Vietnam Journal of Science and Technology ***60*** *(5) (2022) 1149-1161*

doi:10.15625/2525-2518/16896

IMPROVING THE FUZZY EXPERT SYSTEM FOR DIAGNOSING DEPRESSIVE DISORDERS

Mai Thi Nu[1], Nguyen Hoang Phuong[2]

[1]E. Health Administration, Ministry of Health, 135 Nui Truc, Ba Dinh, Ha Noi, Viet Nam
[2]Thang Long University, Nghiem Xuan Yem Road, Hoang Mai District, Ha Noi, Viet Nam

Emails: *1.mainuitmoh@gmail.com, 2.nhphuong2008@gmail.com*

Received: 26 January 2022; Accepted for publication: 12 February 2022

Abstract. This paper proposes an improvement of the knowledge base and inference engine of the PORUL.DEP medical expert system for diagnosing depressive disorders. The knowledge base of PORUL.DEP includes more than 850 positive rules. PORUL.DEP has been tested on more than 260 medical records of depressed patients, giving a correct diagnosis of more than 95 % with light depressive disorder and without depressive disorder, but the remaining depressive disorders are not accurate, reaching only over 24 %. A new expert system, called STRESSDIAG, was developed on combining positive rules (for confirmation of conclusion) and negative rules (for exclusion of conclusion) for diagnosing depressive disorders. STRESSDIAG's knowledge base consists of more than 850 positive rules of PORUL.DEP and more than 120 negative rules. Abelian group operation of MYCIN is used to improve the inference engine based on fuzzy relations. STRESSDIAG gives a correct diagnosis of more than 76 % with 4 depressive disorder types and without depressive disorders, achieving an average percentage of more than 82 %, an increase of nearly 60 % compared to PORUL.DEP.

Keywords: fuzzy expert systems, positive rules, negative rules, diagnosis of depression types.

Classification numbers: 4.8.3, 4.10.2.

1. INTRODUCTION

According to the World Health Organization, depression is a common mental disorder characterized by sadness, loss of interest, feelings of guilt or low self-esteem, sleep disorders, and eating drinking disorders, feeling tired and poor concentration. It affects approximately 264 million people worldwide [1]. Depressive disorder can manifest itself in one of several types, such as light depressive disorder, middle depressive disorder, serious depressive disorder, and serious depressive disorder with mental disorder. Depressive disorder is the fourth leading cause of death worldwide and is predicted to be the second leading cause of death in 2030 [1]. Depressive disorder is one of the most widespread diseases in the world, accounting for about 3 - 5 % of the world's population. In Viet Nam, the number of patients with mental disorders accounts for about 5 - 7 % of the population. A doctor when diagnosing a depressive disorder often faces the problem of how to recognize the right type of depressive disorder and prescribe the right medication for the patient.

There is a relationship either between symptoms and diseases or symptom combinations and diseases. Some symptoms always entail the appearance of some other symptoms or negate the appearance of opposing symptoms. In depressive disorders, for instance, if the patient has "suicidal" symptoms, there is a small possibility that the patient has a light depressive disorder or a middle depressive disorder, or a very low depressive disorder. If a patient has "hallucinations" symptoms, the likelihood that the patient has a light depressive disorder or middle depressive disorder or serious depressive disorder is very low. In short, "suicide" symptoms are called negative symptoms of light depressive disorder and middle depressive disorder; "hallucinations" symptoms are called negative symptoms of light depressive disorder, middle depressive disorder and serious depressive disorder. These negative symptoms will be used to differentiate between the diagnostic criteria for depressive disorders [2].

In the paper, we propose an improvement of the knowledge base and inference engine of the PORUL.DEP medical expert system for diagnosing depressive disorders in order to enhance the accuracy of diagnosing depressive disorders. The contributions of the paper are as follows:

1) Improving the knowledge base of the PORUL.DEP medical expert system by adding and modifying negative rules;
2) Improving the inference engine based on fuzzy relations and Abelian group operation in the MYCIN medical expert system.

The structure of the paper is as follows. Section 2 presents the knowledge base and inference engine of the PORUL.DEP medical expert system and some experimental results. Section 3 proposes improved expert system STRESSDIAG for diagnosing depressive disorders by improving the knowledge base and inference engine of the PORUL.DEP. Finally, conclusions and future research directions are given.

2. THE KNOWLEDGE BASE AND INFERENCE ENGINE OF PORUL.DEP

The main components of an expert system for diagnosing depressive disorders are knowledge base and inference engine. The knowledge base is an important component of an expert system which containing problem-solving knowledge of a particular application. In a rule-based expert system, this knowledge is represented in the form of rules: if… then.... The inference engine has the function of processing and controlling the knowledge represented in the knowledge base in order to respond to questions and user requirements, and to apply the knowledge to solve real-life problems. It is basically an interpreter for the knowledge base [3 - 6].

2.1. The knowledge base

As mentioned above, the knowledge base is a part of the expert systems. The knowledge base of PORUL.DEP contains 857 positive rules, which include 124 rules for diagnosing light depressive disorder, 146 rules for diagnosing middle depressive disorder, 263 rules for diagnosing serious depressive disorder, and 324 rules for diagnosing depressive disorder with mental disorder [2, 7 - 11].

2.2. The Inference Engine

The Inference Engine processes and controls the knowledge represented in the knowledge base to respond to questions and user requests, and apply the knowledge to solve practical problems. In other words, it is an interpreter for the knowledge base.

Let S be the set of symptoms, $S = \{S_1, S_2, ..., S_i, \ldots, S_n\}$, where S_i is the i^{th} symptom. In our case, these symptoms included: complexion reduction; loss of interest and pleasure; reduced energy; reduced attention; reduced self-esteem and self-confidence; having suicidal thoughts; feelings of guilt, unworthiness, feeling gray, self-destructive / thoughts of suicidal behavior; sleep disorders, eating disorders, suicide, delusions, and hallucinations. Symptom S_i takes the value μ_S, in [0,1] which indicates the degree to which the patient exhibits symptom S_i [2, 12, 13].

Intermediate combinations (fuzzy logical combinations of symptoms and diseases) were introduced to the model of the pathophysiological states of patients; and Symptom combinations SC_i are combinations of symptoms, diseases and intermediate combinations. A relationship R_{PSC} is established, defined by $\mu_{R_{PSC}}(P_q, SC_i) = \mu_{SC_i}$ for patient P_q where $SC_i = \{SC_1, \ldots, SC_m\}$ formally describes the symptom combinations observed on the patient.

$$\mu_{R_{PCS}}(P_q,S) = \min\{\mu_{R_{PS_1}}(P_q,S_1), \mu_{R_{PS2}}(P_q,S_2), \ldots, \mu_{R_{PS_i}}(P_q,S_i), \ldots ,\mu_{R_{S_nP}}(P_q,S_n)\} \quad (1)$$

Let D be the set of diseases, $D = \{D_1, D_2,\ldots,D_m\}$, where D_j is the j^{th} depressive disorder. In our case, m = 4, including light depressive disorder, middle depressive disorder, serious depressive disorder, and serious depressive disorder with mental disorder.

A binary fuzzy relationship R_{PS} is established, defined by $\mu_{R_{PS}}(P_q, S_i) = \mu_{S_i}$ for patient P_q, where $P_q = \{P_1, \ldots P_p\}$ and $S_i \in \{S_1, \ldots, S_m\}$. $\mu_{R_{PS}}(P_q, S_i) \in [0,1]$.

A symptom-disease relationship R_{PD} is established, defined by $\mu_{R_{PD}}(P_q, D_j) = \mu_{D_j}$ for patient P_p, where $D_j = \{D_1, \ldots D_n\}$.

A fuzzy relationship R_{SD} is established, defined by $\mu_{R_{SD}}(S,D_j) \in [0,1]$. This value represents the degree of confidence in the likelihood of having or not having D_j disease when a symptom or a set of symptoms S is present. Express the symptom-disease relationship as follows:

$$\text{IF S THEN CONFIRM D WITH (FUZZY DEGREE)} \quad (2)$$

R_{SD} is now a confirming relationship that the patient has D_j disease when there is a symptom or a set of symptoms S. The value $\mu_{R_{SD}}(S,D_j)$ is a fuzzy degree or rule weight.

A fuzzy relationship R_{PD} is established, defined by $\mu_{R_{PD}}(P_q,D_j)$. Determining this relationship also means making a diagnosis of the patient's likelihood. Based on these fuzzy relationships, the MaxMin inference is used to deduce the fuzzy value $\mu_{R_{PD}}(P_q,D_j)$ which indicates the degree of confirmation of disease D_j suffered by patient P_q from the observed symptoms. This MaxMin composition is as follows:

$$R_{PD} = R_{PS} \circ R_{SD} \quad (3)$$

where R_{PS} is relationship of symptom S or combination S ($S_1, S_1, \ldots S_n$) and patient P_q [18 - 19].

$$\mu_{R_{PD}}(P_q,D_j) = \max_{S_i} \min [\mu_{R_{PS}}(P_q, S), \mu_{R_{SD}}(S, D_j)]$$

$$\mu_{R_{PD}}(P_q,D_j,rule_t) = \min [\mu_{R_{PS}}(P_q, S), \mu_{R_{SD}}(S, D_j, rule_t)] \quad (4)$$

$$= \min(\{\mu_{R_{PS}}(S_i, P_q)\}, \mu_{R_{SD}}(S, D_j, rule_t))$$

where $\mu_{R_{SD}}(S,D_j,rule_t)$ is the degree of confirming D_j disease when there is a symptom S or a set of symptoms S on the $rule_t$ (weight of $rule_t$).

$\{ \mu_{R_{PD}}(P_q,D_j, rule_t), \ldots, \mu_{R_{PD}}(P_q,D_j, rule_1), \ldots, \mu_{R_{PD}}(P_q,D_j, rule_n)\}$ t = 1,..,n.

Calculate $\mu_{R_{PD}}(P_q,D_j)$ from the set of $\{ \mu_{R_{PD}}(P_q,D_j,rule_t) \}$ according to formula (3):

$$\mu_{R_{PD}}(P_q,D_j) = \max_{S_i} [\mu_{R_{PD}}(P_q,D_j,rule_1), \ldots, \mu_{R_{PD}}(P_q,D_j,rule_n)] \quad (5)$$

if $\mu_{R_{pd}}(P_q,D_j) = 1$ means absolute confirmation of the conclusion of D_j; $\mu_{R_{pd}}(P_q,D_j) = 0$ means absolute exclusion of the conclusion of D_j; $0 < \mu_{R_{pd}}(P_q,D_j) < 1$ means confirmation of the conclusion of D_j with some fuzzy degree [10].

Example 1. An example of illustrating the calculation of PORUL.DEP. Assuming that patient P_q has symptoms S = $\{S_1, S_2, S_3, S_4, S_5, S_6\}$, with the occurrence of symptoms is, respectively, as follows:

$\mu_{R_{PS}}(P_q, S_1) = 0.87$; $\mu_{R_{PS}}(P_q, S_2)= 0.92$; $\mu_{R_{PS}}(P_q, S_3)= 0.83$; $\mu_{R_{PS}}(P_q, S_4)= 0.94$; $\mu_{R_{PS}}(P_q, S_5)= 0.78$; $\mu_{R_{PS}}(P_q, S_6)= 0.83$.

The diagnostic process will take place in the following steps:

Step 1: List all the rules whose premise is a subset of the set S.

Step 2: Group the rules with the same disease conclusion. Supposing that, with the conclusion that disease D3 = “serious depressive disorder”, we can group a set of rules as follows:

- Rule 1: IF S_1 THEN D_3, 0.3
- Rule 2: IF S_2 THEN D_3, 0.25
- Rule 3: IF S_4 THEN D_3, 0.35
- Rule 4: IF S_5 THEN D_3, 0.17
- Rule 5: IF $S_1 \wedge S_3 \wedge S_5$ THEN D_3, 0.76
- Rule 6: IF $S_2 \wedge S_4$ THEN D_3, 0.33
- Rule 7: IF $S_1 \wedge S_3$ THEN D_3, 0.39

Step 3, Step 4: With the group of rules for positive disease D_3, $\mu_{R_{PD}}(P_q, D_3)$ can be calculated as follows:

- Rule 1: IF S_1 THEN D_3, 0.3
- Rule 2: IF S_2 THEN D_3, 0.25
- Rule 3: IF S_4 THEN D_3, 0.35
- Rule 4: IF S_5 THEN D_3, 0,17
- Rule 5: IF $S_1 \wedge S_3 \wedge S_5$ THEN D_3, 0.76
- Rule 6: IF $S_2 \wedge S_4$ THEN D_3, 0.33
- Rule 7: IF $S_1 \wedge S_3$ THEN D_3, 0.39

Using formula (4), $\mu_{R_{PDrule_h}}(P_q, D_3)$ can be calculated as

$$\mu_{R_{PDrule_h}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_i) , \mu_{R_{SDrule_h}}(S_i, D_3)\}$$

In this example: i = {1, 2, 3, 4, 5}; j=3; h={1, 2, 3, 4, 5, 6, 7}

[Rule 1] $\mu_{R_{PDrule_1}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q , S_1), \mu_{R_{SDrule_1}}(S_1, D_3)\}$

= min {0.87, 0.3} = 0.3

[Rule 2] $\mu_{R_{PDrule_2}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_2), \mu_{R_{SDrule_2}}(S_2, D_3))$

= min {0.92, 0.25} = 0.25

[Rule 3] $\mu_{R_{PDrule_3}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_4), \mu_{R_{SDrule_3}}(S_4, D_3)\}$

= min {0.94, 0.35} = 0.35

[Rule 4] $\mu_{R_{PDrule_4}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_5), \mu_{R_{SDrule_4}}(S_5, D_3)\}$

= min {0.78, 0.17} = 0.17

[Rule 5] $\mu_{R_{PDrule_5}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_1), \mu_{R_{PS}}(P_q, S_3), \mu_{R_{PS}}(P_q, S_5), \mu_{R_{SDrule_5}}(S_{1,3,5}, D_3)\}$= min {0.87, 0.83, 0.78 , 0.76} = 0.76

[Rule 6] $\mu_{R_{PDrule_6}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_2), \mu_{R_{PS}}(P_q, S_4), \mu_{R_{SDrule_6}}(S_{2,4}, D_3)\}$

= min {0.92, 0.94, 0.33} = 0.33

[Rule 7] $\mu_{R_{PDrule_7}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q, S_1), \mu_{R_{PS}}(P_q, S_3), \mu_{R_{SDrule_7}}(S_{1,3}, D_3)\}$

= min {0.87, 0.83 , 0.39} = 0.39

Using formular (5), we can calculate $\mu_{R_{PD}}(P_q, D_3)$ as follows:

$$\mu_{R_{PD}}(P_q, D_3) = \max \{\mu_{R_{PDrule_1}}(P_q, D_3), \ldots, \mu_{R_{PDrule_p}}(P_q, D_3)\}$$

In this example j=3 ; h={1, 2, 3, 4, 5, 6, 7 }

$\mu_{R_{PD}}(P_q, D_3) = \mathrm{Max} \{\mu_{R_{PDrule_1}}(P_q, D_3), \ldots, \mu_{R_{PDrule_h}}(P_q, D_3), \ldots, \mu_{R_{PDrule_7}}(P_q, D_3)\}$

$$\mu_{R_{PD}}(P_q, D_3) = \mathrm{Max} \{0.3; 0.25; 0.35; 0.17; 0.76; 0.33; 0.39\} = 0.76$$

Step 5: Make a final conclusion that patient P_q has serious depressive disorder with a fuzzy degree of 0.76, that is, almost certainly patient P_q has serious depressive disorder.

2.3. Experimental Results

Table 1. Comparison of diagnostic results for each type of depressive disorder (positive rules).

Type of depressive disorder	total	In medical records	PORUL.DEP	rate
light depressive disorder	48	48	46	95.8 %
middle depressive disorder	60	60	Inappropriate	
serious depressive disorder	50	50	Inappropriate	
serious depressive disorder with mental disorder	86	86	Inappropriate	
without depressive disorder	20	20	19	95 %

The tests were performed with a data set of 264 medical records, which included 48 diagnosed with light depressive disorders, 60 diagnosed with middle depressive disorder, 50 diagnosed with serious depressive disorders, 86 medical records diagnosed with serious depressive disorders with mental disorder, and 20 without depressive disorders.

In the test with the patient data set, the author fully updated the disease information of 264 medical records in the expert system software. The diagnostic results of the expert system were compared with the medical records, giving details as shown in the table below.

Table 2. Percentage of correct diagnosis of PORUL.DEP.

Total	In medical records	PORUL.DEP	rate
264	264	65	24.6 %

The above experimental data show that the expert system gives good results for light depressive disorder and without depressive disorder; the remaining depressive disorders are not accurate because the diagnostic standards for these 4 types of depression overlap with some of the symptoms.

It can be seen that the cause of the incorrect diagnoses above is due to the overlapping diagnostic criteria for depressive disorders. To overcome these incorrect diagnoses, PORUL.DEP needs to be improved by adding negative knowledge (negative rules) to clearly distinguish the boundary between diagnostic standards and baseline appropriate inference mechanism for the new knowledge base.

3. IMPROVING FUZZY EXPERT SYSTEM FOR DEPRESSIVE DISORDER DIAGNOSIS

PORUL.DEP was improved by adding negative rules to the knowledge base. This expert system, called STRESSDIAG, is presented for the diagnosis of depressive disorders including positive and negative rules. STRESSDIAG is designed and built with a suitable inference engine for positive rules and negative rules [12].

3.1. Improving knowledge base

The knowledge base's PORUL.DEP is improved by adding the rules of the negative form (rules based on negative knowledge) in the following form:

IF S THEN EXCLUDED WITH FUZZY DEGREE (6)

where "S" is a symptom or a combination of symptoms that are combined by AND without using the NOT operator; "D" is a negative disease; "FUZZY DEGREE" is the rule weight, which indicates the degree of certainty of "Conclusion" with the value in [0,1]. The confidence coefficient shows the relationship between symptom (or symptom combination) and disease.

The introduction of rules that conclude in the negative form is a strong point of this inference engine. It helps the expert system not only simulate the confirmatory diagnosis process of common diseases but also simulate the process of exclusion diagnosis and differential diagnosis (it is very common in medicine). The knowledge base of STRESSDIAG includes 124 negative rules and 857 positive rules [2].

3.2. Improving inference engine

Fuzzy inference system is the most important modeling tool based on fuzzy set theory.

Conventional fuzzy inference systems are built based on field experts that they have been used in automatic control, data classification, decision analysis, and expert systems [14 - 18].

A fuzzy relationship R^t_{PD} is established, defined by μ^t_{RPD} $(P_q,D_j) \in [-1,1]$. This value represents the fuzzy degree D_j of patient P_q.

The relationship R^t_{PD} is composed of two component relations: R_{SD} is positive relationship and R^e_{PD} is exclusion relationship. R_{SD} is defined in section 2, R^e_{PD} is defined as follows:

A fuzzy relationship R^e_{PD} is established, defined by $\mu^e_{R_{PD}}(P_q,D_j) \in [0,1]$.

$\mu^e_{R_{PD}}(P_q,D_j) = 1$ means absolute exclusion of the conclusion of D_j.

$\mu^e_{R_{PD}}(P_q,D_j) = 0$ means no exclusion of the conclusion of D_j.

$0 < \mu^e_{R_{PD}}(P_q,D_j) < 1$ means absolute exclusion of the conclusion of D_j with some fuzzy degree.

$$R^e_{PD} = R_{PS} \circ R^e_{SD} \quad (7)$$

R_{PS} is the relationship of symptom S or combination S $(S_1, S_1, ...S_n)$ and patient P_q, defined by $\mu_{R_{PS}}(P_q,S)$; R^e_{SD} is exclusion relationship of disease D_j, it is defined by $\mu^e_{R_{SD}}(S_i,D_j)$. Based on these fuzzy relationships, the MaxMin inference are used to deduce the fuzzy value $\mu^e_{R_{PD}}(P_q,D_j)$ as follows:

$$\mu^e_{R_{PD}}(P_q,D_j) = \max_{S_i} \min\,[\mu_{R_{PS}}(P_q,S)\,,\ \mu^e_{R_{SD}}(S,D_j)] \quad (8)$$

for each disease D_j, there is a set of negative rules: Rule = $\{rule_1,\ldots,rule_t,\ldots,rule_n)$, with conclusion $D = D_j$, $rule_t$ is t^{th} rule , $t = 1...n$. Then $\mu^e_{R_{SDrule_t}}(P_q, D_j)$ is the negative degree of the possibility of getting disease D_j of patient P_q according to $rule_t$:

$$\mu^e_{R_{PDrule_t}}(P_q, D_j) = \min\,\{\mu^e_{R_{PS}}(P_q, S_i)\,,\ \mu^e_{R_{SDrule_t}}(S_i, D_j)\} \quad (9)$$

where $\{\mu^e_{R_{PS}}(P_q, S_i)\}$ is the set of degrees of symptom S_i appearing in patient P_q; S_i is a symptom in the hypothesis of $rule_t$; $\mu^e_{R_{SDrule_t}}(S_i, D_j)$ is degree of negative disease D_j for symptom S_i according to $rule_t$ (it is weight of $rule_t$).

Obtaining a set of negative degrees of disease D_j corresponding to each $rule_t$

$$\{\ \mu^e_{R_{PDrule_1}}(P_q, D_j),\ldots,\ \mu^e_{R_{PDrule_t}}(P_q, D_j),\ldots\ ,\ \mu^e_{R_{PDrule_n}}(P_q, D_j)\}$$

we can calculate $\mu^e_{R_{PD}}(P_q, D_j)$ from $\{\ \mu^e_{R_{PDrule_t}}(P_q, D_j)\ \}$

$$\mu^e_{R_{PD}}(P_q, D_j) = \max_{S_i}\ \{\mu^e_{R_{PDrule_1}}(P_q, D_j),\ldots,\ \mu^e_{R_{PDrule_n}}(P_q, D_j)\} \quad (10)$$

Operation $\oplus$ is an ordered Abelian group operation on [-1,1]. We can use an operation from the medical expert system MYCIN [10], in which the MYCIN group operation $\oplus$ on [-1,1] is defined as follows:

$$X \oplus Y = X + Y - X.\,Y \text{ where } X\,,\,Y \geq 0;$$

$$X \oplus Y = X + Y + X.\,Y \text{ where } X\,,\,Y \leq 0;$$

$$X \oplus Y = \frac{X + Y}{1 - \min(\,|X|,|Y|\,)} \text{ where } X \text{ x } Y < 0.$$

$\mu^t_{R_{PD}}$ (P_q, D_j) can be calculated as: $\mu^t_{R_{PD}}(P_q,D_j) = \mu_{R_{PD}}(P_q,D_j) \oplus (-\mu^e_{R_{PD}}(P_q,D_j))$.

Because $\mu_{R_{PD}}(P_q,D_j)$ and $-\mu^e_{R_{PD}}(P_q,D_j)$ are opposite inside, so

$$\mu^t_{R_{PD}}(P_q, D_j) = \frac{\mu_{R_{PD}}(P_q,D_j)+(-\mu^e_{R_{PD}}(P_q,D_j))}{1-\min\{|\mu_{R_{PD}}(P_q,D_j)|,|\mu^e_{R_{PD}}(P_q,D_j)|\}} \quad (11)$$

$\mu^t_{R_{PD}}$ (P_q, D_j) = 1 means absolute confirmation of the conclusion of D_j; $0.6 \le \mu^t_{R_{PD}}$ (P_q, D_j) ≤ 1 means almost confirmation of the conclusion of D_j; $\varepsilon \le \mu^t_{R_{PD}}$ (P_q, D_j) < 0.6 means possible confirmation of the conclusion of D_j; $-\varepsilon < \mu^t_{R_{PD}}$ (P_q, D_j) $< \varepsilon$ means "unknown" about confirmation of the conclusion of D_j; $-0.6 < \mu^t_{R_{PD}}$ (P_q, D_j) $\le -\varepsilon$ means possible exclusion of the conclusion of D_j; $-1 < \mu^t_{R_{PD}}$ (P_q, D_j) ≤ -0.6 means almost exclusion of the conclusion of D_j; $\mu^t_{R_{PD}}$ (P_q, D_j) = -1 means absolute exclusion of the conclusion of D_j.

Where ε is a heuristic value. Let's recall that D_j consists of four types of depressive disorder including light depressive disorder, middle depressive disorder, serious depressive disorder, and serious depressive disorder with mental disorder.

Example 2. An example of illustrating the calculation of STRESSDIAG. Assuming that patient P_q has symptoms S = {S_1, S_2, S_3, S_4, S_5, S_6} with the occurrence of symptoms is, respectively, as follows:

$$\mu_{R_{PS}}(P_q, S_1) = 0.87 ; \mu_{R_{PS}}(P_q , S_2)= 0.92 ; \mu_{R_{PS}}(P_q , S_3)= 0.83$$

$$\mu_{R_{PS}}(P_q , S_4)= 0.94 ; \mu_{R_{PS}}(P_q , S_5)= 0.78 ; \mu_{R_{PS}}(P_q , S_6)= 0.83$$

The diagnostic process will take place in the following steps:

Step 1: List all the rules whose premise is a subset of the set S.

Step 2: Group the rules with the same disease conclusion. Supposing that, with the conclusion that disease = D3, we can group a set of rules as follows:

- Rule 1: IF S_1 THEN D_3, 0.3
- Rule 2: IF S_2 THEN D_3, 0.25
- Rule 3: IF S_4 THEN D_3, 0.35
- Rule 4: IF S_5 THEN D_3, 0.17
- Rule 5: IF $S_1 \wedge S_3 \wedge S_5$ THEN D_3, 0.76
- Rule 6: IF $S_2 \wedge S_4$ THEN D_3, 0.33
- Rule 7: IF $S_1 \wedge S_3$ THEN D_3, 0.39
- Rule 8: IF $S_1 \wedge S_2 \wedge S_5$ THEN EXCLUDE D_3, 0.23
- Rule 9: IF $S_3 \wedge S_4$ THEN EXCLUDE D_3, 0.23
- Rule 10: IF S_6 THEN EXCLUDE D_3, 0.34

Step 3: With the group of rules for positive disease D_3

- Rule 1: IF S_1 THEN D_3, 0.3
- Rule 2: IF S_2 THEN D_3, 0.25
- Rule 3: IF S_4 THEN D_3, 0.35

- Rule 4: IF S_5 THEN D_3, 0.17
- Rule 5: IF $S_1 \wedge S_3 \wedge S_5$ THEN D_3, 0.76
- Rule 6: IF $S_2 \wedge S_4$ THEN D_3, 0.33
- Rule 7: IF $S_1 \wedge S_3$ THEN D_3, 0.39

We can calculate $\mu_{R_{PDrule_h}}(P_q, D_3)$ according to formula (4):

$$\mu_{R_{PDrule_h}}(P_q, D_j) = \min \{\mu_{R_{PS}}(P_q , S_i), \mu_{R_{SDrule_h}}(S_i, D_j)\}$$

In this example, i ={1, 2, 3, 4, 5}; j=3; t={1, 2, 3, 4, 5, 6, 7}

[Rule 1] $\mu_{R_{PDrule_1}}(P_q, D_3) = \min \{\mu_{R_{PS}} (P_q , S_1), \mu_{R_{SDrule_1}}(S_1, D_3)\}$

$= \min \{0.87; 0.3\} = 0.3$

[Rule 2] $\mu_{R_{PDrule_2}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q , S_2) , \mu_{R_{SDrule_2}}(S_2, D_3)\}$

$= \min \{0.92; 0.25\} = 0.25$

[Rule 3] $\mu_{R_{PD}rule_3}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q , S_4) , \mu_{R_{SDrule_3}}(S_4, D_3)\}$

$= \min \{0.94; 0.35\} = 0.35$

[Rule 4] $\mu_{R_{PDrule_4}}(P_q, D_3) = \min \{ \mu_{R_{PS}}(P_q , S_5\} , \mu_{R_{SDrule_4}}(S_5, D_3)\}$

$= \min \{0,78; 0,17\} = 0.17$

[Rule 5] $\mu_{R_{PDrule_5}}(P_q, D_3) = \min\{\mu_{R_{PS}}(P_q , S_1), \mu_{R_{PS}}(P_q , S_3), \mu_{R_{PS}}(P_q , S_5), \mu_{R_{SDrule_5}}(S_{1, 3, 5} , D_3)\} = \min\{0.87; 0.83; 0.78 ; 0.76 \} = 0.76$

[Rule 6] $\mu_{R_{PDrule_6}}(P_q, D_3) = \min \{\mu_{R_{PS}}(P_q , S_2), \mu_{R_{PS}}(P_q , S_4), \mu_{R_{SDrule_6}}(S_{2,4} , D_3)\}= \min\{0.92; 0.94; 0.33\} = 0.33$

[Rule 7] $\mu_{R_{PDrule_7}}(P_q, D_3) = \min\{\mu_{R_{PS}}(P_q , S_1), \mu_{R_{PS}}(P_q , S_3) , \mu_{R_{SDrule_7}}(S_{1, 3} , D_3)\}= \min\{0.87; 0.83; 0.39\} = 0.39$

$\mu_{R_{PD}}(P_q, D_3)$ can be calculated according to formula (5):

$$\mu_{R_{PD}}(P_q, D_j) = \max \{\mu_{R_{PDrule_1}}(P_q, D_j), \ldots, \mu_{R_{PDrule_p}}(P_q, D_j)\}$$

In this example j=3 ; t=1,.. 7.

$\mu_{R_{PD}}(P_q, D_3) = \mathrm{Max} \{\mu_{R_{PDrule_h}}(P_q, D_3)\}= \mathrm{Max} \{0.3; 0.25; 0.35; 0.17; 0.76; 0.33; 0.39\}= 0.76$

Step 4: With the group of rules for negative disease D_3

- Rule 8: IF $S_1 \wedge S_2 \wedge S_5$ THEN EXCLUDE D_3, 0.23
- Rule 9: IF $S_3 \wedge S_4$ THEN EXCLUDE D_3, 0.23
- Rule 10: IF S_6 THEN EXCLUDE D_3, 0.34

We can calculate $\mu^e_{R_{PDrule^e_t}}(P_q, D_3)$ according to formula (9):

$$\mu^e_{R_{PDrule^e_t}}(P_q, D_j) = \min \{\mu_{R_{PS}}(P_q, S_i) , \mu^e_{R_{SDrule^e_t}}(S_i, D_j)\}$$

In this example i =1…6; j=3; t= 8…10

[Rule 8] $\mu^e_{R_{PDrule^e_8}}(P_q, D_3) = \min\{\mu_{R_{PS}}(P_q, S_1), \mu_{R_{PS}}(P_q, S_2), \mu_{R_{PS}}(P_q, S_5), \mu^e_{R_{SDrule_8}}(S_{1,2,5}, D_3)\} = \min\{0.87; 0.92; 0.78; 0.23\} = 0.23$

[Rule 9] $\mu^e_{R_{PDrule^e_9}}(P_q, D_3) = \min\{\mu_{R_{PS}}(P_q, S_3), \mu_{R_{PS}}(P_q, S_4), \mu^e_{R_{SDrule^e_9}}(S_{3,4}, D_3)\}$

$= \min\{0.83; 0.94;.0.23\} = 0.23$

[Rule 10] $\mu^e_{R_{PDrule^e_{10}}}(P_q, D_3) = \min\{\mu_{R_{PS}}(P_q, S_6), \mu^e_{R_{SDrule^e_{10}}}(S_6, D_3)\}$

$= \min\{0.83; 0.34\} = 0.34$

Calculate$\mu^e_{R_{PD}}(P_q, D_3)$depend on formula(10)

$$\mu^e_{R_{PD}}(P_q,D_j) = \max\{\mu^e_{R_{PDrule^e_1}}(P_q,D_j), \ldots, \mu^e_{R_{PDrule^e_k}}(P_q,D_j)\}$$

In this example j=3; t={8, 9, 10}

$\mu^e_{R_{PD}}(P_q,D_3) = \max\{\mu^e_{R_{PDrule^e_t}}(P_q,D_3)\} = \max\{0.23; 0.23; 0.34\} = 0.34$

Step 5: From $\mu_{R_{PD}}(P_q,D_3)$ and $\mu^e_{R_{PD}}(P_q,D_3)$, we can calculate $\mu^t_{R_{PD}}(P_q,D_3)$ according to formula (11):

$$\mu^t_{R_{PD}}(P_q, D_j) = \frac{\mu_{R_{PD}}(P_q,D_j)+(-\mu^e_{R_{PD}}(P_q,D_j))}{1-\min\{|\mu_{R_{PD}}(P_q,D_j)|,|\mu^e_{R_{PD}}(P_q,D_j)|\}}$$

$\mu^t_{R_{PD}}(P_q,D_3) = (0.76 - 0.34) / (1 - \min\{|0.76|, |-0.34|\})$

$= (0.76 - 0.34) / (1 - 0.34) = 0.63$

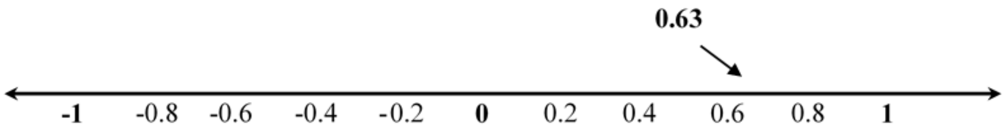

Figure1. Description of the results of the illustration of STRESSDIAG.

Step 6: Similar calculation with other diseases D_j in steps 2 to 5.

Step 7: Make a final conclusion that patient P_q has serious depressive disorder with a fuzzy degree of 0.63, that is, almost certainly patient P_q has serious depressive disorder.

3.3. Experimental Results

In these tests, with a data set of 264 medical records as tested in section 2, the diagnostic results of STRESSDIAG were compared with the medical records, giving details as shown in the table below.

The above experimental results show that STRESSDIAG gives good results for 4 depressive disorder types and without depressive disorder. The average rate of correct diagnosis exceeds 82 %, meaning nearly 60 % more than PORUL.DEP. In particular, STRESSDIAG correctly diagnosed middle depressive disorder, serious depressive disorder, and serious depressive disorder with mental disorder.

Table 3. Comparison of diagnostic results for each type of depressive disorder (positive rules and negative rules).

Type of depressive disorder	total	In medical records	STRESSDIAG	rate
Light depressive disorder	48	48	46	95.8 %
Middle depressive disorder	60	60	46	80 %
Serious depressive disorder	50	50	38	76 %
Serious depressive disorder with mental disorder	86	86	68	81.4 %
without depressive disorder	20	20	19	95 %

Table 4. Comparison of diagnostic results between the expert system software and the medical record (positive rules and negative rules).

Total	In medical records	STRESSDIAG	rate
264	264	217	82.19 %

5. CONCLUSION

The paper proposed an improvement of the PORUL.DEP expert system by adding and modifying the knowledge base of negative rules and improving the inference engine based on fuzzy relations and Abelian group operation in the MYCIN medical expert system. Abelian group operation ($\oplus$) of Mycin is used to provide good diagnostic results for all types of depressive disorders. The experimental results show that the proposed STRESSDIAG medical expert system gives good results for 4 depressive disorder types and without depressive disorder and improves the accuracy of diagnosing compared with the traditional PORUL.DEP expert system.

To achieve better diagnostic results, it requires time and expertise by repeated "trial and error" to determine the complete values and functions for each specific problem. This is also a limitation of building the knowledge base and inference engine in medical diagnostic specialist systems. In the coming time, the author will continue to research and improve the problem, considering the importance of symptoms for depressive disorders.

CRediT authorship contribution statement. Mai Thi Nu: made the survey, drafted the manuscript. Nguyen Hoang Phuong: conceivedthe study, edited, reviewed, and gave guidance on the theoretical and mathematical issues and contexts. Both authors provided critical feedback and corrections.

Declaration of competing interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. https://www.who.int/mental_health/management/depression/en/.

2. Olawale O. O., Francis A. O., Abasiubong F., Adebayo R. E. - Detection of mental disorders with the Patient Health Questionnaire in primary care settings in Nigeria, J. Mental Illness **2** (2010), doi: 10.4081/mi.2010.e10.
3. Abu Naser S. S., Shaath M. Z. - Expert system urination problems diagnosis, World Wide Journal of Multidisciplinary Research and Development **2** (2016) 9-19.
4. Salman F. M., AbuNaser S. S. - Expert System for Castor Diseases and Diagnosis, International Journal of Engineering and Information Systems (IJEAIS) **3** (2019) 1-10.
5. Supriya T., Neha G., Nitin P. - AI Based expert system to Aid patients with depresion disorder, International Journal of Latest Trends in Engineering and Technology **8** (2017) 18-23.
6. Yanase J., Triantaphyllou E. - Seven Key Challenges for the Future of Diagnostics in Computer Aided Medicine, International Journal of Health Informatics **129** (2019) 413-422.
7. AbuNasser B. - Medical Expert Systems Survey, International Journal of Engineering and Information Systems (IJEAIS) (2017) 218-22.
8. Oguoma S., Uka K., Chukwu C., Nwaoha E. - An Expert System for Diagnosis and Treatment of Mental Ailment, Open Access Library Journal **7** (2020) 1-22.
9. Victor E., Ekong, Uyinomen O., Uwadiae, Enobakhare E., Emmanuel A. - A fuzzy infeference system for predicting depresion risk levels, African Journal of Mathematics and Computer Science Research **6** (2013) 197-204.
10. Hamer M., Fraser-Smith N., Lesperance F. - Depressive Symptoms and 24-Hour Ambulatory Blood Pressure in Africans: The SABPA Study, Int. Journal of Hypertension **6** (2012) doi: 10.1155/2012/426803
11. Mai Thi Nu, Nguyen Hoang Phuong, Hirota K. - Modeling a Fuzzy Rule Based Expert System combining Positive and Negative Knowledge for Medical Consultations using the importance of Symptoms, In Proc. of IFSA-SCIS’2017 **164** (2017).
12. Hildrum B., Romild U., Holman J. - Anxiety and depression lowers blood pressure: 22 year follow up of the population based HUNT study, Norway, BMC Journal of Public Health (2011), doi: 10.1186/1471-2458-11-601.
13. Abu Naser S., El-Hissi H.,AbuRass M., El-Khozondar N. - An expert system for endocrine diagnosis and treatments using JESS, Journal of Artificial Intelligence – Scialert **3** (2010) 239-251.
14. Selvachandran G., Quek S. G., Lan L. T. H., Son L. H., Giang N. L., Ding W., Abdel-Basset M., Albuquerque V. H. - A New Design of Mamdani Complex Fuzzy Inference System for Multi-attribute Decision Making Problems”, IEEE Transactions on Fuzzy Systems **29** (4) (2021) 716-730.
15. Son L. H., Ngan R. T., Mumtaz Ali, Hamido Fujita, Mohamed Abdel-Basset,Giang N. L., Gunasekaran Manogaran, Priyan M. K. - A New Representation of Intuitionistic Fuzzy Systems and Their Applications in Critical Decision Making, IEEE Intelligent Systems **35** (1) (2020) 6-17.
16. Lan L. T. H., Tuan T. M., Ngan T. T., Son L. H., Giang N. L., Ngoc V. T. N, Hai P. V. - A New Complex Fuzzy Inference System with Fuzzy Knowledge Graph and Extensions in Decision Making, IEEE ACCESS **8** (2020) 164899-164921.

17. Tuan T. M., Lan L. T. H., Chou S. Y., Ngan T. T., Son L. H., Giang N. L., Mumtaz Ali - M-CFIS-R: Mamdani Complex Fuzzy Inference System with Rule Reduction Using Complex Fuzzy Measures in Granular Computing, Mathematics **8** (5) (2020) 707-731.
18. Phuong H. T., Giang N. L. - Fuzzy Distance based Filter-Wrapper Incremental Algorithms for Attribute Reduction when Adding or Deleting Attribute Set, Vietnam Journal of Science and Technology **59** (2) (2021) 261-274.