Object-driven Action Rules and their Application to Hypernasality Treatment

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, Dept. of Multimedia, 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 [6] have been applied so far in various domains, including banking, music, and medicine. In this paper, medical data are also taken into account, but we propose extraction of rules for data representing changes of medical parameters over time. The extracted rules are called object-driven action rules. The database used in experiments describes observations of hypernasality treatment in children, and our goal is to search for actions which can be taken to move current diagnostic state of the patients to less severe one. The algorithm of rules extraction is presented, and the obtained results are discussed in this paper.

Keywords: action rules, temporal data, hypernasality

1 Introduction and Background

Starting from the introduction of action rules in [6], they have been applied to a variety of data [5, 7, 9, 11]. The goal of extracting action rules is to obtain hints how to reclassify data from one category to another, more desirable, through performing changes of attribute values according to action rules. This way action rules can help to increase the profit of a company, through certain actions triggering reclassification of clients from less to more profitable category, or they can help to improve the health of the treated patients, through certain actions triggering their reclassification to less severe disease categories [6, 9]. Strategies for discovering action rules can be divided into two types: rule-based and object-based. Rule-based approach to action rules discovery consists of two main steps: (1) a standard learning method is used to detect interesting patterns in the form of classification rules, association rules, or clusters; (2) the second step is to use an automatic or semi-automatic strategy to inspect such results and derive possible action strategies. Object-based approach assumes that action rules are extracted directly from a database [4, 10, 12].

In this paper, we deal with complex medical data representing short time series, i.e. several observations of various medical parameters, changing over time, for numerous patients. The data describe medical tests collected through years 1999–2010 in the Childrens Memorial Health Institute in Warsaw, Poland, for patients diagnosed with various cases of nasality [2]. This dataset shows how the condition of the patients has changed during treatment, and the examination of each patient was done at least twice, up to 11 times. In our proposed methodology, a short time series for each patient is treated as a sequence of ordered pairs, where each pair represents changes between 2 examinations, the second one any time later after the first one (not necessarily directly after - there can be any number of examinations between these observations). Therefore, this methodology can be applied to series of any length.

The strategy presented in this paper, called object-driven approach to action rules discovery, is introduced for the purpose of handling temporal data more efficiently and it is a modification of the object-based approach.

2 Object-driven Action Rules

In this section, we introduce the notion of object-driven action rules, which are extracted from temporal series of object's properties. The proposed theory is then applied to hypernasality data, where objects represent patients (Section 3).

2.1 Action Rules

Action rules are rules extracted from an information system that describe the required change of values for a particular attribute(s), which will cause one feature (attribute), called the decision attribute, to change from one state to another. Let $S = (X, A, V)$ denotes an information system [3], where:

1. X is a nonempty, finite set of instances x,
2. A is a nonempty, finite set of attributes a;
   $a : X \to V_a$ is a function for any $a \in A$, where V_a is called the domain of a,
3. $V = \bigcup\{V_a : a \in A\}$.

Furthermore, we assume that $A = A_{St} \cup A_{Fl}$, where A_{St} is the set of *stable* attributes, and A_{Fl} is the set of *flexible* attributes. By *stable* attributes, we refer to the attributes that we have no control over; their values cannot be changed by the user of our system. An example of a *stable* attribute is the age of the patient. 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.

An *atomic action set* is an expression that defines a change of state for a distinct attribute. For example, $(a, a_1 \to a_2)$ is an atomic action set which defines a change of the value of the attribute a from a_1 to a_2, where $a_1, a_2 \in V_a$. 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 remained b_1, where $b_1 \in V_b$.

Action sets are defined as the smallest collection of sets such that:

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.

The *domain* of an action set t is the set of attributes of all the 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, $r = [[(a, a_1 \rightarrow a_2) \wedge (b, b_2)] \Rightarrow (d, d_1 \rightarrow d_2)]$ means that by changing the state of attribute a from a_1 to a_2, and by keeping the state of attribute b as b_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*.

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

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

To extract action rules, we use the algorithm described in [4]. The algorithm starts with 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 value of 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* [4] 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.

Starting from the atomic action set, we generate all frequent atomic action sets, which we will call k-element frequent action sets, where $k = 1$ in the case of atomic action sets. Then we iteratively undertake the following two steps:

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

We keep iterating these steps until no new frequent action sets are generated.

From the final set of frequent action sets, we can construct action rules that are guaranteed to have support greater than or equal to λ_1.

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

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

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 in the decision attribute, e.g. in changing the decision attribute d from state d_1 to d_2. Therefore, we will only consider r_2.

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.

2.3 Temporal Constraint

In this section, we will start with a brief description of our medical dataset, then we will illustrate the modifications applied to our action rules extraction algorithm, while communicating the rationale behind it.

Our dataset is of temporal type and it contains information about patients' visits to the Children's Memorial Health Institute in Warsaw, Poland. The number of visits for each patient is ranging from 2 to 11. Patients are seen as objects represented in our dataset by minimum two and maximum eleven instances.

In a typical scenario of action rules extraction, we have no additional information about instances in a dataset besides values of their attributes. However, in the case of our medical dataset, we also assume that:

1. For each instance, we have a unique patient ID, which is utilized to extract object-driven action rules, a term we will define in Subsection 2.4.
2. We also know that the visits are ordered for each patient, so for example, the $(y)^{th}$ visit for a specific patient, where $y > 1$, has occurred immediately after the $(y-1)^{th}$ visit; this information will allow us to add an ordered pairing restriction which we will call the temporal constraint.

The strategy of action rules construction, presented in the previous section, will be slightly modified to take into account the temporal nature of our dataset. This way, we believe, more refined action rules will be built leading to their higher accuracy and better generalization property.

Temporal Constraint: Here, we make the assumption that the only valid change of attribute value, is the change that happens between two instances of the same object. Accordingly, the standard interpretation N_s^{TC} that complies with the temporal constraint of an action set in $S = (X, A, V)$ is redefined as follows:

1. If $(a, a_1 \rightarrow a_2)$ is an atomic action set, then
$$N_s^{TC}((a, a_1 \rightarrow a_2)) = [\{x_1 \in X : a(x_1) = a_1\}, \{x_2 \in X : a(x_2) = a_2\}] ,$$
where $\forall\, x_1 \,\exists\, x_2$ such that instance x_2 is after x_1.
2. If $t_1 = (a, a_1 \rightarrow a_2) \wedge t$ and $N_s^{TC}(t) = [Y_1, Y_2]$, then
$$N_s^{TC}(t_1) = [Y_1 \cap \{x_1 \in X : a(x_1) = a_1\}, Y_2 \cap \{x_2 \in X : a(x_2) = a_2\}] ,$$
where $\forall\, x_1 \,\exists\, x_2$ such that instance x_2 is after x_1.

The definition of support of an action set and the definitions of support and confidence of an action rule are all the same as in the previous subsection.

2.4 Object-driven Action Rules

Let **O** be the set of object ID's which instances belong to X. We define object-driven action rules to be rules that are extracted from a subsystem $S_p = (X_p, A, V)$ of S, where X_p contains all instances in X of the object p, $p \in \mathbf{O}$.

Table 1. Information System S. Each one of the two employees was observed four different times

	ObjectID	*Observation*	*Loyalty*	*Income*	*Children*
x_0	1	1	High	High	More than 3
x_1	1	2	High	High	More than 3
x_2	1	3	Low	Medium	More than 3
x_3	1	4	Low	Medium	More than 3
x_4	2	1	High	Medium	Less than or equal to 3
x_5	2	2	High	Medium	Less than or equal to 3
x_6	2	3	Low	Low	Less than or equal to 3
x_7	2	4	Low	Low	Less than or equal to 3

Here we provide an example to demonstrate the advantages of object-driven action rules extraction, using the information system S shown in Table 1. As a company, we are interested in extracting action rules that change the state of attribute *Loyalty* from *Low* to *High*. We have two objects with 4 instances each.

Two action rules will be extracted from the system:

1. $r_1 = [[(Income, Low \to Medium) \wedge (Children, \leq 3)] \Rightarrow (Loyalty, Low \to High)]$; $conf(r_1) = 100\%$.
2. $r_2 = [[(Income, Medium \to High) \wedge (Children, > 3)] \Rightarrow (Loyalty, Low \to High)]$; $conf(r_2) = 100\%$.

Now let us assume that the attribute *Children* is missing in S. The following action rules will be extracted:

1. $\wp_1 = [(Income, Low \to Medium) \Rightarrow (Loyalty, Low \to High)]$;
 $conf(\wp_1) = 50\%$.
2. $\wp_2 = [(Income, Low \to High) \Rightarrow (Loyalty, Low \to High)]$;
 $conf(\wp_2) = 100\%$.
3. $\wp_3 = (Income, Medium) \Rightarrow (Loyalty, Low \to High)$;
 $conf(\wp_3) = 25\%$.
4. $\wp_4 = [(Income, Medium \to High) \Rightarrow (Loyalty, Low \to High)]$;
 $conf(\wp_4) = 50\%$.

We can observe that the rules $\wp_1, \wp_4$ are weaker than r_1, r_2 and also the condition and decision part in both rules $\wp_2$ and $\wp_3$ are referring to different objects, so we should not consider them valid. By using object-driven action rules extraction, we would not get $\wp_2$ and $\wp_3$.

Coming back to the action sets, we define the p^{th} standard interpretation $N_{s(p)}$, where p is the object's unique ID, of action sets in $S = (X, A, V)$ as follows:

1. If $(a, a_1 \to a_2)$ is an atomic action set, then
$$N_{s(p)}((a, a_1 \to a_2)) = [\{x \in X_p : a(x) = a_1\}, \{x \in X_p : a(x) = a_2\}] \ ,$$
 where X_p is the set of all instances of the p^{th} object.
2. If $t_1 = (a, a_1 \to a_2) \wedge t$ and $N_{s(p)}(t) = [Y_1, Y_2]$, then
$$N_{s(p)}(t_1) = [Y_1 \cap \{x \in X_p : a(x) = a_1\}, Y_2 \cap \{x \in X_p : a(x) = a_2\}] \ ,$$
 where X_p is the set of all instances of the p^{th} object.

Following both the definition of the **T**emporal **C**onstraint standard interpretation N_s^{TC} and the p^{th} standard interpretation $N_{s(p)}$, it becomes apparent that the definition of the p^{th} standard interpretation that complies with the Temporal Constraint, $N_{s(p)}^{TC}$, where p is the object's unique ID, of action sets in $S = (X, A, V)$ can be defined as follows:

1. If $(a, a_1 \to a_2)$ is an atomic action set, then
$$N_{s(p)}^{TC}((a, a_1 \to a_2)) = [\{x_1 \in X_p : a(x_1) = a_1\}, \{x_2 \in X_p : a(x_2) = a_2\}] \ ,$$
 where $\forall x_1 \, \exists x_2$ such that x_2 is after x_1, and X_p is the set of all instances for the p^{th} object.

2. If $t_1 = (a, a_1 \rightarrow a_2) \wedge t$ and $N^{TC}_{s(p)}(t) = [Y_1, Y_2]$, then

$$N^{TC}_{s(p)}(t_1) = [Y_1 \cap \{x_1 \in X_p : a(x_1) = a_a\}, Y_2 \cap \{x_2 \in X_p : a(x_2) = a_2\}] ,$$

where $\forall\, x_1 \,\exists\, x_2$ such that x_2 is after x_1, and X_p is the set of all instances for the p^{th} object.

If t is an action set and $N^{TC}_{s(p)}(t) = \{Y_1, Y_2\}$, then the support of t in S is defined as: $supp^{TC}_p(t) = min\{card(Y_1), card(Y_2)\}$.

Let $r = [t_1 \Rightarrow t_2]$ be an action rule, where $N^{TC}_{s(p)}(t_1) = [Y_1, Y_2], N^{TC}_{s(p)}(t_2) = [Z_1, Z_2]$. The p^{th} support $supp^{TC}_p(r)$ and the p^{th} confidence $conf^{TC}_p(r)$ of r are defined as follows:

$$supp^{TC}_p(r) = min\{card(Y_1 \cap Z_1), card(Y_2 \cap Z_2)\},$$
$$conf^{TC}_p(r) = \left[\frac{card(Y_1 \cap Z_1)}{card(Y_1)}\right] \cdot \left[\frac{card(Y_2 \cap Z_2)}{card(Y_2)}\right].$$

Now, after all objects-driven action rules are extracted and their p^{th} support and p^{th} confidence are computed for each $p \in \mathbf{O}$, we then calculate their total support $supp^{TC}_{\mathbf{O}}(r)$ (called support) and total confidence $conf^{TC}_{\mathbf{O}}(r)$ (called confidence) following the definitions below:

$$supp^{TC}_{\mathbf{O}}(r) = \sum_{p \in \mathbf{O}} supp^{TC}_p(r),$$
$$conf^{TC}_{\mathbf{O}}(r) = \sum_{p \in \mathbf{O}} \left(\frac{supp^{TC}_p(r) \cdot conf^{TC}_p(r)}{supp^{TC}_{\mathbf{O}}(r)}\right).$$

3 Experimental Data: Hypernasality Data Set

Distortions of the velopharyngeal closure, resulting in speech hypernasality or hyponasality, may cause speech disorders in children [2]. 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.

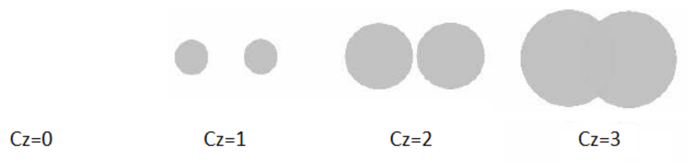

Fig. 1. Czermak's mirror fogging test, rating the degree of the patient's nasal air escape on a 4-point scale: none = 0; small = 1, medium = 2, large = 3 [2].

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, so they are not present in our table. 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, 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 hypertrophy of adenoids and possibly palatine tonsils, and 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/

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

Attribute	Description
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
*diagnosis*2	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]

- vowel of sustained phonation), and *bdg* (high pressure consonants /b, d, g/); SAMPA coding of phonetic alphabet is used [8]. 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 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 if to find actions rules suggesting how to move the diagnosis to less severe category of hypernasality.

The purpose of action rules is to provide hints referring to doctor's intervention, by showing 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 to Hypernasality Data

In addition to the stored attributes shown in Table 2, eight new attributes were derived. For each two visits of a given patient, we calculated the difference and rate of change for *yeaoui*, *i - long*, *bdg*, and *motility* as follows:

1. The difference of values of these attributes for these 2 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 can be equal to the value of bdg for the $(k+1)^{th}$ visit minus the value for the k^{th} visit (but the visits do not have to be consecutive).
2. The rate of change a_2 for these 2 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 these 2 visits.

After calculating the derived attributes, we used the *Rough Set Exploration System* [1] to discretize our real attributes wrt. our decision attribute. Next, our object-driven action rule discovery system 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, then intermediate value was assigned. Therefore, the decision values are {0, 0.5, 1, 1.5, 2, 2.5, 3}.

5 Results and Discussion

Using the attributes described in Table 2, along with 8 new attributes described above, our object-driven action rules extraction system found a number of rules. Few samples are given below.

Rule 1. $r_1 = [(tonsils, 0) \wedge (i - long, \geq 9.5) \wedge (i_2 - long, \geq 5.5 \rightarrow < 5.5)] \Rightarrow (\textit{Czermak's mirror test}, 2 \rightarrow 1.5)$; $supp(r_1) = 2$, $conf(r_1) = 66.7\%$.

Rule 1 means that if our patient has no hypertrophied adenoids nor palatine tonsils (value 0), and *i – long* (nasalization for /i/-long) is more than 9.5, then decreasing the rate of change of *i – long* (*i_2 – long)* from more than or equal to 5.5, to less than 5.5, would improve patient's *Czermak's mirror test* from 2 to 1.5.

Rule 2. $r_2 = [(tonsils, 1) \wedge (i - long, \geq 9.5) \wedge (i_2 - long, \geq 5.5 \rightarrow < 5.5)] \Rightarrow (\textit{Czermak's mirror test}, 2 \rightarrow 1.5)$; $supp(r_2) = 2$, $conf(r_2) = 66.7\%$.

Rule 2 means that if our patient has a bit hypertrophied adenoids and possibly palatine tonsils (value 1), and *i – long* is more than 9.5, then decreasing the rate of change of *i – long* (*i_2 – long)* from more than or equal to 5.5, to less than 5.5, would improve patients *Czermak's mirror test* from 2 to 1.5.

Rule 3. $r_3 = (i - long, \geq 9.5 \rightarrow [2.5, 7.5)) \Rightarrow (Czermak\text{'}s\ mirror\ test, 1 \rightarrow .5)$; $supp(r_3) = 2$, $conf(r_3) = 66.7\%$.

Rule 3 means that if we decrease the value of $i - long$ from greater than 9.5 to [2.5, 7.5), we would improve the patient's *Czermak's mirror test* from 1 to .5. Decreasing of $i - long$, i.e. the decrease of nasalization for /i/-long can be achieved, to some extend, through speech therapy, or (eventually) surgically through a nasopharynx surgery, moving posterior pharyngeal wall forward, towards soft palate. However, none of the examined patients underwent this surgery.

These rules suggest decreasing adenoids and possibly palatine tonsils, and decreasing the nasalization coefficient for /i/-long. This confirms the importance of these attributes, which was expected, but not so obvious in the case of $i - long$. However, decreasing adenoids and possibly palatine tonsils causes only slight change of the Czermak's test, and this is also a confirmation for physicians that surgery might not be particularly beneficiary for patients. Adenoids and tonsils tend to regrow after surgery; at the same time, hypernasality can be treated, to some extent, through speech therapy aiming at decreasing nasalization for /i/-long.

Another interesting observation in our resulted action rules is that we may desire different transitions for the same attribute, depending on the current value of the *Czermak's mirror test*. For example, if our patient's *Czermak's mirror test* is 2, we would want for $i - long$ to stay greater than 9.5 while changing $i_2 - long$. However, if our patient's *Czermak's mirror test* is 1, we would want to decrease the value of $i - long$ from greater than 9.5 to [2.5 7.5).

Rule 4. $r_4 = [(tonsils, < 2) \wedge (i_2 - long, < 5.5) \wedge (motility, [4.5, 5.5)) \wedge (i - long, [1.5, 2.5) \rightarrow < 1.5)] \Rightarrow (Czermak\text{'}s\ mirror\ test, 0.5 \rightarrow 0)$; $supp(r_4) = 3$, $conf(r_4) = 60\%$.

Rule 5. $r_5 = [(tonsils, < 2) \wedge (bdg_1, < 6.5) \wedge (motility, [4.5, 5.5)) \wedge (i - long, [1.5, 2.5) \rightarrow < 1.5)] \Rightarrow (Czermak\text{'}s\ mirror\ test, 0.5 \rightarrow 0)$; $supp(r_5) = 3$, $conf(r_5) = 60\%$.

The above rules mean that very slight hypernasality might be completely removed by decreasing the nasalization coefficient for /i/-long, provided the indicated accompanying conditions are fulfilled.

Rule 6. $r_6 = [(sleep\ apnoea, < 2 \rightarrow > 2) \wedge (bdg, > 8.5 \rightarrow [6.5, 8.5)] \Rightarrow (Czermak\text{'}s\ mirror\ test, 1.5 \rightarrow 0)$; $supp(r_6) = 2$, $conf(r_6) = 100\%$.

This rule is interesting because it shows that worsening of sleep apnoea (along with decreasing the nasality of /bdg/) suppresses light-medium hypernasality.

Rule 7. $r_7 = [(tonsils, < 2) \wedge (motility, < 3.5 \rightarrow [4.5, 5.5)) \wedge (difference\ level\ F1 - F2, [5.5, 6.5) \rightarrow < 4.5)] \Rightarrow (Czermak\text{'}s\ mirror\ test, 1.5 \rightarrow 0)$; $supp(r_7) = 2$, $conf(r_7) = 100\%$.

Rules 6 and 7 are remarkable because their confidence is 100%, and they both induce a significant shift in Czermak's test, from 1.5 to 0, i.e. curing light-medium hypernasality completely, through procedures increasing the motility of the soft palate, and decreasing the difference for the first two formants of the vocal tract for /i/-long.

6 Summary and Conclusions

We presented a new object-driven action rules extraction strategy for temporal datasets based on extraction of rules for each object separately and then summing the obtained results. Its application to hypernasality along with a number of action rules generated by our system is presented. Our ultimate goal is to build a recommender system for treatment of nasality-related speech disorders in children with a collection of object-driven action rules as its kernel.

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 medical diagnoses.

This project was partially supported by the Research Center of PJIIT, supported by the Polish Ministry of Science and Higher Education.

References

1. Bazan, J. and Szczuka, M.: The rough set exploration system. In: Transactions on rough sets III, 25–42 (2005)
2. 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), 165–175 (2007)
3. Pawlak, Z.: Information systems - theoretical foundations. Information Systems Journal 6, 205–218 (1981)
4. Raś, Z.W., Dardzińska, A., Tsay, L.S., Wasyluk, H.: Association action rules. In: IEEE International Conference on Data Mining Workshops, 283–290 (2008)
5. Raś, Z., Dardzińska, A.: From Data to Classification Rules and Actions. International Journal of Intelligent Systems 26(6), Wiley, 572–590 (2011)
6. Raś, Z.W., Wieczorkowska, A.: Action-Rules: How to increase profit of a company. In: Zighed, D.A., Komorowski, J., Zytkow, J. (eds.) PKDD 2000. LNAI, vol. 1910, Springer, Heidelberg, 587–592 (2000)
7. 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, Springer, LNAI, vol. 4944, 196–208 (2008)
8. SAMPA - computer readable phonetic alphabet, `http://www.phon.ucl.ac.uk/home/sampa/`
9. Wasyluk, H., Raś, Z.W., Wyrzykowska, E.: Application of Action Rules to HEPAR Clinical Decision Support System. Exp. Clin. Hep. 4(2), (2008)
10. Wang, K., Jiang, Y., Tuzhilin, A.: Mining actionable patterns by role models. In: Proc. 22nd International Conference on Data Engineering, 3-7 April, 16–26 (2006)
11. 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, 620–625 (2010)
12. Zhang, H., Zhao, Y., Cao, L., Zhang, C.: Combined association rule mining. In: Washio, T., Suzuki, E., Ting, K.M., Inokuchi, A. (eds.) PAKDD 2008. LNCS, vol. 5012, Springer, 1069–1074 (2008)