_{St}, A_{Fl}, and $\{d\}$ will represent the set of stable attributes, the set of flexible attributes, and the decision attribute, respectively. Hence, the set of attributes A can be redefined as $A = A_{St} \cup A_{Fl} \cup \{d\}$.

An *atomic action set* is an expression that defines a change of state for a single distinct attribute. For example, $(a, a_1 \rightarrow a_2)$ is an atomic action set which defines a change of state for the attribute a from a_1 to a_2, where $a_1, a_2 \in V_a$. Clearly, in this case, the attribute a is a flexible attribute, since it changes its state from a_1 to a_2. In the case when there is no change, we omit the right arrow sign, so for example, (b, b_1) means that the value of attribute b remains b_1, where $b_1 \in V_b$.

An *action set* is defined as follows:

1. If t is an atomic action set, then t is an action set.
2. If t_1, t_2 are action sets and $\wedge$ is a 2-argument functor called composition, then $t_1 \wedge t_2$ is a candidate action set.
3. If t is a candidate action set and for any two atomic action sets $(a, a_1 \rightarrow a_2), (b, b_1 \rightarrow b_2)$ contained in t we have $a \neq b$, then t is an action set.
4. No other sets are called action sets.

The *domain* $Dom(t)$ of an action set t is the set of attributes of all atomic action sets contained in t. For example, $t = (a, a_1 \rightarrow a_2) \wedge (b, b_1)$ is an action set that consists of two atomic action sets, namely $(a, a_1 \rightarrow a_2)$ and (b, b_1). Therefore, the domain of t is $\{a, b\}$.

Action rules are expressions that take the following form: $r = [t_1 \Rightarrow t_2]$, where t_1, t_2 are action sets. The interpretation of the action rule r is that by applying the action set t_1, we would get, as a result, the changes of states in action set

t_2. We also assume that $Dom(t_1) \cup Dom(t_2) \subseteq A$, and $Dom(t_1) \cap Dom(t_2) = \phi$.

For example, $r = [[(a, a_1 \to a_2) \wedge (b, b_2)] \Rightarrow (d, d_1 \to d_2)]$ means that by changing the state of the attribute a from a_1 to a_2, and by keeping the state of the attribute b as b_2, we would observe a change in the attribute d from the state d_1 to d_2, where d is commonly referred to as the *decision attribute*.

Standard interpretation N_s of action sets in S is defined as follows:

1. If $(a, a_1 \to a_2)$ is an atomic action set, then $N_s((a, a_1 \to a_2)) = [\{x \in X : a(x) = a_1\}, \{x \in X : a(x) = a_2\}]$.
2. If $t_1 = (a, a_1 \to a_2) \wedge t$ and $N_s(t) = [Y_1, Y_2]$, then $N_s(t_1) = [Y_1 \cap \{x \in X : a(x) = a_1\}, Y_2 \cap \{x \in X : a(x) = a_2\}]$.

Let us define $[Y_1, Y_2] \cap [Z_1, Z_2]$ as $[Y_1 \cap Z_1, Y_2 \cap Z_2]$ and assume that $N_s(t_1) = [Y_1, Y_2]$ and $N_s(t_2) = [Z_1, Z_2]$. Then, $N_s(t_1 \wedge t_2) = N_s(t_1) \cap N_s(t_2)$.

If t is an action set and $N_s(t) = [Y_1, Y_2]$, then the support of t in S is defined as $supp(t) = min\{card(Y_1), card(Y_2)\}$.

Let $r = [t_1 \Rightarrow t_2]$ be an action rule, $supp(t_1) > 0, N_s(t_1) = [Y_1, Y_2]$, and $N_s(t_2) = [Z_1, Z_2]$. Support $supp(r)$ and confidence $conf(r)$ of r are defined as:

$$supp(r) = min\{card(Y_1 \cap Z_1), card(Y_2 \cap Z_2)\},$$

$$conf(r) = \left[\frac{card(Y_1 \cap Z_1)}{card(Y_1)}\right] * \left[\frac{card(Y_2 \cap Z_2)}{card(Y_2)}\right].$$

2.2 Action Rules Extraction

There have been a considerable amount of research on various methodologies for extracting action rules from information systems [5, 7, 13]. In general however, we can categorize all methodologies into two groups; the first one being when classification rules are required prior to the construction of action rules [12], [15], and the second, more recent approach, being when action rules are directly extracted from an information system [10].

To extract pair-based object-driven action rules, we used the algorithm described in [10]. The idea of the algorithm is to start by constructing all possible action sets that have occurred more than a pre-defined number, called the minimum support. Then, in accordance to our desired change in the decision attribute, action rules are formed.

Let t_a be an action set, where $N_s(t_a) = [Y_1, Y_2]$ and $a \in A$. We say that t_a is a *frequent action set* [10] if $card(Y_1) \geq \lambda_1$ and $card(Y_2) \geq \lambda_1$, where λ_1

is the minimum support. Another way of interpreting the frequent action sets would be that all frequent action sets have support greater than or equal to the minimum support λ_1. By specifying λ_1, we make sure that the extracted action rules have support greater than or equal to the minimum support λ_1. Algorithm presented below is similar to [1].

To extract action rules, we start with generating atomic action sets that have support greater than or equal to the minimum support value λ_1 pre-defined by the user; we will refer to this set as *1-element frequent action set.* The term *frequent* will be used to indicate that an action set has support greater than or equal to the minimum support, and the term *k-element* will be used to indicate the number of elements (or atomic action terms) in an action set. Both *frequent atomic action sets* and *1-element frequent action set* refer to exactly the same set, since from the definition of action sets, they consist of only one element.

After generating all frequent atomic action sets, we undertake the following two-step process initially for $k = 1$:

1. **Merge step:** Merge pairs (t_1, t_2) of k-element action sets into all $(k+1)$-element candidate action sets.
2. **Delete step:** Delete all $(k+1)$-element candidate action sets that are either not action sets, or contain a non-frequent k-element action set, or that have support less than the minimum support λ_1.

We keep iterating the above two steps until we cannot generate new frequent action sets anymore. At this point, we have generated all $(k+1)$-element frequent action sets, which will allow us to generate action rules that are guaranteed to have support greater than or equal to the minimum support λ_1. Last step is to further filter the desired action rules based on their confidence, where we only consider action rules with confidence greater than or equal to a pre-defined minimum confidence λ_2.

For example, from the frequent action set $t_1 = (a, a_1 \to a_2) \wedge (d, d_1 \to d_2)$, we can generate the following two action rules:

1. $r_1 = [(a, a_1 \to a_2) \Rightarrow (d, d_1 \to d_2)]$.
2. $r_2 = [(d, d_1 \to d_2) \Rightarrow (a, a_1 \to a_2)]$.

where both r_1 and r_2 have support greater than or equal to the minimum support λ_1. However, we will only be interested in specific changes of the decision attribute, e.g. in changing the decision attribute d from state d_1 to d_2. Therefore, we will only consider r_1.

2.3 Temporal Constraint and Pair-based Approach

As defined previously, temporal object-driven datasets consist of numerous unique objects, where each object is comprised of multiple instances that have assigned corresponding timestamps. Previously in [4], the object p based standard interpretation of an action set $t = (a, a_1 \rightarrow a_2)$ was defined as the pair of two sets $[Y_1, Y_2]$ where Y_1 is the set of instances of the object p that satisfy the left side, or condition side, of the action set, and Y_2 is the set of instances of the object p that satisfy the right side, or decision side, of the action set, with the addition that for every instance in Y_1, there exist a matching instance in Y_2 that occurred after it. This definition resembles the definition of standard interpretation for classical action rules while restricting valid transitions to only one direction. In this paper however, we argue that the nature of the object-driven temporal dataset allows us to redefine the standard interpretation into a more intuitive pair-based structure which we believe is more appropriate for object-driven temporal systems.

Let us first assume that $I_s(p)$ denotes the set of all instances of the object p in an information system S. Also, the relation $\angle \subseteq I_s(p)$ is defined as:

$(p_1, p_2) \in \angle$ iff p_2 has occurred after p_1 .

The pair-based standard interpretation $N^{TC}_{s(p)}$ in $S = (X, A, V)$ for an object p is redefined as:

1. If $(a, a_1 \rightarrow a_2)$ is an atomic action set, then
$N^{TC}_{s(p)}((a, a_1 \rightarrow a_2)) = \{(p_1, p_2) \in \angle : a(p_1) = a_1, a(p_2) = a_2\}$
where $\angle \subset I_s(p)$.
2. If $t_1 = (a, a_1 \rightarrow a_2) \wedge$ and $N^{TC}_{s(p)}(t) = Y_1$, then
$N^{TC}_{s(p)}(t_1) = Y_1 \cap \{(p_1, p_2) \in \angle : a(p_1) = a_1, a(p_2) = p_2\}$
where $\angle \subset I_s(p)$.

In other words, our standard interpretation will consist of all valid transitions from the left side of an action set to the right side, represented as pairs. The motivation behind this new interpretation is due to the fact that the instances within one object are not observed independently, which will allow us to relax the minimum assumption previously used. Our object-independency assumption states that the whole system consists of multiple independent subsystems, each one marked by a unique object. Although it confines the system to extract action rules only from instances of the same object, it provides more flexibility to be applied within unique objects.

If t is an action set and $N^{TC}_{s(p)}(t) = Y_1$, then the support of t in S is defined as: $supp^{TC}_p = card(Y_1)$.

Let $r = [t_1 \Rightarrow t_2]$ be an action rule, where $N^{TC}_{s(p)}(t_1) = Y_1, N^{TC}_{s(p)}(t_2) = Y_2$. The p^{th} support $supp^{TC}_p(r)$ and the p^{th} confidence $conf^{TC}_p(r)$ of r are defined as follows:

$$supp^{TC}_p(r) = card(Y_1 \cap Y_2),$$

$$conf_p^{TC}(r) = \left[\frac{card(Y_1 \cap Y_2)}{card(Y_d)}\right].$$

To define Y_d, let us first assume that $\zeta(Y)$ denotes the set of first elements of the set of pairs Y. For instance, if $Y = \{(p_1, p_3), (p_3, p_4), (p_1, p_2)\}$, then $\zeta(Y) = \{p_1, p_3\}$. We define $Y_d = \{(p_1, p_2) \in Y_1 : p_1 \in \zeta(Y_2)\}$. The interpretation of this definition means that to calculate the confidence of the action rule $r = (a, a_1 \to a_2) \Rightarrow (d, d_1 \to d_2)$, the pairs that we are considering are the ones that have first elements that satisfy $a = a_1$ and $d = d_1$. Since the transition from a_1 to a_2 could possibly trigger other states of decision attribute d, we are only interested in the states of our action rule.

After all object-driven action rules are extracted and their p^{th} support and p^{th} confidence are computed for all $p \in X$, we then calculate their total support $supp_X^{TC}(r)$ (called support) and total confidence $conf_X^{TC}(r)$ (called confidence) following the definition below:

$$supp_X^{TC}(r) = \sum_{p \in X} supp_p^{TC}(r) \ ,$$

$$conf_X^{TC}(r) = \sum_{p \in X} \left(\frac{supp_p^{TC}(r) * conf_p^{TC}(r)}{supp_X^{TC}(r)}\right).$$

If the denominator in the formula for calculating confidence is equal to zero, then the confidence is equal to zero by definition.

Example to demonstrate pair-based object driven support and confidence: Here, we provide an example to demonstrate how we calculate the support and the confidence for the whole system S shown in Table 1. We assume that for all 3 objects in X their instances x_i, where $1 \leq i \leq 10$, have chronological order.

Referring to our information system S shown in Table 1, we calculate the support $supp_X^{TC}(r)$ and the confidence $conf_X^{TC}$ for the following rule:

$$r = [(a_1 \to a_2) \wedge (c, c_1) \Rightarrow (d, d_1 \to d_2)].$$

We first calculate the p^{th} standard interpretation for each object p (e.g. for a patient) for both the condition and the decision parts in the action rule r:

$N_{s(1)}^{TC}((a, a_1 \to a_2) \wedge (c, c_1)) = \{(x_0, x_1), (x_0, x_2), (x_0, x_4), (x_3, x_4)\} \cap$
$\{(x_0, x_1), (x_0, x_3), (x_0, x_4), (x_1, x_3), (x_1, x_4), (x_3, x_4)\}$
$= \{(x_0, x_1), (x_0, x_4), (x_3, x_4)\}$,

Table 1. Information System S

	$objectID$	a	b	c	d
x_0	1	a_1	b_1	c_1	d_1
x_1	1	a_2	b_1	c_1	d_1
x_2	1	a_2	b_2	c_2	d_2
x_3	1	a_1	b_2	c_1	d_1
x_4	1	a_2	b_1	c_1	d_2
x_5	2	a_1	b_2	c_1	d_2
x_6	2	a_2	b_1	c_1	d_1
x_7	3	a_1	b_2	c_1	d_1
x_8	3	a_2	b_2	c_1	d_2
x_9	3	a_1	b_1	c_1	d_1
x_{10}	3	a_2	b_1	c_1	d_2

$$N^{TC}_{s(1)}(d, d_1 \to d_2) = \{(x_0, x_2), (x_0, x_4), (x_1, x_2), (x_1, x_4), (x_3, x_4)\} \ ,$$

$$N^{TC}_{s(2)}((a, a_1 \to a_2) \wedge (c, c_1)) = \{(x_5, x_6)\} \ ,$$

$$N^{TC}_{s(2)}(d, d_1 \to d_2) = \phi \ ,$$

$$N^{TC}_{s(3)}((a, a_1, \to a_2) \wedge (c, c_1)) = \{(x_7, x_8), (x_7, x_{10}), (x_9, x_{10})\} \cap$$
$$\{(x_7, x_8), (x_7, x_9), (x_7, x_{10}), (x_8, x_9), (x_8, x_{10}), (x_9, x_{10})\}$$
$$= \{(x_7, x_8), (x_7, x_{10}), (x_9, x_{10})\} \ ,$$

$$N^{TC}_{s(3)}(d, d_1 \to d_2) = \{(x_7, x_8), (x_7, x_{10}), (x_9, x_{10})\} \ .$$

Using the temporal constraint and the object-driven assumptions, the pair-based support and confidence for each object is calculated as follows:

$$sup^{TC}_1(r) = card(\{(x_0, x_4), (x_3, x_4)\}) = 2 \ ,$$

$$conf^{TC}_1(r) = \left[\frac{card(\{(x_0, x_4), (x_3, x_4)\})}{card(\{(x_0, x_1), (x_0, x_4), (x_3, x_4)\})}\right] = \frac{2}{3} \ ,$$

$$sup^{TC}_2(r) = card(\phi) = 0 \ ,$$

$$conf^{TC}_2(r) = 0 \ ,$$

$$sup_3^{TC}(r) = card(\{(x_7, x_8), (x_7, x_{10}), (x_9, x_{10})\}) = 3 \ ,$$

$$conf_3^{TC}(r) = \left[\frac{card(\{(x_7, x_8), (x_7, x_{10}), (x_9, x_{10})\})}{card(\{(x_7, x_8), (x_7, x_{10}), (x_9, x_{10})\})}\right] = \frac{3}{3} = 1 \ .$$

Now we calculate the overall support and confidence for the whole system:

$$sup_X^{TC}(r) = 5, \ conf_X^{TC}(r) = \left(\frac{2 * \frac{2}{3}}{5}\right) + \left(\frac{3 * 1}{5}\right) = \frac{4.33}{5} = .87 \ .$$

3 Experimental Data: Hypernasality Data Set

Distortions of the velopharyngeal closure, resulting in speech hypernasality or hyponasality, may cause speech disorders in children [3]. The patient's nasopharynx disorders have been examined in the Children's Memorial Health Institute in Warsaw for many years. The gathered data also include general information on the patient's condition if it can be of importance, e.g. cerebral palsy, neurology, or myopathy. This way a reach collection of complex data describing hypernasality was gathered, in close cooperation with one of the co-authors, Prof. Ryszard Gubrynowicz, who is a speech scientist and expert in this area; the data were collected when he was working in the Children's Memorial Health Institute.

3.1 Velum Malfunction in Children

Hypernasality can be examined by means of Czermak's mirror test of nasal air escape, see Figure 1. The child is asked to repeat several times a syllable composed of a plosive consonant and an open vowel, e.g. /pa/-/pa/-/pa/, and the sizes of the fogging circles appearing on the mirror are rated on 4–point scale, from 0 (no hypernasality) to 3 (most severe hypernasality). Therefore, $Czermak's\ mirror\ test$ was used as a decision attribute in the nasality data set. All attributes, representing various medical conditions in the examined children, are listed in Table 2. More explanations about these attributes are given below.

Each patient was examined several times. Personal data were recorded (first name and last name, sex), and for each examination the age of the child was marked. Personal data were removed before further processing, and replaced with ID data, representing the patient's ID combined with the sequential number of this patient's visit.

During each visit, the articulation of selected vowels and consonants was recorded, and the recording date was marked (*recording date* attribute). The data stored in columns marked as *diagnosis* and *diagnosis*2 describe patient's condition related to nasality; only one diagnosis is stored in each of these columns,

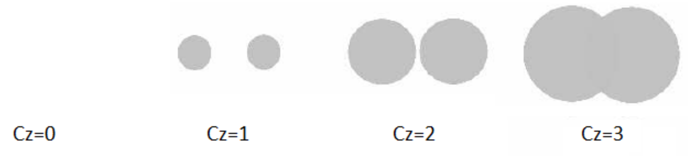

Fig. 1. Czermak's mirror fogging test, rating the degree of the patient's nasal air escape on a 4-point scale: none = 0; small = 1, medium = 2, large = 3 [3].

so *diagnosis*2 represents additional diagnosis, if there is more than one. The following diagnoses are described in these columns: R - cleft, RP - cleft palate, OR - after cleft palate surgery, WKP - congenital short velum, NO - hypernasality, NZ - hyponasality, BR - no diagnosis, PRP - submucous cleft palate, AT - after tonsillectomy, DKP - quite short palate, RJ - cleft uvula, III - hypertrophy of adenoids and possibly palatine tonsils, MP - hypertrophy of palatine tonsils, MPDz - cerebral palsy, AD - after adenotomy, ADT - after adenotonsillectomy, UK - larynx after injury/trauma, NS - hypoacusis, ORM - retarded speech development, NEU - neurology, ONR - after neurological surgery. If NO (hypernasality) is diagnosed and marked in the column *diagnosis*, it represents the most severe case of hypernasality. The numbers 0–3 in *diagnosis*2 refer to sleep apnoea, i.e. temporary cessation of respiration during sleep. 0 means no apnoea, 3 - very often. Sleep apnoea is also represented as a separate attribute, but the values assessed for the same patient may differ significantly, so they were kept in both columns. Generally, physicians may differ in their opinions, this is why we must be prepared to deal with some inconsistencies in the data. More of diagnostic details are given in the column *comments*, but these comments are not taken into account in the current version of our action rule software.

Other physical conditions recorded in the database include the degree of hypertrophy of adenoids and possibly palatine tonsils, and the degree of motility of the soft palate, represented as *tonsils* and *motility* attributes. The assessment of the patient's recorded speech is represented in the following attributes: *yeaoui* (vowels /I, e, a, o, u, i/ - a sequence of short vowel sounds spoken in isolation), $i - long$ (long vowel /i/ - vowel of sustained phonation), and *bdg* (high pressure consonants /b, d, g/); SAMPA coding of phonetic alphabet is used [14]. These attributes describe the measure of nasalization (coefficient of nasalization), calculated from the analysis of mouth and nose signals (separately recorded), as the ratio of the nose signal level to the sum of the level of the nose and mouth signals for the phonemes indicated in each attribute. *difference level* $F1 - F2$ describes the vocal tract's first 2 resonances as the difference level of the 1^{st} and the 2^{nd} formant, measured for /i/-long.

The best diagnosis we are interested in is when the parameters' values are in normal ranges. Our decision attribute is Czermak's mirror test, so its values are most important in our research. The most desired value of our decision

Table 2. Attributes in the Hypernasality Data Set. Expansions of acronyms are given in the text, see Section 3.1.

Attribute	Description
ID	Patient's ID, with the sequential number of his/her visit
age	Age [years, months]
sex	Sex {M, F}
$recording\ date$	Recording Date [yyyy.mm.dd]
$diagnosis$	Diagnosis {AD, ADT, AT, BR, III, myopathy, MPDz, NEU, NO, ONR, OR, ORM, RJ, RP, UK, WKP}
$comments$	Comments, details of the diagnosis
$diagnosis2$	Diagnosis {0, 1, 2, 3, DKP, RJ, WKP}
$sleep\ apnoea$	Sleep apnoea {0, 1, 2, 3}
$tonsils$	Hypertrophy of adenoids and possibly palatine tonsils {0, 1, 2, 3}
$Czermak's\ mirror\ test$ - decision attribute	Mirror-fogging test {0, 1, 2, 3}
$yeaoui$	Measure of nasalization for vowels /I, e, a, o, u, i/ [0, 100]
$i - long$	Measure of nasalization for vowel /i/-long [0, 100]
bdg	Measure of nasalization for high pressure consonants /b, d, g/ [0, 100]
$motility$	Motility of the soft palate [0, 12]
difference level $F1 - F2$	The difference level of 1^{st} & 2^{nd} formant measured for /i/-long [-14, 26]

attribute is when it is equal to 0. The diagnosis is worse when Czermak's test value equals 2, next worse case is when Czermak's test value equals 3, and this is the most severe case. The lower the Czermak's test value, the better the diagnosis is. Therefore, we are interested in action rules indicating how to decrease the Czermak's test value. The goal of our system is to find action rules which purpose is to provide hints referring to doctor's interventions. They show how values of certain attributes need to be changed (through various medical procedures, according to the physician's order), so the patient's condition will get improved.

4 Application of Object-driven Action Rules

In this work, we derived a new set of attributes in accordance to [4]. In addition to our attributes shown in Table 2, for each of the following four attributes: *yeaoui, i - long, bdg*, and *motility*, two new attributes were derived, resulting in eight new attributes. The two derived attributes are the difference, and the rate of change for every two consecutive instances, which we calculated as follows:

1. The difference of values for *yeaoui, i - long, bdg* and *motility* for every two consecutive visits is calculated, thus constituting the following new attributes: $yeaoui_1, i_1 - long, bdg_1$ and $motility_1$. For example, the value of

bdg_1 equals to the value of bdg for the $(k+1)^{th}$ visit minus the value for the k^{th} visit.

2. The rate of change a_2 for every two consecutive visits is defined as:

$$a_2 = arctan\left(\frac{a_1}{\text{age difference in months}}\right)$$

where a_1 is the difference of values of the attribute a for the two visits.

After calculating the derived attributes, we used the Rough Set Exploration System [2] to discretize our real-valued attributes wrt. our decision attribute. Next, our temporal object-driven action rule discovery system, presented in Section 2.3, was applied to the discretized data.

Our decision attribute *Czermak's mirror test* was not discretized. Moreover, when a physician could not decide between two neighboring Czermak's test values, an intermediate value was assigned. Therefore, the decision values are {0, .5, 1, 1.5, 2, 2.5, 3}.

5 Results and Discussion

In this section we show a sample of the results after running our proposed pair-based approach to extract object-driven action rules from temporal systems. We show that by using pair-based approach, not only we were able to extract a dramatically larger set of action rules, but also we were able to extract action rules that provide more dramatic decrease of patient severity than the rules extracted in [4]. For an action rule to be eligibly used on a patient, the pre-conditions of the action rule and the patient's current condition have to match, meaning that only a subset of our patients will benefit from each particular action rule. Having said that, using our pair-based approach to extract action rules will generate a significant amount of action rules that can be appropriately used for various sets of patients.

Rule 1. $r_1 = (\textit{difference level F1-F2}, \geq 9.5 \rightarrow [6.5, 9.5))$
$\Rightarrow (\textit{Czermak's mirror test}, 3 \rightarrow 2);\quad supp(r_1) = 2,\quad conf(r_1) = 100\%$.

This rule means that by decreasing the difference between the first two formants of the vocal tract for /i/ - long, we would notice a decent shift of the Czermak's mirror test, decreasing from 3 to 2. In [4], we extracted a similar action rule that also indicated the importance of *difference level F1-F2* attribute. However, this action rule is exclusive to the work described in this paper.

Rule 2. $r_2 = (i_2 - long, \geq 5.5 \rightarrow < 5.5) \Rightarrow (\textit{Czermak's mirror test}, 3 \rightarrow 2)$; $supp(r_2) = 3,\quad conf(r_2) = 66.7\%$.

This rule means that decreasing the value of $i - long$ in a short period of time, since $i_2 - long$ is defined as the rate of change, will result in a similar decrease of the Czermak's mirror test from 3 to 2. Again, this rule affirms the importance of the attribute $i - long$.

Rule 3. $r_3 = (i_2 - long, \geq 5.5 \rightarrow < 5.5) \wedge (bdg, \geq 8.5)$
$\Rightarrow (Czermak\text{'s mirror test}, 2.5 \rightarrow 2);\quad supp(r_3) = 2,\quad conf(r_3) = 100\%$.

This rule is similar to Rule 1. It confirms the effect of decreasing the rate of change of the nasalization measured for /i/ - long, but also adds an additional condition concerning the nasality of /bdg/, that is, this rule only applies to patients suffering from high nasality for /bdg/ (≥ 8.5).

Rule 4. $r_4 = (tonsils, < 2) \wedge (i_2 - long, \geq 5.5 \rightarrow < 5.5) \wedge (motility, [4.5, 5.5))$
$\Rightarrow (Czermak\text{'s mirror test}, 2 \rightarrow 1.5);\quad supp(r_4) = 2,\quad conf(r_4) = 100\%$.

This rule states that when a patient is experiencing a little hypertrophied adenoids and possibly palatine tonsils ($tonsils < 2$), we can slightly improve his condition from Czermak's mirror test 2 to 1.5 by decreasing the rate of change in /i/ - long, and if the motility of the soft palate does not change.

Rule 5. $r_5 = (bdg, \geq 8.5 \rightarrow [6.5, 8.5)) \Rightarrow (Czermak\text{'s mirror test}, 1 \rightarrow .5);$ $supp(r_5) = 3,\quad conf(r_5) = 66.7\%$.

This rule states that by only decreasing the nasality of /bdg/, we would be able to shift the patients' Czermak's mirror test state from 1 to .5.

Rule 6. $r_6 = (i - long, \geq 9.5 \rightarrow [2.5, 7.5)) \Rightarrow (Czermak\text{'s mirror test}, 1 \rightarrow .5);\quad supp(r_6) = 2,\quad conf(r_6) = 100\%$.

Although the support of this action rule is not high, the rule is rather interesting. It states that by decreasing only one attribute; /i/ - long, there is a 100% chance that the Czermak's mirror test will shift from 1 to .5.

Rule 7. $r_7 = (motility, < 3.5 \rightarrow [4.5, 5.5)) \wedge (diagnosis, OR)$
$\Rightarrow (Czermak\text{'s mirror test}, 1 \rightarrow 0);\quad supp(r_7) = 3,\quad conf(r_7) = 100\%$.

This rule states that if a patient has gone through a cleft palate surgery (OR), then increasing the motility of the soft palate would significantly improve the patient's condition, to the level where the patient is entirely cured, which will result in shifting the Czermak's mirror test value from 1 to 0.

Rule 8. $r_8 = (i_2 - long, \geq 5.5 \rightarrow < 5.5) \Rightarrow (Czermak\text{'s mirror test}, .5 \rightarrow 0);$ $supp(r_8) = 7,\quad conf(r_8) = 71\%$.

This rule has a relatively high support. It states that decreasing the rate of change of $i - long$ from greater than or equal to 5.5, to less than 5.5, will result in curing a light hypernasality.

Rule 9. $r_9 = (i_2 - long, \geq 5.5 \rightarrow < 5.5) \wedge (sleep\ apnoea, < 2)$
$\Rightarrow (Czermak\text{'s mirror test}, .5 \rightarrow 0);\quad supp(r_9) = 6,\quad conf(r_9) = 83\%$.

This rule is a similar, but more specific action rule than rule 8. By expanding the condition side of action rules, we are able to generate action rules with higher confidence. Rule 8 states that by only decreasing the rate of change of $i - long$, we would have a 71% chance of shifting the Czermak's mirror test from .5 to 0. However, rule 9 states that by decreasing the rate of change of $i - long$ and maintaining a low value of sleep apnoea, we would have an 83% chance of shifting Czermak's mirror test from .5 to 0.

Rule 10. $r_{10} = (tonsils, \geq 2 \rightarrow < 2) \Rightarrow (\mathit{Czermak's\ mirror\ test}, .5 \rightarrow 0)$; $supp(r_9) = 5, \quad conf(r_9) = 100\%$.

This rule states, with absolute certainty (confidence 100%), that by decreasing the hypertrophied adenoids and possibly palatine tonsils that the patient is experiencing, the Czermak's mirror test will shift from .5 to 0. Although the improvement does not appear to be significant, the high support and high confidence make this rule highly valuable.

In our hypernasality dataset, most of the patients were experiencing slight to no hypernasality speech (Czermak's mirror test .5 or 0). As a consequence, the last three action rules had a much higher support compared to the others.

6 Summary and Conclusions

In this paper we presented a new approach to examine temporal and object-driven information systems. We proposed a novel pair-based extraction approach that extends the work proposed by Ayman et al. in [4]. In addition to extracting stronger action rules in the same domain of hypernasality speech treatment, we were able to extract a dramatically larger set of action rules that is highly diversified and can be applied to various cases of patients.

One of the authors is still collaborating with physicians, so the outcome of this research can be implemented and tested in practice. This confirms that the obtained rules are in concordance with experience, and they help speech scientists to recapitulate their practical knowledge.

Acknowledgments. The authors would like to thank Dr. Danuta Chojnacka-Wądołowska and Dr. Cecylia Konopka from Children's Memorial Health Institute in Warsaw for their help with data collection and providing medical diagnoses.

This project was partially supported by the Research Center of PJIIT, supported by the Polish Ministry of Science and Higher Education. It is also based upon work supported by the National Science Foundation under Grant No. OISE-0730065.

References

1. Agrawal, R., Imielinski, T., Swami, A.: Mining Association Rules Between Sets of Items in Large Database, Proceedings of ACM SIGMOD International Conference on Management of Data, 1993, 207–216
2. Bazan, J., Szczuka, M.: The rough set exploration system. In: Transactions on Rough Sets III, 2005, 25–42
3. Gubrynowicz, R.; Chojnacka-Wądołowska, D., Konopka, C.: Assessment of Velum Malfunction in Children Through Simultaneous Nasal and Oral Acoustic Signals Measurements. Archives of Acoustics 32(1), 2007, 165–175
4. Hajja, A., Wieczorkowska, A., Raś, Z. W., Gubrynowicz, R.: Object-driven Action Rules and their Application to Hypernasality Treatment. In: Proceedings of ECML-PKDD Workshop on New Frontier in Mining Complex Patterns, Bristol, UK, September 24-28, 2012, 104-115
5. He, Z., Xu, X., Deng, S., Ma, R.: Mining Action Rules from Scratch. In: Expert Systems with Applications, Vol. 29, No. 3, Elsevier, 2005, 691–699
6. Raś, Z.W., Wieczorkowska, A. (Eds): Advances in Music Information Retrieval, Studies in Computational Intelligence, Vol. 274, Springer, 2010
7. Paul, R., Hoque, A.S.: Mining Irregular Association Rules Based on Action and Non-action Type Data, Proceedings of the Fifth International Conference on Digital Information Management (ICDIM), 2010, 63–68
8. Pawlak, Z.: Information systems - theoretical foundations. Information Systems Journal 6, 1981, 205–218
9. Raś, Z., Dardzińska, A.: From Data to Classification Rules and Actions. International Journal of Intelligent Systems 26(6), Wiley, 2011, 572–590
10. Raś, Z.W., Dardzińska, A., Tsay, L.S., Wasyluk, H.: Association action rules. In: IEEE International Conference on Data Mining Workshops, 2008, 283–290
11. Raś, Z.W., Wieczorkowska, A.: Action-Rules: How to increase profit of a company. In: Zighed, D.A., Komorowski, J., Zytkow, J. (eds.) Proceedings of PKDD 2000, LNAI, Vol. 1910, Springer, 2000, 587–592
12. Raś, Z.W., Wyrzykowska, E., Wasyluk, H.: ARAS: Action Rules discovery based on Agglomerative Strategy. In: Mining Complex Data, Post-Proceedings of 2007 ECML/PKDD Workshop (MCD 2007), LNAI, Vol. 4944, Springer, 2008, 196–208
13. Rauch, J., Simunek, M.: Action Rules and the GUHA Method: Preliminary Considerations and Results, Proceedings of ISMIS 2009, LNAI 5722, Springer, 2009, 76–87
14. SAMPA - computer readable phonetic alphabet, `http://www.phon.ucl.ac.uk/home/sampa/`
15. Tsay, L.-S, Raś, Z.W.: Action rules discovery system DEAR3. In: Foundations of Intelligent Systems, Proceedings of ISMIS, Springer, LNAI, Vol. 4203, 2006, 483-492
16. Wasyluk, H., Raś, Z.W., Wyrzykowska, E.: Application of Action Rules to HEPAR Clinical Decision Support System. Experimental and Clinical Hepatology 4(2), 2008, 46–48
17. Zhang, X., Raś, Z.W., Jastreboff, P.J., Thompson, P.L.: From Tinnitus Data to Action Rules and Tinnitus Treatment. In: Proceedings of 2010 IEEE Conference on Granular Computing, Silicon Valley, CA, IEEE Computer Society, 2010, 620–625