Granular Rules for Medical Diagnosis

Shusaku Tsumoto(✉) and Shoji Hirano

Department of Medical Informatics, Faculty of Medicine, Shimane University, 89-1 Enya-cho, Izumo 693-8501, Japan
{tsumoto,hirano}@med.shimane-u.ac.jp
http://www.med.shimane-u.ac.jp/med_info/tsumoto/

Abstract. This paper discusses granular models of medical diagnostic rules which is an extension of rough set rule model. Medical diagnostic reasoning is characterized by three processes: focusing mechanism, differential diagnosis and detection of complications. First, focusing mechanism uses a set of symptoms which are always observed by almost all the cases of a candidate and if a case does not include any one of them, the candidate will be rejected. Second, from selected candidates, a set of symptoms which are highly observed in the cases are used for confirming the differential diagnosis. Finally, detection of complications is a set of symptoms whose occurrence of a candidate is very low but are very important for diagnosis of other diseases. These rule models can be easily described by an extension of rough set model: supporting sets of the first two sets of symptoms correspond to upper and lower approximations of a target concept. The final one is described by interrelations between a target concept and other concepts, which will be a new type of information granules.

1 Introduction

Classical medical diagnosis of a disease assumes that a disease is defined as a set of symptoms, in which the basic idea is *symptomatically*. Symptomatically had been a major diagnostic rules before laboratory and radio-logical examinations. Although the power of symptomatology for differential diagnosis is now lower, it is true that change of symptoms are very important to evaluate the status of chronic status. Even when laboratory examinations cannot detect the change of patient status, the set of symptoms may give important information to doctors.

Symptomatological diagnostic reasoning is conducted as follows. First, doctors make physical examinations to a patient and collect the observed symptoms. If symptoms are observed enough, a set of symptoms give some confidence to diagnosis of a corresponding disease. Thus, correspondence between a set of manifestations and a disease will be useful for differential diagnosis. Moreover, similarity of diseases will be inferred by sets of symptoms.

The author has been discussed modeling of symptomatological diagnostic reasoning by using the core ideas of rough sets since [12]: selection of candidates (screening) and differential diagnosis are closely related with diagnostic rules

R. B. Kearfott et al. (Eds.): IFSA 2019/NAFIPS 2019, AISC 1000, pp. 681–691, 2019.
https://doi.org/10.1007/978-3-030-21920-8_60

obtained by upper and lower approximations of a given concept. Thus, this paper discusses formalization of medical diagnostic rules which is closely related with rough set rule model. The important point is that medical diagnostic reasoning is characterized by focusing mechanism, composed of screening and differential diagnosis, which corresponds to upper approximation and lower approximation of a target concept. Furthermore, this paper focuses on detection of complications, which can be viewed as relations between rules of different diseases.

The paper is organized as follows. Section 2 shows characteristics of medical diagnostic process. Section 3 introduces rough sets and basic definition of probabilistic rules. Section 4 gives two style of formalization of medical diagnostic rules. The first one is a deterministic model, which corresponds to Pawlak's rough set model. And the other one gives an extension of the above ideas in probabilistic domain, which can be viewed as application of variable precision rough set model [14]. Section 5 proposes a new rule induction model, which includes formalization of rules for detection of complications. Finally, Sect. 6 concludes this chapter

2 Background: Medical Diagnostic Process

This section focuses on medical diagnostic process as rule-based reasoning. The fundamental discussion of medical diagnostic reasoning related with rough sets is given in [11].

2.1 RHINOS

RHINOS is an expert system which diagnoses clinical cases on headache or facial pain from manifestations. In this system, a diagnostic model proposed by Matsumura [1] is applied to the domain, which consists of the following three kinds of reasoning processes: exclusive reasoning, inclusive reasoning, and reasoning about complications.

First, exclusive reasoning excludes a disease from candidates when a patient does not have a symptom which is necessary to diagnose that disease. Secondly, inclusive reasoning suspects a disease in the output of the exclusive process when a patient has symptoms specific to a disease. Finally, reasoning about complications suspects complications of other diseases when some symptoms which cannot be explained by the diagnostic conclusion are obtained.

Each reasoning is rule-based and all the rules needed for diagnostic processes are acquired from medical experts in the following way.

Exclusive Rules

These rule correspond to exclusive reasoning. In other words, the premise of this rule is equivalent to the necessity condition of a diagnostic conclusion. From the discussion with medical experts, the following six basic attributes are selected which are minimally indispensable for defining the necessity condition: *1. Age, 2. Pain location, 3. Nature of the pain, 4. Severity of the pain, 5. History since onset, 6. Existence of jolt headache.* For example, the exclusive rule of common migraine is defined as:

```
In order to suspect common migraine,
the following symptoms are required:
pain location: not eyes,
nature :throbbing or persistent or radiating,
history: paroxysmal or sudden and
jolt headache: positive.
```

One of the reasons why the six attributes are selected is to solve an interface problem of expert systems: if all attributes are considered, all the symptoms should be input, including symptoms which are not needed for diagnosis. To make exclusive reasoning compact, we chose the minimal requirements only. It is notable that this kind of selection can be viewed as the ordering of given attributes, which is expected to be induced from databases. This issue is discussed later in Sect. 6

Inclusive Rules

The premises of inclusive rules are composed of a set of manifestations specific to a disease to be included. If a patient satisfies one set, this disease should be suspected with some probability. This rule is derived by asking the medical experts about the following items for each disease: *1. a set of manifestations by which we strongly suspect a disease. 2. the probability that a patient has the disease with this set of manifestations: SI (Satisfactory Index) 3. the ratio of the patients who satisfy the set to all the patients of this disease: CI (Covering Index) 4. If the total sum of the derived CI (tCI) is equal to 1.0 then end. Otherwise, goto 5. 5. For the patients with this disease who do not satisfy all the collected set of manifestations, goto 1.* Therefore a positive rule is described by a set of manifestations, its satisfactory index (SI), which corresponds to *accuracy measure*, and its covering index (CI), which corresponds to *total positive rate.* Note that SI and CI are given empirically by medical experts.
For example, one of three positive rules for common migraine is given as follows.

```
If history: paroxysmal, jolt headache: yes,
nature: throbbing or persistent,
prodrome: no, intermittent symptom: no,
persistent time: more than 6 hours,
and location: not eye,
then common migraine is suspected with
accuracy 0.9 (SI=0.9) and this rule covers
60 percent of the total cases (CI=0.6).
```

Disease Image: Complications Detection

This rule is used to detect complications of multiple diseases, acquired by all the possible manifestations of the disease. By the use of this rule, the manifestations which cannot be explained by the conclusions will be checked, which suggest complications of other diseases. For example, the disease image of common migraine is:

```
The following symptoms can be explained by
common migraine: pain location: any or
depressing: not or jolt headache: yes or ...
```

Therefore, when a patient who suffers from common migraine is depressing, it is suspected that he or she may also have other disease.

2.2 Focusing Mechanism

The most important process in medical differential diagnosis shown above is called a focusing mechanism [7,13]. Even in differential diagnosis of headache, medical experts should check possibilities of more than 100 candidates, though frequent diseases are 5 or 6. These candidates will be checked by past and present history, physical examinations, and laboratory examinations. In diagnostic procedures, a candidate is excluded one by one if symptoms necessary for diagnosis are not observed.

Focusing mechanism consists of the following two styles: exclusive reasoning and inclusive reasoning. Relations of this diagnostic model with another diagnostic model are discussed in [5,11], which is summarized in Fig. 1: First, exclusive reasoning excludes a disease from candidates when a patient does not have symptoms that is necessary to diagnose that disease. Second, inclusive reasoning suspects a disease in the output of the exclusive process when a patient has symptoms specific to a disease. Based on the discussion with medical experts, these reasoning processes are modeled as two kinds of rules, negative rules (or exclusive rules) and positive rules; the former corresponds to exclusive reasoning, the latter to inclusive reasoning [1]. [1]

3 Basics of Rule Definitions

3.1 Rough Sets

In the following sections, we use the following notation introduced by Grzymala-Busse and Skowron [4], based on rough set theory [2]. Let U denote a nonempty finite set called the universe and A denote a nonempty, finite set of attributes,

[1] Implementation of detection of complications is not discussed here because it is derived after main two process, exclusive and inclusive reasoning. The way to deal with detection of complications is discussed in Sect. 5.

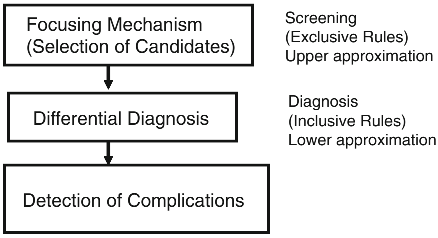

Fig. 1. Focusing mechanism.

i.e., $a : U \to V_a$ for $a \in A$, where V_a is called the domain of a, respectively. Then a decision table is defined as an information system, $A = (U, A \cup \{d\})$. The atomic formulas over $B \subseteq A \cup \{d\}$ and V are expressions of the form $[a = v]$, called descriptors over B, where $a \in B$ and $v \in V_a$. The set $F(B, V)$ of formulas over B is the least set containing all atomic formulas over B and closed with respect to disjunction, conjunction, and negation.

For each $f \in F(B, V)$, f_A denotes the meaning of f in A, i.e., the set of all objects in U with property f, defined inductively as follows:

1. If f is of the form $[a = v]$, then $f_A = \{s \in U | a(s) = v\}$.
2. $(f \wedge g)_A = f_A \cap g_A$; $(f \vee g)_A = f_A \vee g_A$; $(\neg f)_A = U - f_a$.

3.2 Classification Accuracy and Coverage

3.2.1 Definition of Accuracy and Coverage

By use of the preceding framework, classification accuracy and coverage, or true positive rate are defined as follows.

Definition 1. Let R and D denote a formula in $F(B, V)$ and a set of objects that belong to a decision d. Classification accuracy and coverage(true positive rate) for $R \to d$ is defined as:

$$\alpha_R(D) = \frac{|R_A \cap D|}{|R_A|}(= P(D|R)), \tag{1}$$

$$\kappa_R(D) = \frac{|R_A \cap D|}{|D|}(= P(R|D)), \tag{2}$$

where $|S|$, $\alpha_R(D)$, $\kappa_R(D)$, and $P(S)$ denote the cardinality of a set S, a classification accuracy of R as to classification of D, and coverage (a true positive rate of R to D), and probability of S, respectively.

It is notable that $\alpha_R(D)$ measures the degree of the sufficiency of a proposition, $R \to D$, and that $\kappa_R(D)$ measures the degree of its necessity. For example, if $\alpha_R(D)$ is equal to 1.0, then $R \to D$ is true. On the other hand, if $\kappa_R(D)$ is equal to 1.0, then $D \to R$ is true. Thus, if both measures are 1.0, then $R \leftrightarrow D$.

3.3 Probabilistic Rules

By use of accuracy and coverage, a probabilistic rule is defined as:

$$R \stackrel{\alpha,\kappa}{\to} d \quad s.t.\ R = \wedge_j [a_j = v_k], \alpha_R(D)\, \delta_\alpha \text{ and } \kappa_R(D)\, \delta_\kappa, \tag{3}$$

where D denotes a set of samples that belong to a class d. If the thresholds for accuracy and coverage are set to high values, the meaning of the conditional part of probabilistic rules corresponds to the highly overlapped region. This rule is a kind of probabilistic proposition with two statistical measures, which is an extension of Ziarko's variable precision model (VPRS) [14].[2]

It is also notable that both a positive rule and a negative rule are defined as special cases of this rule, as shown in the next sections.

4 Formalization of Medical Diagnostic Rules

4.1 Deterministic Model

4.1.1 Positive Rules

A positive rule is defined as a rule supported by only positive examples. Thus, the accuracy of its conditional part to a disease is equal to 1.0. Each disease may have many positive rules. If we focus on the supporting set of a rule, it corresponds to a subset of the lower approximation of a target concept, which is introduced in rough sets [2]. Thus, a positive rule is defined as:

$$R \to d \quad s.t. \qquad R = \wedge_j [a_j = v_k], \quad \alpha_R(D) = 1.0 \tag{4}$$

where D denotes a set of samples that belong to a class d.

This positive rule is often called a deterministic rule. However, we use the term, positive (deterministic) rules, because a deterministic rule supported only by negative examples, called a negative rule, is introduced below.

4.1.2 Negative Rules

The important point is that a negative rule can be represented as the contrapositive of an exclusive rule [13]. An exclusive rule is defined as a rule whose supporting set covers all the positive examples. That is, the coverage of the rule to a disease is equal to 1.0. That is, an exclusive rule represents the necessity condition of a decision. The supporting set of an exclusive rule corresponds to the upper approximation of a target concept, which is introduced in rough sets [2]. Thus, an exclusive rule is defined as:

$$R \to d \quad s.t. \qquad R = \vee_j [a_j = v_k], \quad \kappa_R(D) = 1.0, \tag{5}$$

where D denotes a set of samples that belong to a class d.

[2] This probabilistic rule is also a kind of *rough modus ponens* [3].

Next, let us consider the corresponding negative rules in the following way. An exclusive rule should be described as:

$$d \to \vee_j [a_j = v_k],$$

because the condition of an exclusive rule corresponds to the necessity condition of conclusion d. Since a negative rule is equivalent to the contrapositive of an exclusive rule, it is obtained as:

$$\wedge_j \neg [a_j = v_k] \to \neg d,$$

which means that if a case does not satisfy any attribute value pairs in the condition of a negative rule, then we can exclude a decision d from candidates.

Thus, a negative rule is represented as:

$$\wedge_j \neg [a_j = v_k] \to \neg d \quad s.t. \quad \forall [a_j = v_k] \kappa_{[a_j = v_k]}(D) = 1.0, \tag{6}$$

where D denotes a set of samples that belong to a class d.

Negative rules should also be included in a category of deterministic rules, because their coverage, a measure of negative concepts, is equal to 1.0. It is also notable that the set supporting a negative rule corresponds to a subset of negative region, which is introduced in rough sets [2].

In summary, positive and negative rules correspond to positive and negative regions defined in rough sets. Figure 2 shows the Venn diagram of those rules.

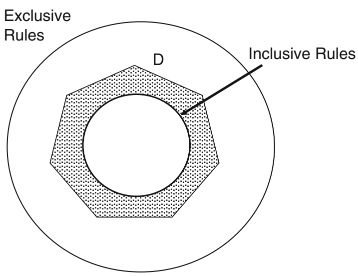

Fig. 2. Venn diagram of exclusive and positive rules.

4.2 Probabilistic Model

Although the above deterministic model exactly corresponds to original Pawlak rough set model, rules for differential diagnosis is strict for clinical setting, because clinical diagnosis may include elements of uncertainty.[3] Tsumoto [5]

[3] However, deterministic rule induction model is still powerful in knowledge discovery context as shown in [8].

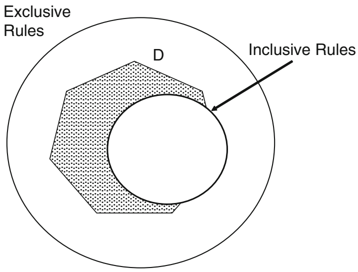

Fig. 3. Venn diagram of exclusive and inclusive rules.

relaxes the condition of positive rules and defines an inclusive rules, which models the inclusive rules of RHINOS model. The definition is almost the same as probabilistic rules defined in Sect. 3, except for the constraints for accuracy: the threshold for accuracy is sufficiently high. Thus, the definitions of rules are summarized as follows.

4.2.1 Exclusive Rules

$$R \rightarrow d \quad s.t.\ R = \vee_j [a_j = v_k], \quad (s.t. \quad \kappa[a_j = v_k](D) > \delta_\kappa) \quad \kappa_R(D) = 1.0. \quad (7)$$

4.2.2 Inclusive Rules

$$R \stackrel{\alpha,\kappa}{\rightarrow} d \quad s.t.\ R = \wedge_j [a_j = v_k], \quad \alpha_R(D) > \delta_\alpha \text{ and } \kappa_R(D) > \delta_\kappa. \quad (8)$$

In summary, positive and negative rules correspond to positive and negative regions defined in variable rough set model [14]. Figure 3 shows the Venn diagram of those rules.

Tsumoto introduces an algorithm for induction of exclusive and inclusive rules as PRIMEROSE-REX and conducted experimental validation and compared induced results with rules manually acquired from medical experts [5]. The results show that the rules do not include components of hierarchical diagnostic reasoning. Medical experts classify a set of diseases into groups of similar diseases and their diagnostic reasoning is multi-staged: first, different groups of diseases are checked, then final differential diagnosis is performed with the selected group of diseases. In order to extend the method into induction of hierarchical diagnostic rules, one of the authors proposes several approach to mining taxonomy from a dataset in [6,9,10].

5 Detection of Complications

The former rule induction models do not include reasoning about detection of complications, which is introduced as *disease image* as shown in Sect. 1. The core idea is that medical experts detect the symptoms which cannot be frequently occurred in the final diagnostic candidates. For example, let us assume that a patient suffering from muscle contraction headache, who usually complains of persistent pain, also complains of paroxysmal pain, say he/she feels a strong pain every one month. The situation is unusual and since paroxysmal pain is frequently observed by migraine, medical experts suspect that he/she suffers from muscle contraction headache and common migraine. Thus, a set of symptoms which are not useful for diagnosis of a disease may be important if they belong to the set of symptoms frequently manifested in other diseases. In other means, such the set of symptoms will be elements of detection of complications. Based on these observations, complications detection rules can be defined as follows:

5.1 Rules for Detection of Complications

Complications detection rule of diseases are defined as a set of rules each of which is included into inclusive rules of other diseases.[4]

$$\{R \rightarrow d \quad s.t.\ R = [a_i = v_j],\ \alpha_R(D) > \delta_\alpha, \kappa_R(D) > \delta_\kappa\} \tag{9}$$

Figure 4 depicts the relations between exclusive, inclusive and complications detection rules.

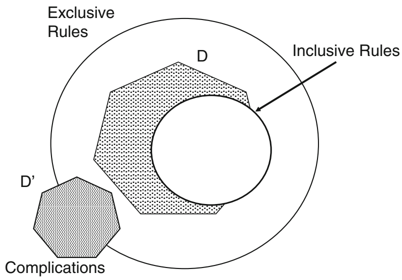

Fig. 4. Venn diagram of exclusive, inclusive and complications detection rules.

5.2 Classification of Observations

The relations between three types of rules can be visualized in a two dimensional plane, called $(\alpha, \kappa) - plane$, as shown in Fig. 5. The vertical and horizontal axis denotes the values of accuracy and coverage, respectively. Then, each rule can be plotted in the plane with its accuracy and coverage values. The region for

[4] The first term $R = [a_i = v_j]$ may not be needed theoretically. However, since deriving conjunction in an exhaustive way is sometimes computationally expensive, here this constraint is imposed for computational efficiency.

inclusive rules is shown in upper right, whereas the region for candidates of detection of complications is in lower left. When a rule of that region belongs to an inclusive rule of other disease, it is included into complications detection rule of the target diseases.

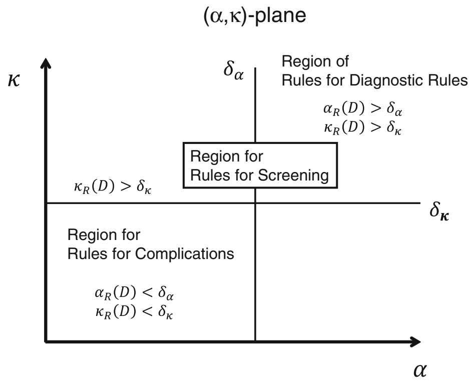

Fig. 5. Two dimensional plot: $(\alpha, \kappa) - plane$

6 Conclusion

Formalization of medical diagnostic reasoning based on symptomatology is discussed. Reasoning consists of three processes, exclusive reasoning, inclusive reasoning and complications detection, the former two of which belongs to a focusing mechanism. In exclusive reasoning, a disease is ruled out from diagnostic candidates when a patient does not have symptoms necessary for diagnosis. The process corresponds to screening. Second, in inclusive reasoning, a disease out of selected candidates is suspected when a patient has symptoms specific to a disease, which corresponds to differential diagnosis. Finally, if symptoms which are rarely observed in the final candidate, complication of other diseases will be suspected.

Previous studies are surveyed: one of the author concentrate on the focusing mechanism. First, in a deterministic version, two steps are modeled as two kinds of rules obtained from representations of upper and lower approximation of a given disease. Then, he extends it into probabilistic rule induction, which can be viewed as an application of VPRS.

Then, the authors formalize complications detection rules in this paper. The core idea is that the rules are not simply formalized by the relations between a set of symptoms and a disease, but by those between a symptoms, a target disease and other diseases. The next step will be to introduce an efficient algorithm to generate complication detection rules from data.

Acknowledgments. The author would like to thank past Professor Pawlak for all the comments on my research and his encouragement. Without his influence, one of the authors would neither have received Ph.D on computer science, nor become a professor of medical informatics. The author also would like to thank Professor Jerzy Grzymala-Busse, Andrezj Skowron, Roman Slowinski, Yiyu Yao, Guoyin Wang, Wojciech Ziarko for their insightful comments.

References

1. Matsumura, Y., Matsunaga, T., Maeda, Y., Tsumoto, S., Matsumura, H., Kimura, M.: Consultation system for diagnosis of headache and facial pain: "rhinos". In: Wada, E. (ed.) LP. Lecture Notes in Computer Science, vol. 221, pp. 287–298. Springer (1985)
2. Pawlak, Z.: Rough Sets. Kluwer Academic Publishers, Dordrecht (1991)
3. Pawlak, Z.: Rough modus ponens. In: Proceedings of International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems 98. Paris (1998)
4. Skowron, A., Grzymala-Busse, J.: From rough set theory to evidence theory. In: Yager, R., Fedrizzi, M., Kacprzyk, J. (eds.) Advances in the Dempster-Shafer Theory of Evidence, pp. 193–236. Wiley, New York (1994)
5. Tsumoto, S.: Automated induction of medical expert system rules from clinical databases based on rough set theory. Inf. Sci. **112**, 67–84 (1998)
6. Tsumoto, S.: Extraction of experts' decision rules from clinical databases using rough set model. Intell. Data Anal. **2**(3), 215–227 (1998)
7. Tsumoto, S.: Modelling medical diagnostic rules based on rough sets. In: Polkowski, L., Skowron, A. (eds.) Rough Sets and Current Trends in Computing. Lecture Notes in Computer Science, vol. 1424, pp. 475–482. Springer, Heidelberg (1998)
8. Tsumoto, S.: Automated discovery of positive and negative knowledge in clinical databases based on rough set model (2000)
9. Tsumoto, S.: Extraction of hierarchical decision rules from clinical databases using rough sets. Inf. Sci. (2003)
10. Tsumoto, S.: Extraction of structure of medical diagnosis from clinical data. Fundam. Inform. **59**(2–3), 271–285 (2004)
11. Tsumoto, S.: Rough sets and medical differential diagnosis. In: Skowron, A., Suraj, Z. (eds.) Rough Sets and Intelligent Systems. Intelligent Systems Reference Library, vol. 42, pp. 605–621. Springer (2013)
12. Tsumoto, S., Tanaka, H.: Induction of probabilistic rules based on rough set theory. In: Jantke, K.P., Kobayashi, S., Tomita, E., Yokomori, T. (eds.) Algorithmic Learning Theory, 4th International Workshop, ALT 1993, Tokyo, Japan, 8-10 November 1993, Proceedings. *Lecture Notes in Computer Science*, vol. 744, pp. 410–423. Springer (1993). https://doi.org/10.1007/3-540-57370-4
13. Tsumoto, S., Tanaka, H.: Automated discovery of medical expert system rules from clinical databases based on rough sets. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining 96, pp. 63–69. AAAI Press, Palo Alto (1996)
14. Ziarko, W.: Variable precision rough set model. J. Comput. Syst. Sci. **46**, 39–59 (1993)