Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Mobile Triage Support System for Pediatric Emergencies

Wojtek Michalowski (wojtek@admin.uottawa.ca)
University of Ottawa
Steven Rubin
Children's Hospital of Eastern Ontario
Roman Slowinski, Szymon Wilk
Poznan Technical University

Abstract

This paper describes the process and methodology of designing and developing a mobile support system to triage abdominal pain in the emergency room of a hospital. Application of rough sets theory and fuzzy measures to data collected at Children's Hospital of Eastern Ontario allows us to identify the most relevant clinical symptoms and signs while evaluating an abdominal pain patient. This information was used to develop a multi-level *clinical algorithm* that forms the reasoning module of a triage support system. We describe a front-end system called *MAT* that is installed on Palm handheld and that can be used to triage a child irrespective of the available information. We present *MAT*'s functions allowing for the electronic data capture and wireless data transfer. Such a system's design and implementation supports triage at a bed side and facilitates fast and reliable patients' record transfer and storage.

1. Introduction

Abdominal pain in childhood is a highly prevalent symptom in modern society caused by a range of organic diseases, psychosocial disturbances and emotional disorders. In many cases, the exact cause is never determined. Many of these young patients when in severe pain end up in hospital emergency departments where medical staff must focus on identifying the very small portion of cases with a serious organic disease that are in need of urgent treatment. For the majority of such patients where the diagnosis of *appendicitis* is in doubt, investigations and repeated assessments conducted by different physicians are time consuming and may be painful. There is empirical evidence that highlights an obvious advantage of the rapid triage of patients with abdominal pain (Fioravanti *et al.* 1998). However, the central difficulty of such triage is accurate initial assessment based on a limited number of clinical symptoms and signs (attributes) that in combination contribute the most to the diagnosis and management.

This paper describes the development of an easy to use and caregiver-friendly mobile abdominal pain triage (*MAT*) computer system to be used by Emergency Room (ER) personnel for the evaluation of patients with abdominal pain. Prior to developing such a system one needs to evaluate and describe in the form of a *clinical algorithm* the diagnostic practice that is applied to triage a child. Results of our earlier research (Michalowski *et al.* 2001; Rubin *et al*. 2000) focused on the evaluation of alternative knowledge discovery methodologies, and confirm that development of a *clinical algorithm* for triage of the child with abdominal pain is feasible. Building on these results, we analyzed a large sample of

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

patients' data in order to develop a multi-level *clinical algorithm* in the form of decision rules. This algorithm constitutes a reasoning module of the *MAT* system implemented on the Palm handheld. A client system installed on the Palm allows the ER personnel (Nurse Practitioner (NP) or physician) to triage the child appropriately, and to gather pertinent patient information. Clinical information gathered for the triage of a patient can be shared easily among various caregivers using wireless communication (IrDA port of the Palm handheld).

The *MAT* system was designed following a general framework developed for the medical aid tools for patient management support (Sutton 1989), but at the same time it significantly expands the suggested scope of these tools by:

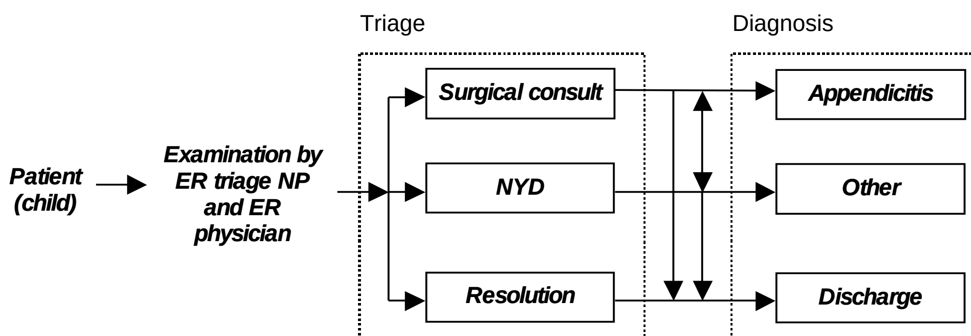
- Supporting early triage;
- Allowing for structured data capture, and
- Providing "triage on demand" at the patient's bed side.

The paper is organized as follows. We start by describing the process of ER management of a child with abdominal pain and data that is collected at that time. The methodology is described in Section 3, and the results are presented in Section 4. A mobile triage system implemented on Palm handheld is described in Section 5. The paper ends with the conclusions.

2. ER Management of Child with Abdominal Pain

The typical process of ER management of a child complaining of abdominal pain is illustrated on Figure 1. This also establishes a framework for use of a mobile triage system for such patients.

Figure 1. ER management of child with abdominal pain



At present, a child is examined first by the ER triage NP to evaluate his/her condition. This assessment is followed by a detailed examination conducted by the ER physician. The possible outcomes of this evaluation are: *resolution*, *surgical consults*, and *not-yet-diagnosed* (*NYD*). *Resolution* indicates that the abdominal pain spontaneously subsides in the absence of a clinical diagnosis. Such a patient can be discharged to the care of his/her regular physician. *Surgical consult* implies that *acute appendicitis* is suspected and a surgeon is called for further evaluation. *NYD* indicates the need for further in-hospital clinical evaluation.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Triage is the first stage in the process of patient management. The final (actual) diagnosis is known when the patient is discharged, and with the patients triaged as *surgical consult*, the final diagnosis is obtained from the post- surgery pathology report.

There is no consensus in the medical literature regarding the most effective management of abdominal pain patients. Several studies advocate use of some form of a scoring system as a clinical support tool (Anatol *et al*. 1995; Hallam *et al*. 1997), while others argue that only an experienced member of the pediatric surgical team is capable of the reliable diagnosis of acute appendicitis (Simpson and Smith 1996). Current practice in teaching hospitals shows an ER triage accuracy of 50-60% achieved by an ER triage NP in the case of diagnosing *appendicitis*. This observation prompted us to focus our attention on a triage stage of the patient management process.

We began our research with the development of the data set. It contains the records of 647 patients with abdominal pain seen during the 1997 – 2000 period in the ER of the Children's Hospital of Eastern Ontario (Ottawa, Ontario). For each patient the 12 attributes (clinical symptoms and signs)[1] and their values as given in Table 1, were extracted from the charts. It should be noticed that not every caregiver was able to conduct all the examinations necessary to obtain the values for every attribute. The ER triage NP is trained to acquire the values of some of the attributes (for example elements of a medical history), while the ER physician is trained to carry out a complete examination.

Most of the considered attributes have nominal values (for example gender or type of pain), some have numerical values (for example temperature or white blood cell count), and some indicate the location of a condition on the patient's abdomen (for example location of the pain or site of tenderness). Values of these latter attributes were collected with the help of special abdomen pictograms, on which the ER physician marked the exact location. Numerical values were discretized according to medical practice, and "location" attributes were assigned a value using an algorithm developed by surgeons and ER physicians for that specific purpose.

[1] These are the attributes recommended in current medical texts for the diagnosis of patients with abdominal pain.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Table 1. Clinical attributes and their domains

Attribute Code	Description	Domain
Age	Number of years	0–6 years, ≥ 7 years
Sex	Gender	Male, female
AbdPainDuration	Length of time pain occurs	≤ 24 hours, 1–7 days, > 7 days
AbdPainSite	The site of maximal pain	RLQ, lower abdomen, other
AbdPainType	Type of maximal pain	Continuous, other
Vomiting	Vomiting occurred	Yes, no
PrevVis	Previous visits to the ER for abdominal pain (irrespective of the site) in the last 48 hours	Yes, no
Tempr	Fever	< 37 °C, 37–39 °C, ≥ 39 °C
AbdTend	Site of maximal tenderness	RLQ, lower abdomen, other
Guarding	Localized muscle sustained contraction noted when palpating the abdomen	Absent, present
LocAbdRebTend	Pain felt at site of maximal tenderness, produced by altering intra-abdominal pressure	Absent, present
WBC	White blood cell count	≤ 4000, 4000–12000, ≥ 12000

According to medical practice, patient records were classified into 3 decision classes (triage outcomes): *resolution*, *surgical consult*, and *NYD*. The detailed partition of a data set is given in Table 2. The data set is unbalanced – 61% of the objects belong to the *resolution* class (and this defines the default accuracy), the *surgical consult* class includes 30% of all records, and the remaining 9% belong to the *NYD*.

Table 2. Partition of patient records into decision classes

Decision class	# of charts	% of charts
Resolution	394	60.9
Surgical consult	195	30.1
NYD	58	9.0
Total	647	100.0

Data was collected as part of a prospective chart study and thus it contains a significant number of missing values. Detailed information about these values is presented in Table 3. Values of only one attribute – Age – are known for all patients. Five attributes have more than 10% of missing values, and three of them (Guarding, WBC, and LocAbdRebTend) have more than 20% of missing values.

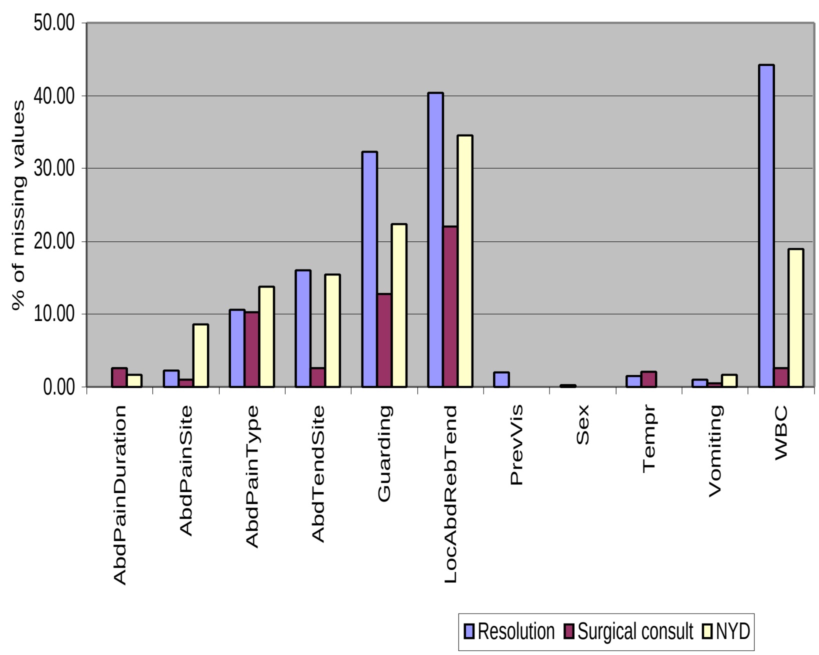
Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Table 3. Missing values of the attributes (in %)

Attribute Code	Decision class		
	Resolution	*Surgical consult*	*NYD*
Age	0.00	0.00	0.00
AbdPainDuration	0.00	2.56	1.72
AbdPainSite	2.28	1.03	8.62
AbdPainType	10.66	10.26	13.79
AbdTendSite	15.99	2.56	15.52
Guarding	32.23	12.82	22.41
LocAbdRebTend	40.36	22.05	34.48
PrevVis	2.03	0.00	0.00
Sex	0.25	0.00	0.00
Tempr	1.52	2.05	0.00
Vomiting	1.02	0.51	1.72
WBC	44.16	2.56	18.97

Figure 2 reveals an interesting pattern of missing values among the decision classes. Except for two attributes (AbdPainDuration, and Tempr), the largest ratio of missing values appears in the *resolution* class, and the smallest in the *surgical consult*. This may be explained by the fact that if, after conducting a few basic examinations the most likely triage is *resolution*, then other symptoms and signs are not checked. Clearly patients triaged as *surgical consult* or *NYD* are examined more thoroughly.

Figure 2. Distribution of the missing values among decision classes



Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

An important issue to be addressed is the treatment of the attributes with a significant number of missing values. Clearly, they should not be discarded (a common approach in other studies is to remove from further analysis the attributes with a number of missing values above a certain threshold), as their importance depends on the context in which they are collected. The WBC attribute is a very good example of such a situation. It has 44% of missing values for the *resolution* class, but less than 3% for the *surgical consult* class. One can conclude that triage of almost half of the patients was so clear that it did not require any further investigations and thus ordering the WBC wasn't necessary. Thus, the WBC in this particular context is not an important attribute. On the other hand, when a patient is triaged as a *surgical consult*, the WBC results are very important for that patient's management[2].

3. Methodology

The data set created from the patients' charts was analyzed for regularities using rough sets theory and fuzzy measures. In this section we describe the basic notions behind the rough sets theory and its extensions (Pawlak 1991; Pawlak and Slowinski 1994, Slowinski 1995), and also Shapley fuzzy measure (Shapley 1953) that was used for additional evaluation of the attributes.

3.1. Rough Sets Theory

For rough sets analysis the data is supplied in the form of an *information table*, in which rows represent objects (patients' charts) and columns represent attributes (clinical symptoms and signs and triage outcomes recorded on the charts). Each cell of the table indicates an evaluation (quantitative or qualitative) of an object represented by the corresponding row by means of an attribute represented by the corresponding column. Formally, an information table is the 4-tuple $S = \langle U,Q,V,f \rangle$, where U is a finite set of objects (universe), Q is a finite set of attributes, $V = \bigcup_{q \in Q} V_q$ and V_q is the domain of the attribute q, f: $U \times Q \rightarrow V$ is a function called *information function* such that $f(x,q) \in V_q$ for each $q \in Q$, $x \in U$. The set Q is divided into a set C of *condition* attributes, and a set D of *decision* attributes (for the data set described in this paper it is a singleton – the triage outcome).

Each object x of U is described by a vector of evaluations (attribute values), called *description* of x in terms of the attributes of Q. This vector represents available information about x. Objects having the identical description are called *indiscernible*. In general, the *indiscernibility relation* on U, denoted by I_p, is associated with every (non-empty) subset of attributes $P \subseteq Q$

$$I_P = \{(x,y) \in U \times U : f_q(x) = f_q(y), \text{ for each } q \in P\} \qquad (3.1.1)$$

Clearly, the relation (3.1.1) is an equivalence relation (reflexive, symmetric and transitive); thus, it partitions set U into equivalence classes called *P-elementary sets.* The family of all the equivalence classes of relation I_P is denoted by $U \mid I_P$ and the equivalence class containing an object $x \in U$ is denoted by $I_p(x)$. If $(x,y) \in I_P$, then objects x and y are *P-indiscernible*.

[2] The 3% of missing values should be attributed to incorrect entry of the WBC information on the patient's chart in the ER.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

The partitions of set U induced by subsets of condition attributes and subsets of decision attributes represent *knowledge* about U.

The key idea of rough sets is related to an approximation of knowledge expressed by decision attributes using knowledge that is expressed by condition attributes. Rough sets theory answers several questions related to such approximation:

(a) Is the data contained in the information table consistent?

(b) What are the non-redundant subsets of condition attributes (*reducts*) ensuring the same quality of approximation as the whole set of condition attributes?

(c) What are the condition attributes (*core*) that cannot be eliminated from the approximation without decreasing the quality of the approximation?

(d) What minimal "*if ..., then ...*" decision rules can be induced from the approximations?

Several important aspects of rough sets theory made it of particular interest to the researchers evaluating real data sets describing various decision situations (Lin and Cercone 1997). With respect to the input information, rough sets allow us to analyze both quantitative and qualitative data, and data inconsistencies need not be removed prior to the analysis, as they are dealt with by separating certain and doubtful knowledge extracted from the information table. With reference to the output information, rough sets allow us to acquire *a posteriori* information regarding the relevance of a particular attribute (or subsets of attributes) for the quality of approximation. Moreover, the final result presented in the form of "*if..., then...*" decision rules that are based on the most relevant attributes, is easy to interpret.

For a demonstration of the basic concepts of rough sets, let us assume that X is a non-empty subset of U, for example an equivalence class with respect to set D of decision attributes, and $P \subseteq C$ is a subset of condition attributes.

We say that object $x \in X$ *belongs certainly* to X if all objects from the P-elementary set $I_p(x)$ also belong to X, i.e. $I_p(x) \subseteq X$. Then, for a given P, information about object x is *consistent* with information about other objects from U.

We say, moreover, that object $x \in U$ *could belong* to X if at least one object from the P-elementary set $I_p(x)$ belongs to X, i.e. $I_p(x) \cap X \neq \emptyset$. If $\emptyset \neq I_p(x) \cap X \neq I_p(x)$ then, for a given P, information about object x is *inconsistent* with information about other objects from $I_p(x)$.

For $P \subseteq C$, the set of all objects belonging certainly to X constitutes the P-*lower approximation* of X, denoted by $\underline{P}(X)$, and the set of all objects that could belong to X constitutes the P-*upper approximation* of X, denoted by $\overline{P}(X)$:

$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\}, \tag{3.1.2}$$

$$\overline{P}(X) = \{x \in U : I_P(x) \cap X \neq \emptyset\}. \tag{3.1.3}$$

The difference between the upper and lower approximations of X is called the P-boundary of X:

$$Bn_P(x) = \overline{P}(X) - \underline{P}(X). \tag{3.1.4}$$

The P-boundary of X is composed of inconsistent objects that belong to X with some ambiguity. The following relations hold: $\underline{P}(X) \subseteq X \subseteq \overline{P}(X)$, $\underline{P}(X) = U - \overline{P}(U - X)$.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

The family of all the sets $X \subseteq U$ having the same lower and upper approximations is called a *rough set*.

The rough approximations of a subset $X \subseteq U$ can be extended to partitions of U; in particular to the partition induced by decision attributes from D. This partition corresponds to the classification of objects into decision classes – the lower approximations of decision classes represent certain knowledge, upper approximations represent possible knowledge, and the boundaries represent doubtful knowledge about the classification, expressed in terms of condition attributes from $P \subseteq C$.

Given a partition of U into decision classes, $\mathbf{Cl}=\{Cl_t, t \in T\}$, $T=\{1,...,n\}$, the P-boundary with respect to k>1 classes $\{Cl_{t1},\dots,Cl_{tk}\} \subseteq \{Cl_1,\dots,Cl_n\}$ is defined as

$$Bd_P(\{Cl_{t1},...,Cl_{tk}\}) = \left(\bigcap_{t=t1,...,tk} Bn_P(Cl_t)\right) \cap \left(\bigcap_{t\neq t1,...,tk} \left(U - Bn_P(Cl_t)\right)\right) \quad (3.1.5)$$

The objects from $Bd_P(\{Cl_{t1},...,Cl_{tk}\})$ can be assigned to one of the classes $Cl_{t1},\dots,Cl_{tk}$, however, P and all its subsets do not provide enough information to do this assignment precisely.

Using the rough approximations of decision classes, it is possible to induce *decision rules* describing the classification represented by examples contained in the information table. These are logical statements (implications) of the type "*if ..., then...*" where the antecedent (condition part) is a conjunction of the elementary conditions concerning particular condition attributes and the consequence (decision part) is a disjunction of possible assignments to particular classes of a partition of U induced by decision attributes. Given a partition $\mathbf{Cl}$ of U, the syntax of the rule is the following:

"*if* $f(x,q_1) = r_{q1}$ *and* $f(x,q_2) = r_{q2}$ *and ...* $f(x,q_p) = r_{qp}$,
then x is assigned to Cl_{t1} *or ...* Cl_{tk}", (3.1.6)

where $\{q_1,...,q_p\} \subseteq C$, $(r_{q1},...,r_{qp}) \in V_{q1}\times...\times V_{qp}$ and $\{Cl_{t1},...,Cl_{tk}\} \subseteq \{Cl_1,...,Cl_n\}$. If the consequence is univocal, i.e. k=1, then the rule is *exact*, otherwise it is *approximate* or *uncertain*.

Let us observe that for any $Cl_t \in \{Cl_1,...,Cl_n\}$ and $P \subseteq C$, the definition (3.1.2) of P-lower approximation of Cl_t can be rewritten as

$$\underline{P}(Cl_t) = \{x \in U: \text{for each } y \in U, \text{ if } yI_Px, \text{ then } y \in Cl_t\} \quad (3.1.2')$$

Thus, the objects belonging to the lower approximation $\underline{P}(Cl_t)$ can be considered as prototypes for the induction of exact decision rules.

Therefore, the statement "*if* $f(x,q_1) = r_{q1}$ *and* $f(x,q_2) = r_{q2}$ *and ...* $f(x,q_p) = r_{qp}$, *then* x is assigned to Cl_t", is accepted as an exact decision rule if and only if there exists at least one object $y \in \underline{P}(Cl_t)$, $P=\{q_1,\dots,q_p\}$, such that $f(y,q_1) = r_{q1}$ and $f(y,q_2) = r_{q2}$ and … $f(y,q_p) = r_{qp}$.

Given $\{Cl_{t1},\dots,Cl_{tk}\} \subseteq \{Cl_1,...,Cl_n\}$ we can write

$$Bd_P(\{Cl_{t1},\dots,Cl_{tk}\}) = \{x \in U: \text{for each } y \in U, \textit{ if } yI_Px, \textit{ then } y \in Cl_{t1} \textit{ or } \dots\ Cl_{tk}\}. \quad (3.1.2'')$$

Thus, the objects belonging to the boundary $Bd_P(\{Cl_{t1},\dots,Cl_{tk}\})$ can be considered as a basis for the induction of approximate decision rules.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

The analysis of large information tables shows that the calculation of approximations according to (3.1.2) and (3.1.3) may result in a large *P*-boundary of *X*. Consequently, it leads to weak decision rules (supported by few objects from lower approximations). In such a case it seems reasonable to relax the conditions for the assignment of objects into lower approximations by allowing it to include some inconsistent objects. This relaxation is called a *variable precision model* (VPM) and is described in (Ziarko 1993). VPM defines lower approximations using a limited number of counterexamples that is controlled by pre-defined level of certainty β (0 < β ≤ 1). In VPM, the *P*-lower approximation of *X* in *U* is defined as:

$$\underline{P}(X) = \left\{ x \in U : \frac{card\left(I_P(x) \cap X\right)}{card(X)} \geq \beta \right\}. \qquad (3.1.7)$$

If β is set to 1, then the VPM model is equivalent to (3.1.2).

Decision rules induced from the lower approximations of decision classes defined by (3.1.7) have univocal consequences (decisions); however, the *confidence index* of each rule (defined as the number of objects matching both the condition and decision part of the rule to the number of objects matching the condition part only) varies from β to 1.

The *accuracy* of the approximation of $X \subseteq U$ by the attributes from *P* is given as the ratio:

$$\alpha_P(X) = \frac{card\left(\underline{P}(X)\right)}{card\left(\overline{P}(X)\right)}. \qquad (3.1.8)$$

The *quality* of the approximation of $X \subseteq U$ by the attributes from *P* is given as the ratio:

$$\gamma_P(X) = \frac{card\left(\underline{P}(X)\right)}{card(X)}. \qquad (3.1.9)$$

where $0 \leq \gamma_P(X) \leq 1$ and the quality represents the relative frequency of the objects correctly classified by means of the attributes from *P*.

The *quality of the approximation of classification* ***Cl*** by set of attributes *P* is given as:

$$\gamma_P(\boldsymbol{Cl}) = \frac{\sum_{i=1}^{n} card\left(\underline{P}(Cl_i)\right)}{card(U)}. \qquad (3.1.10)$$

It is called in short *quality of classification* and specifies the ratio of all *P*-correctly classified objects to all objects in the information table.

Each minimal subset $P \subseteq C$ such that $\gamma_P(\boldsymbol{Cl}) = \gamma_C(\boldsymbol{Cl})$ is called a *reduct* of *S* and denoted by $RED_{\boldsymbol{Cl}}(C)$. An information table can have more than one reduct. The intersection of all reducts is called a *core* and is denoted by $CORE_{\boldsymbol{Cl}}(C)$. The core is composed of indispensable attributes that cannot be removed from the information table without decreasing the quality of classification. Condition attributes that do not belong to any reduct are called superfluous.

3.2. Missing Values

Classical rough sets theory is not well suited for dealing with missing values in data (empty cells in the condition portion of the information table) so they must be converted by:

a) Replacing the missing value with a special value (for example, N/A) and then treating it as a known value.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

b) Replacing the missing value with a known one (for example, with the average or most frequent value of the corresponding condition attribute in a whole data set or in a given decision class).

Both approaches have serious shortcomings. Using the former one, it is possible to obtain decision rules with conditions based on missing values that are represented by a selected special value. Such rules are difficult to interpret (for example, it is unreasonable to interpret a rule that in a condition part has an attribute with N/A value). The latter approach may falsify the data (by assigning for example, the most frequent value without any sound justification), especially when the number of missing values is large.

To address these shortcomings, an extension to rough sets theory was proposed (Greco *et al*. 2000), and it is briefly discussed here.

The definition of the information table (S = <U, Q, V, f>) is extended by assuming that the set V is augmented to include the missing value (indicated by “*”).

Instead of the indiscernibility relation I_P, a new type of relation, denoted by I_P^* is introduced. For each object $x, y \quad U$ and for each subset of attributes $P \quad Q$, y I_P^* x means that $f(x, q)$ = f(y, q), or f(x, q) = , or f(y, q)= , for every $q \quad P$. Let $I_P^*(x) = \{y \quad U\!: y \; I_P^* \; x\}$ for each $x \quad U$ and for each $P \quad Q$. I_P^* is a reflexive and symmetric but not transitive binary relation. Finally, let $U_P^* = \{x \quad U\!: f(x, q) \neq *$ for at least one $q \quad P\}$.

Using I_P^* the definitions of the P-lower and P-upper approximation of X become:

$$\underline{I}_P^*(X) = \{x \quad U_P^* : I_P^*(x) \quad X\}, \tag{3.2.1}$$

$$\overline{I}_P^*(X) = \{x \quad U_P^* : I_P^*(x) \cap X \neq \quad \}. \tag{3.2.2}$$

P-lower approximation can be also calculated using the VPM model as

$$\underline{I}_P^*(X) = \left\{ x \quad U_P^* : \frac{card\left(I_P^*(x)\right)}{card(X)} \geq \beta \right\} \tag{3.2.3}$$

The approximations defined in (3.2.1) and (3.2.2) are further used to calculate the P-boundary of X, accuracy of approximation of X, and the quality of the approximation of X.

Given the partition ***Cl*** of U, one can calculate the quality of the classification of ***Cl*** and use this measure to find *reducts* and *core* of attributes.

Using the rough approximations (3.2.1) and (3.2.2), it is possible to induce a generalized description of the examples contained in the information table in terms of *decision rules* (see 3.1).

Since each decision rule (3.1.6) is an implication (similar to (3.1.1’) and (3.1.2’) for I_P^*), a *minimal* decision rule represents a unique implication in the sense that there is no other implication having a subset of elementary conditions and the same consequent.

We say that $y \quad U$ *supports* the exact decision rule "*if* $f(x,q_1) = r_{q1}$ *and* $f(x,q_2) = r_{q2}$ *and* ... $f(x,q_p) = r_{qp}$, *then* x is assigned to Cl_j", if [$f(y,q_1) = r_{q1}$ and/or $f(y,q_1)$ =] and [$f(y,q_2) = r_{q2}$ and/or $f(y,q_2)$ =] ... and [$f(y,q_p) = r_{qp}$ and/or $f(y,q_p)$ =] and $y \quad Cl_t$.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Similarly, we say that y U *supports* the approximate decision rule "*if* $f(x,q_1) = r_{q1}$ *and* $f(x,q_2) = r_{q2}$ *and ...* $f(x,q_p) = r_{qp}$, *then* x is assigned to Cl_{t1} *or ...* Cl_{tk}", if [$f(y,q_1) = r_{q1}$ and/or $f(y,q_1) =$] and [$f(y,q_2) = r_{q2}$ and/or $f(y,q_2) =$] ... and [$f(y,q_p) = r_{qp}$ and/or $f(y,q_p) =$] and y $Bd_C^*(\{Cl_{t1},\ldots,Cl_{tk}\})$.

Decision rules induced from lower approximations defined by (3.2.3) have univocal consequences (decisions); however, the confidence index of each rule (see 3.1) varies from β to 1.

3.3. Fuzzy Measure

The quality of classification (calculated using either I_P or I_P^*) satisfies the properties of the set functions called *fuzzy measures*. Such measures can be used for modeling the importance of coalitions (Grabisch 1997), or as proposed in (Greco *et al*. 1998, 2001) to assess the relative value of the information supplied by each attribute and to analyze the interactions among the attributes (using the quality of classification calculated according to (3.1.9)). Let us explain this point in greater detail.

Let $C=\{q_1,...,q_n\}$ be a finite set, whose elements could be the players in a game, condition attributes in an information table, different criteria in a multicriteria decision problem, *etc*. Let $PS(C)$ denote the power set of C, i.e. the set of all subsets of N. A *fuzzy measure* on C is a set function μ: $PS(C)\rightarrow[0,1]$ satisfying the following axioms:

a) $\mu(\;\;)=0$, $\mu(N) \leq 1$,

b) A B implies $\mu(A)\leq\mu(B)$, for all A,B $PS(C)$.

Within game theory, the function $\mu(A)$ is called the characteristic function and represents the payoff obtained by the coalition A C in a cooperative game (Shapley 1953); in a multi-attribute classification $\mu(A)$ can be interpreted as the conjoint importance of the attributes from A C.

In game theory some indices were proposed as specific solutions of cooperative games. The most important of those is the *Shapley value* (Shapley 1953), defined for every element q_i C as:

$$\phi_S(q_i) = \sum_{K \quad C-\{q_i\}} \frac{(n-|K|-1)!\,|K|!}{n!}[\mu(K \quad \{q_i\})-\mu(K)]. \qquad (3.3.1)$$

The Shapley value can be interpreted as an average contribution of the element q_i to all the possible coalitions (combinations) of the elements from C.

For ***Cl*** being a partition of U, $\mu(K) = \gamma_K(\boldsymbol{Cl})$ for every K C; the value of $\mu(C)$ is shared among the elements of C, i.e. $\sum_{i=1}^{n}\phi_S(q_i) = \gamma_C(\boldsymbol{Cl})$.

Thus, the Shapley value $\phi_S(q_i)$ can be used to assess the contribution of a single attribute q_i to the quality of a classification. Those attributes with higher value of $\phi_S(q_i)$ are considered to explain better relationships in a data set.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

4. Results

The medical data set described in section 2 was analyzed using ROSE software (Predki and Wilk 1999). Due to the large number of missing values, an extended rough sets methodology was used to:

- Find the clinical symptoms and signs that are the most relevant for classifying a patient as *resolution*, *surgical consult*, or *NYD*.
- Induce the set of decision rules based on the attributes selected in the previous step that ensure a high classification accuracy of patients in the ER.

Initial analysis identified one reduct containing all the condition attributes, and thus we used the fuzzy measure (3.3.1) to establish a partial ordering of the attributes.

Table 4 gives the Shapley values for all condition attributes. The table also gives the percentages of missing values for each attribute, and the quality of classification associated with the set of attributes containing attributes starting from the "Tempr", up to the attribute in a specific row of Table 4. For example, the quality of classification of 0.094 for the AbdPainSite was achieved for the set of attributes containing Tempr, AbdPainDuration, Sex, Vomiting, and AbdPainSite.

Table 4. The attributes sorted in a descending order according to Shapley values

Attribute Code	Shapley value	Quality of classification	% of missing values
Tempr	0.079	0.000	1.5
AbdPainDuration	0.072	0.002	0.9
Sex	0.069	0.002	0.2
Vomiting	0.067	0.006	0.9
AbdPainSite	0.066	0.094	2.5
PrevVis	0.065	0.196	1.2
AbdTendSite	0.053	0.280	11.9
AbdPainType	0.046	0.337	10.8
Age	0.045	0.433	0.0
Guarding	0.031	0.539	25.5
WBC	0.030	0.601	29.4
LocAbdRebTend	0.017	0.640	32.3

It is worthwhile to note that the order of the attributes in Table 4 closely follows the order of the percentage of missing values, especially for those that have more than 10% of missing data. It is consistent with an intuitive observation, that the less information the attribute bears, the less relevant it is according to the Shapley value ranking[3]. The only exception is "Age", which has no missing values, but its low importance may be explained by the very general discretization (see Table 1) that most likely entails the loss of information.

The choice of the most relevant attributes on the basis of the information given in Table 4 is not obvious; thus after consulting with the ER physicians we decided to use several thresholds, beginning with the set of the top five attributes (Tempr, AbdPainDuration, Sex,

[3] The low rank of the WBC attribute can be explained by the fact that the original data set is not well balanced, with the *resolution* class being the dominant one. Therefore, the large number of missing WBC values for that class (the blood test was simply not administered) had an overwhelming impact on the WBC rank.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Vomiting, and AbdPainSite), for which a considerable increase in the quality of classification was observed. Then we iteratively enlarged this set with the remaining attributes according to their Shapley value ranking. The LocAbdRebTend was appended last, so we finished with the set containing all 12 condition attributes.

For each set of attributes we tested the classification accuracy of the corresponding decision rules. The classification accuracy was estimated using 10-fold cross-validation tests (Mitchell 1997). In order to get more reliable results, the validation tests were repeated five times and their results were averaged over all repetitions. The decision rules were induced using the *Explore* algorithm (Stefanowski and Vanderpooten 2001) that generates the set of *satisfactory rules* that meet pre-defined requirements (maximum number of elementary conditions in a rule (rule length), the minimum rule strength, and the minimum confidence index). In our analysis the length of the generated rules was not restricted, the minimum confidence index was set to 0.8, and the minimum strength was later modified during the tests. This modification procedure is described in details in (Stefanowski and Wilk 2001). Running *Explore* with a minimum confidence index of 0.8 is equivalent to inducing rules from the lower approximations, calculated in the VPM with β coefficient equal to the confidence index.

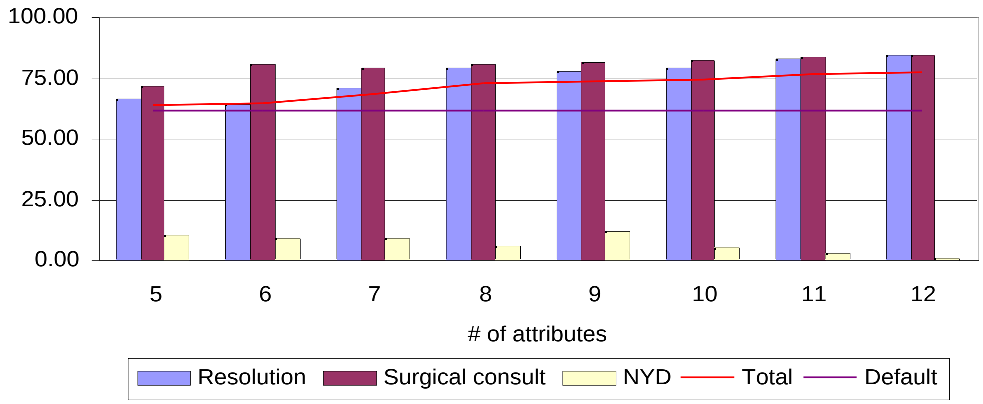
Handling of missing values required some changes in the rule induction algorithm. Firstly, we assumed that the rule condition will be satisfied by value equal to the value in a condition or by a missing value (for example, condition Vomiting = 'yes' will be met by the value 'yes' and also by a missing value). Secondly, the algorithm was modified to ensure robustness of generated rule so it covers at least one object that has a non-missing (known) value for the condition attributes. The classification accuracy resulting from the 10-fold cross validation tests is reported in Table 5 and illustrated in Figure 3.

Table 5**.** Accuracy of classification

# of attributes in a set	Accuracy of classification [%]			
	Total	*Resolution*	*Surgical consult*	*NYD*
5	63.46	66.57	71.48	10.67
6	64.01	64.05	80.56	8.73
7	67.91	71.00	79.30	8.60
8	72.76	78.75	80.35	6.07
9	72.95	77.67	81.47	12.07
10	73.64	79.45	82.13	5.46
11	76.01	83.13	83.25	3.13
12	76.88	84.57	83.99	0.67

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Figure 3**.** Accuracy of classification



The rules induced by the *Explore* algorithm gave the best and most consistent accuracy for the *surgical consult* class (see Figure 3). This suggests that it is possible to generate "accurate" rules (i.e. ensuring high classification accuracy for the patients that require *surgical consult*) using smaller subsets of condition attributes. The accuracy for the *resolution* decision class increases as the set of considered attributes increases. This may suggest that additional information is helpful when making the discharge decision. The accuracy of classification for the *NYD* patients is not acceptable. The detailed analysis of cross-validation test results revealed that in most cases the *NYD* patients are classified to the *resolution* category. This type of misclassification is not desirable as it may result in a potentially sick patient being discharged from the ER. The misclassifications occurring between *resolution* and *NYD* classes suggest that either ER physicians rely on information not recorded in the charts while differentiating between the *resolution* and *NYD* patients, or that it is impossible to distinguish patients from these two classes with acceptable accuracy, and the ER medical staff decides to keep all questionable patients in hospital for observation. Our experience with the patients belonging to the *NYD* class is confirmed by other medical studies (Williams *et al*. 1998) where it was reported that differentiation between appendicitis and non-specific abdominal pain (one of the most common complaints among the *NYD* patients) is very difficult on the basis of regularly evaluated symptoms and signs.

Information about the number of rules generated by the *Explore* algorithm for the sets of attributes of different size is given in Table 6.

Table 6. Average number of rules generated in 10-fold cross validation tests

# of attributes in a set	# of rules
5	5.6
6	7.3
7	13.2
8	19.5
9	22.0
10	38.8
11	68.5
12	110.6

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Analysis of the results followed by the consultations with the pediatric surgeons resulted in the selection of three sets of attributes – containing 5, 8 and 11 attributes respectively – to be used for development of the *clinical algorithm* implemented in the *MAT* system. These rules are to be used at various stages of child management in the ER, depending on the amount of available information (thus, the rules created for 5 attributes could be used to suggest the initial triage shortly after the patient is admitted to the ER, while the rules generated for 11 attributes can be applied when the necessary examinations have been concluded).

For illustrative purposes in Tables 7 and 8 we present the rules generated for the sets of 5 and 8 attributes respectively. The relative strength of a rule explains that rule's coverage (for example, the first rule in Table 7 having the relative strength of 54.6% is capable of matching 54.6% patients from the *resolution* class).

Table 7. Decision rules generated for 5 attributes

Triage	AbdPainDuration	AbdPainSite	Sex	Tempr	Vomiting	Relative strength [%]
Resolution		other				54.6
Resolution				<37	absent	38.6
Surgical consult	≥7 hours <5 days	RLQ	male		present	28.7
Surgical consult		RLQ	male	<37	present	16.9
NYD	<7 hours			>39		1.7
NYD		other	male	>39	absent	3.4

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Table 8. Decision rules generated for 8 attributes

Triage	AbdPainDuration	AbdPainSite	AbdPainType	AbdTendSite	PrevVis	Sex	Tempr	Vomiting	Relative strength [%]
Resolution		other							54.6
Resolution				other					53.6
Resolution			intermittent				<37		38.8
Resolution			intermittent					absent	37.8
Resolution							<37	absent	38.6
Surgical consult	≥7 hours <5 days	RLQ				male		present	28.7
Surgical consult		RLQ	constant			male		present	33.3
Surgical consult		RLQ		RLQ		male		present	34.4
Surgical consult		RLQ			no	male		present	30.3
Surgical consult		RLQ				male	<37	present	16.9
Surgical consult	≥7 hours <5 days	RLQ	constant		no	male			33.8
Surgical consult		RLQ	constant	RLQ	no	male			41.0
Surgical consult		RLQ	constant	RLQ	no		≥37 ≤39		41.0
Surgical consult		RLQ	constant		no	male	≥37 ≤39		23.1
Surgical consult			constant	RLQ	no	male	≥37 ≤39		24.6
Surgical consult			constant	RLQ	no	male		present	31.3
Surgical consult				RLQ	no	male	≥37 ≤39	present	17.4
NYD		other		other		male	>39		6.9
NYD		other				male	>39	absent	3.4
NYD				other		male	>39	absent	3.4
NYD	<7 hours		constant	other			≥37 ≤39	present	3.4

5. Mobile Triage System

The management of a patient with abdominal pain in the ER requires an initial assessment of his/her condition, repeatedly evaluating the patient's history, physical findings and laboratory and radiological tests. On the basis of such an assessment and a triage, a final diagnosis is reached and an appropriate management is selected. Thus, any medical management support system needs to have the ability to follow this process and to provide a caregiver with an

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

appropriate level of support at its every stage. It means that such a system should have information gathering facilities, the ability to triage, and finally should "follow a patient" as his/her management is transferred between different caregivers. The *MAT* system discussed here maintains all of these features. It allows for the collection of information about the patient's condition; it uses a three-level *clinical algorithm* (for 5, 8, and 11 attributes) to support the triage irrespective of the amount of available information about a patient. It also provides "triage on demand" by "following the patient" – a feature achieved by its implementation on a mobile device such as a Palm handheld. The system has been developed according to the principles of client-server architecture (Ford, 1998). The client module runs on a Palm handheld with *PalmOS 3.5* (or later), and the server is implemented on a PC running *Windows NT/2000*. Communication among Palm handhelds and between the Palm and the PC is maintained using a wireless infrared (IrDA) port or a cradle adapter respectively. The desired functionality of the *MAT* system is accomplished through a clear division of the tasks to be performed on the server and on the client side.

The Palm handheld client is responsible for the following tasks:

- Collecting patient data.
  The client is used for entering data about examined patients and storing it in a local database (electronic data capture).
- Supporting the triage decision.
  Using information available in a local database, the client applies three-level *clinical algorithm* to triage a patient.
- Synchronizing patient's data with other clients.
  The client is capable of transmitting data from a local database to other clients using a wireless port. It is also capable of receiving data from other clients and storing it in a local database.
- Transferring data to a PC server.
  The client is able to transfer the contents of a local database to a server using a cradle adapter or infrared port.

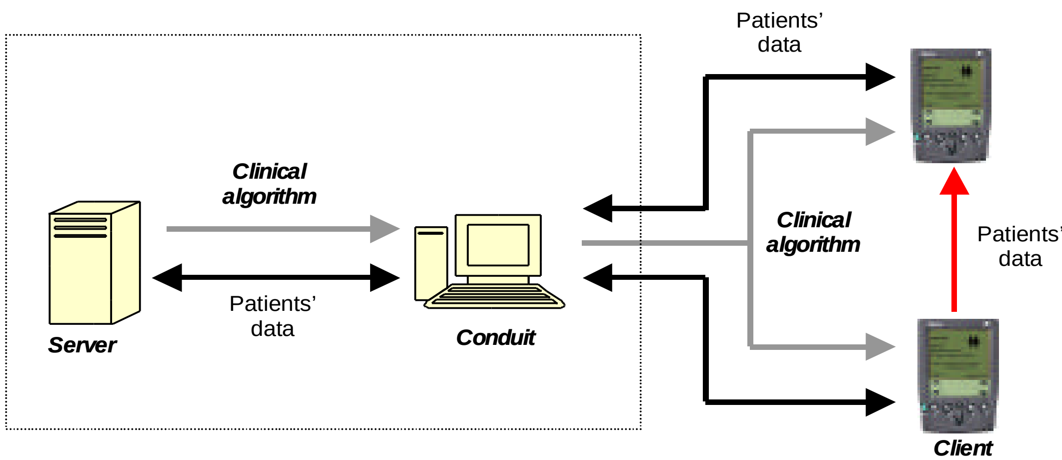
The PC server is responsible for the following tasks:

- Managing and synchronizing a centralized database
  Through merge and update operations, the server manages a centralized database containing all the patients' records collected from the Palm handheld clients.
- Periodic analysis of the data stored in a centralized database
  The server periodically analyzes the data stored in the centralized database to evaluate and possibly calibrate the *clinical algorithm*.
- Updating the *clinical algorithm* used by the Palm handheld clients
  The server automatically updates the *clinical algorithm* residing on the Palm handheld client through a *HotSync* function.

The overall system's architecture is illustrated in Figure 4.

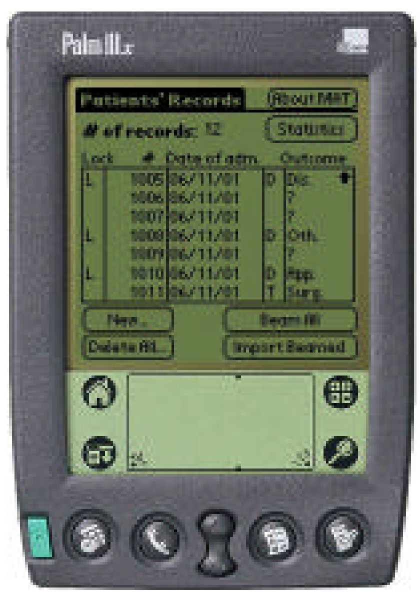
Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Figure 4. System's architecture



Use of the system begins with the entry of the patient's information (including unique PIN) at the moment of the patient's admission to the ER. Initial data entry screen is presented in Figure 5.

Figure 5. Patients' records entry screen



One of the unique features of the system is its adaptability to different end-user needs while capturing the patient's information. For example, the results of the physical examination of the abdomen are entered using the pictograms (see Figure 6). Information about the type of maximal pain is gathered using a pre-defined selection list (see Figure 7), while the information about the duration of the pain is entered using a free-entry format (see Figure 8).

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

There is reported evidence that such structured data collection may contribute to the improved triage and diagnosis of a patient (Korner *et al*. 1998).

Figure 6. Graphical data capture

Figure 7. Pre-defined selection data capture

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Figure 8. Free-entry data capture

The system's triage function can be invoked at any time and it uses the most current data to give a triage recommendation (see Figure 9 for triage advice screen). For most of the clinical attributes the system allows also for the free entry of any additional comments about the patient's condition (see "comments button" in the Figures 6 and 7) as deemed necessary by the attending health care professional.

Figure 9. Invoking triage function

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Depending on the information available to a caregiver, the *MAT* system invokes the most suitable level of the clinical algorithm, that is, a set of rules providing the best overall match. Even if in a given set there are no rules exactly matching the available data, the system will consult the most closely matching rules but diminishing at the same time a confidence associated with such consultation. Presentation of this information is illustrated in Figure 10.

Figure 10. Confidence value of a triage

The triage confidence value is calculated on the basis of rules fired during classification. First, for all invoked rules the matching ratios are calculated (matching ratio is a function of the number of fully matched conditions in the rule, rule's strength, and rule confidence index). Matching ratios are then summed up within each decision class for all possible rules (for a given set of the attributes) to get total matching ratios. Finally, these total ratios are normalized relative to their maximum values. The normalized ratios are presented to end-used as the triage confidence values. A decision class that acquired the highest triage confidence is then presented as a "suggested triage". If for all possible triage outcomes the normalized matching rations are equal to zero, then the *NYD* class is presented as a suggested triage.

Once the care of a patient has been transferred to another caregiver, all information gathered so far (including triage recommendation) can be beamed (wireless transfer) to another Palm handheld. When new information about a patient becomes available, the triage function of the system might be invoked again and *MAT* will use updated information to re-evaluate the latest triage decision. At the end of the process, all the pertinent patient data is transferred to the PC server thus either creating or updating each patient's record in a centralized database.

6. Conclusions

The purpose of the research described in this paper was to develop an easy-to-use and mobile triage system that can support the triage decision for a patient with abdominal pain. The medical data set we analyzed was characterized by the large number of missing values for

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

some of the attributes. The analysis discussed here would not be possible without a hybrid approach that allowed us to consider the differential information content of the attributes, taking into account the number of missing values and their distribution among the triage classes. Once the information content of the attributes was analyzed, it was possible to calculate the fuzzy measure that was used to rank the attributes according to their triage relevance. The generation of decision rules and the development of a multilevel clinical algorithm to triage patients in the ER followed this analysis. Development of the algorithm was the result of consultations with the domain experts, resulting in the identification of three subsets of triage attributes containing 5, 8 and 11 attributes respectively. These sets were used in the generation of decision rules by a methodology enhanced with the ability to handle the missing values. While developing a computer system we implemented a modified way of rules' induction allowing for matching rules with incomplete antecedents.

The *MAT* system that was developed follows the basic requirements expected for the effective support of a patient management process:

- It allows entry of the relevant patient information (history, tests, and physical examination).
- It allows storage and transmission of this information.
- It has the ability to invoke the triage function independently of the stage in the management process.

All the above can be accomplished in a coordinated manner at a patient's bedside. Currently, we are working on a limited clinical testing of the system in a teaching hospital in Ontario. We are also expanding the system's reasoning capabilities to cover the management of a scrotal pain condition in childhood.

Acknowledgement

The research reported in this paper was supported by the grants from the Polish Committee for Scientific Research, the Foundation for Polish Science, and the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank Eric Payne for help in developing a data set from the ER admission charts, and Ann Burgess for editorial assistance.

The *MAT* system was developed using *PUMATECH Satellite Forms Enterprise Edition* application development environment.

References

Anatol, T.I., Y. Holder. 1995. A scoring System for Use in the Diagnosis of Acute Abdominal Pain in Childhood. *W.I. Med. J.* **44** 67-69.

Fioravanti, M., F. Di Cesare, L. Ramelli, F. La Torre, A. Nicastro, S. Messinetti, R.Lazarro. 1988. Presurgery Information and Psychological Adjustment to Enterostomy. *Italian Journal of Surgical Science* **18** 55-61.

Ford, N. 1998. *Web Developer.Com Guide to Building Intelligent Web Sites with Java Script*. J. Wiley

Grabisch, M. 1997. k-Order Additive Discrete Fuzzy Measures and Their Representation. *Fuzzy Sets and Systems* **89** 445-456.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Greco, S., B. Matarazzo, R. Slowinski. 1998. Fuzzy Measures as a Technique for Rough Set Analysis. Proc. 6th European Congress on Intelligent Techniques and Soft Computing (EUFIT'98). Aachen, 99-103.

Greco, S., B. Matarazzo, R. Slowinski. 2000. Dealing with Missing Data in Rough Set Analysis of Multi-Attribute and Multi-Criteria Decision Problems. S.H. Zanakis, G. Doukidis, C. Zopounidis, eds. *Decision Making: Recent Developments and Worldwide Applications*. Kluwer Academic Publishers 295-316.

Greco, S., B. Matarazzo, R. Slowinski. 2001. Rough Sets Theory for Multicriteria Decision Analysis. *European Journal of Operational Research* **129** 1-47.

Hallan, S., T. Edna. 1997. Estimating the Probability of Acute Appendicitis using Clinical Criteria of a Structured Record Sheet: The Physician Against the Computer. *European Journal of Surgery* **163** 427-432.

Korner, H., K. Sondenaa, J.A. Soreide. 1998. Structured Data Collection Improves the Diagnosis if Acute Appendicitis. *British Journal of Surgery* **85** 341-344.

Lin, T., N. Cercone eds. 1997. *Rough Sets and Data Mining*. Kluwer Academic Publishers.

Michalowski, W., S. Rubin, R. Slowinski, Sz. Wilk. 2001. Triage of the Child With Abdominal Pain: A Clinical Algorithm for Emergency Patient Management. *Paediatrics and Child Health* **6** 23-28.

Mitchell, T. 1997. *Machine Learning*. McGraw Hill.

Owens, D.K., H.C. Sox, Jr. 1990. Medical Decision Making: Probabilistic Medical Reasoning. E.H. Shortliffe, L.E. Perreault, L.W. Fagan, G. Wiederhold, eds., *Medical Informatics: Computer Applications in Health Care*. Addison-Wesley 70-116.

Pawlak, Z. 1991. *Rough Sets: Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers.

Pawlak, Z., R. Slowinski. 1994. Rough Set Approach to Multi-attribute Decision Analysis. *European Journal of Operational Research* **72** 443-459.

Predki, B., Sz. Wilk. 1999. Rough Set Based Data Exploration Using ROSE System. Z.W. Ras, A. Skowron, eds. *Foundations of Intelligent Systems*. Springer-Verlag 172-180.

Rubin, S., W. Michalowski, R. Slowinski, Sz. Wilk. 2000. Computer-assisted Triage of Abdominal Pain in Childhood: A Suggested Clinical Algorithm. Proc. Conference on Simulation in Health Sciences, 19-24.

Shapley, L.S. 1953. A Value for n-Person Games. H.W. Kuhn, A.W. Tucker, eds. *Contributions to the Theory of Games II*, Princeton University Press 307-317.

Simpson, E.T., A. Smith. 1996. The Management of Acute Abdominal Pain in Children. *Journal of Paediatric Child Health* **32** 110-112.

Slowinski, R. 1995. Rough Sets Approach to Decision Analysis. *AI Expert Magazine* **10** 8-25.

Stefanowski, J., D. Vanderpooten. 2001. Induction of Decision Rules in Classification and Discovery-oriented Perspectives. *International Journal of Intelligent Systems* **16** 13-28.

Published in Journal of Decision Support Systems, vol. 36 no. 2, 2003, 161-176

Stefanowski, J., Sz. Wilk. 2001. Evaluating Business Credit Risk by Means of Approach Integrating Decision Rules and Case Based Reasoning. *International Journal of Intelligent Systems in Accounting, Finance and Management* (to appear).

Williams, N.M.A., J.M. Johnstone, N.W. Everson. 1998. The Diagnostic Value of Symptoms and Signs in Childhood Abdominal Pain. *Journal of the Royal College of Surgery* **43** 390-392.

Ziarko, W. 1993. Variable Precision Rough Sets Model. *Journal of Computer and Systems Sciences* **1** 39-59.