Wahana Matematika dan Sains: Jurnal Matematika, Sains, dan Pembelajarannya,
Vol. 14 No 1, April 2020
e-ISSN: 2549-6727 , p-ISSN: 1858-0629

Expert System to Diagnose Diseases of Mental Health with Forward Chaining and Certainty Factor

N K Ariasih[1] and I G A A D Indradewi[2]

[1,2] Informatics Engineering Department, STMIK STIKOM Indonesia, Jl. Tukad Pakerisan No. 97 Denpasar, Indonesia

Corresponding author: kdariasih@stiki-indonesia.ac.id

Abstract

The latest developments in the medical world use computers to help diagnoses and treat diseases. This study aims to develop an expert system which can be used to diagnose mental health problem, especially in the category of depression. User can diagnose based on the symptoms felt by patients. The knowledge representation used in this study is the production rule. The inference method used to get conclusions is using Forward Chaining the Certainty Factor method, and the platform used is a web-based system. After analyzing and applying the Forward Chaining and calculation of Certainty Factor method into the expert system, the conclusion is that the Forward Chaining and Certainty Factor methods can solve the problem of depression psychiatric diagnosis by prioritizing the value of the certainty of users and experts so that the value is close to certainty from diagnoses that are made to psychiatric illness depression.

Keywords: *Expert System; Certanity Factor; Foward Chaining; Diagnoses diseases*

INTRODUCTION

Mental health is an integral part of health and is a condition that allows the optimal physical, mental and social development of an individual, and which is in harmony with the development of others. But it cannot be denied with the development of the current era, many things can affect one's mental health so that it can cause a disorder or a disease. Depression is a condition of mental disorders that causes feelings of sadness and loss of interest that persists. Disruption of one's feelings, ways of thinking and behaving, as well as the emergence of various emotional and physical problems are the effects that arise due to someone suffering from depression. Although the causes of depression are still quite difficult to ascertain, the symptoms of someone who has depression can be known and studied. Consulting with an expert such as a psychiatrist is one way to find out whether symptoms that arise from a person can be diagnosed as depression or not. In 2017, World Health Organization (WHO) states that according to the latest estimates, more than 300 million people are now living with depression, an increase of more than 18% between 2005 and 2015. Lack of support for people with mental disorders, coupled with a fear of stigma, prevent many from accessing the treatment they need to live healthy, productive lives. At its worst, depression can lead to suicide. Close to 800.000 people die due to suicide every year. Suicide is the second leading cause of death in 15-29-year-olds [1].

As science develops, the expertise of a psychiatrist can be transferred into a form of technology called an expert system. Expert systems are computer-based systems that use knowledge, facts, and reasoning techniques in solving problems that can usually only be solved by an expert in the field [2]. This of course will make it easier for someone who wants to know whether the symptoms he feels or feels by those around him are depressed or not. Knowing someone's depression earlier will certainly help in handling it further, and this can be one of the precautions so that things that are harmful or undesirable from someone who has symptoms of depression do not occur.

Expert systems can be built using various types of reasoning methods. One of the reasoning method that can be used is Forward Chaining, which is a method of reasoning based on existing facts (data driven) to draw conclusions. In the health sector, Forward Chaining is used to build an expert system that can diagnose diseases in children under five years old (toddlers). This expert system was built through three stages, namely the collection of data and information from the Manajemen Terpadu Balita Sakit (Integrated Management of Toddler Sickness) and midwives, the creation of rules for diseases, and the implementation of an expert system with features of disease diagnosis, history of diagnosis, and providing advice on handling [3]. An expert system must be able to work in uncertainty. One method that can be used to resolve uncertainty problems is the certainty factor. Research related to the use of certainty factors in expert systems, one of which is research conducted by [4], which is an expert system for diagnosing cholesterol in adolescents. In this study the system is able to display the amount of confidence in the symptoms of the possibility of illness by the user.

Based on the two previous studies that have been presented, then in this study, the expert system that was built will use the method of forward chaining and certainty factors so that the expert system that will be generated can overcome the uncertainty problem of the diagnosis of depression. Considering so many problems regarding depression, this research is only limited to the scope of vegetative depression, dysrhythmic depression, agitation depression, and psychotic depression.

METHODS

A. Forward Chaining

Forward chaining method is a reasoning that starts from the facts to get conclusions (conclusions) from these facts [5]. This reasoning is based on existing facts (data driven), where this method is carried out by gathering existing facts to draw conclusions. In other words, the process starts from the facts through the interface fact process towards a goal. This method is also called using the if-then rule where the premise (if) leads to a conclusion (then) [6].

An example of forward chaining using five rules is as follows:

R1 : IF (Y and D) THEN Z
R2 : IF (X and B and E) THEN Y
R3 : IF A THEN X
R4 : IF C THEN L
R5 : IF (L and M) THEN N
Facts: A, B, C, D, and E are of true value
Goal: Determine whether Z is true or false

Facts in the database that can trigger predetermined rules are A, if fact A is included in the rules it will trigger R3's rules and generate new facts in the form of fact X that is entered into the database. This step is followed by fact C which triggers R4 rule as a trait forward and will produce new facts into the database, namely fact L. In the database has new facts namely X and L so that the facts in the database become A, B, C, D, E, X, and L. Furthermore facts X, B, and E are included in the rules that trigger R2 and produce new facts, namely Y. The facts in the database become A, B, C, D, E, X, L, and Y. Furthermore the facts Y and D are entered into the rules and trigger R1 and produce new facts Z. At this point, the step stops because the destination sought, Z, have found.

B. Certainty Factor

According to Sutojo in [4], certainty factor was introduced by Shortliffe Buchanan in making MYCIN. Certainty factor is the value of clinical parameters given by MYCIN to show the amount of trust. In dealing with a problem often found answers that do not have full certainty. This uncertainty can be a probability or probability that depends on the outcome of an event. Uncertain results are caused by two factors, namely uncertain rules and uncertain user answers to questions raised by the system. This is very easy to see in the disease diagnosis system, where experts can not define the relationship between symptoms and their causes with certainty, and patients can not feel a symptom with certainty as well. Eventually, there were many possible diagnoses.

According to Sutojo in [4], the expert system must be able to work in uncertainty. A number of theories have been found to resolve uncertainties, including classical probability, Bayesian probability, Hartley theory based on classical sets, Shannon theory based on probability (Shannon theory based on probability), Dempster-Shafer theory (Dempster-Shafer theory), Zadeh's fuzzy theory (Zadeh's fuzzy theory), and certainty factor. The certainty factor is a method used to measure a person's beliefs.

Probability and Certainty Factor. Certainty factor is defined as in equation (1):

$$CF(H,E) = MB(H,E) - MD(H,E) \tag{1}$$

CF (H, E) is a certainty factor of the hypothesis H that is influenced by the evidence (evidence) E. The amount of CF ranges from -1 to 1. A value of -1 indicates absolute mistrust, while a value of 1 indicates absolute trust. MB (H, E) is a measure of increase in confidence (measure of increased belief) of H hypothesis that is influenced by symptoms of E. MD (H, E) is a measure of distrust (measure of increased disbelief) of hypothesis H which is influenced by symptoms of E.
The basic form of the certainty factor equation is a rule if E then H as shown as in equation (2):

$$CF(H,e) = CF(E,e) * CF(H,E) \tag{2}$$

CF (H, e) is a certainty factor hypothesis that is influenced by evidence e. CF (E, e) is a certainty factor of evidence E that is influenced by evidence e. CF (H, E) is a certainty factor hypothesis with the assumption of evidence known with certainty, namely when CF (E, e) = 1. If the evidence on the antecedent is known with certainty then the equation will become as in equation (3):

$$CF(E,e) = CF(H,E) \tag{3}$$

The value of CF (rule) is obtained from the expert's "term" interpretation which is changed to a certain CF value according to the Table 1:

Table 1. The Factor Of Certainty Value.

Uncertain Term	CF
Certainly not	-1.0
Almost certainly not	-0.8
Most likely not	-0.6
Probably not	-0.4
Do not know	-0.2 to 0.2

Probably	0.4
Most likely	0.6
Almost certain	0.8
Certainly	1.0

Calculation of Combined Certainty Factors. Suppose another rule also concludes the same hypothesis, but with a different certainty factor. The certainty factors of rules concluding the same hypothesis are calculated from the combining function for certainty factors defined as in equation (4),

$$CF_{combine}(CF_1, CF_2) = \begin{cases} CF_1 + CF_2(1 - CF_1) & both > 0 \\ \frac{CF_1 + CF_2}{1 - min(|CF_1|, |CF_2|)} & one < 0 \\ CF_1 + CF_2(1 + CF_1) & both < 0 \end{cases} \quad (4)$$

where the formula for $CF_{combine}$ used depends on wether the individual certainty factors are positive or negative. The combining function for more than two certainty factors is applied incrementally. That is, the $CF_{combine}$ is calculated for two CF values, and then the $CF_{combine}$ is combined using equation (4) with the third CF value, and so forth (Giarratano, 1998).

Design and Implementation

System Overview

Expert system to diagnose this category of depression requires knowledge and inference engine to diagnose the user's disease. This knowledge base contains the factors needed by the system. While the inference engine is used to analyze the factors entered by the user so that a conclusion can be found. The knowledge base needed by the system consists of disease symptoms, types of disease and therapy. Figure 1 is a general description of the expert system to diagnose diseases of mental health with Forward Chaining and Certainty Factor.

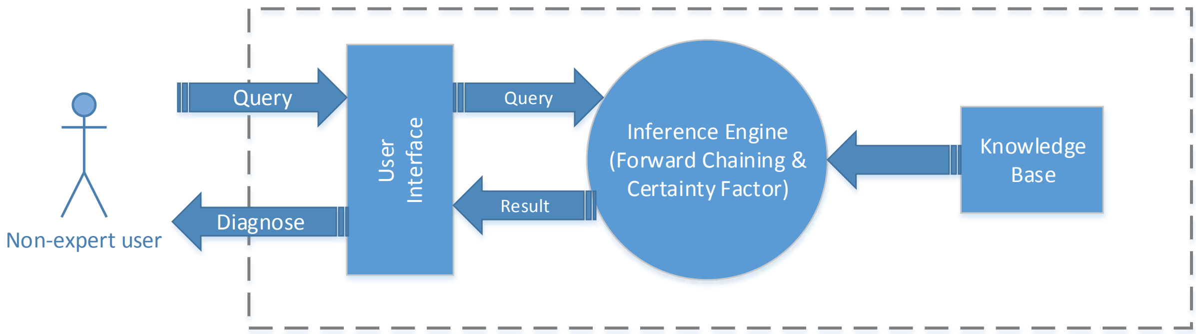

Figure 1. The Expert System Overview

Through the expert system interface, non-expert user answer a number of questions related to symptoms that appear or are felt. Answers in the form of symptoms felt by the non-expert user are accompanied by a certainty value for each symptom. The certainty value of each symptom felt by the non-expert user is processed by the Certainty Factor method, then the reasoning is processed by the Forward Chaining method. The knowledge base is the core of an expert system, which is in the form of knowledge representation from the experts. The knowledge base is composed of facts and rules. Fact is information about object, event and situation. Rule is a way to generate a new fact from facts that are already known. In the inference engine there is a process to manipulate and direct the rules, models, and facts stored in the knowledge base in order to reach a solution or conclusion.

Implementation

Implementation of the expert system is carried out by compiling a knowledge base that will later support the inference engine in generating diagnoses. Table 2 shows the list of depressive symptom. There are 23 list of symptoms related to vegetative depression, dysrhythmic depression, agitation depression, and psychotic depression. The weight of each symptom is determined by an expert based on the weighting conditions shown in Table 1.

Table 2. The Weight Of Depression Symptoms.

Symptom Code	Name of The Symptom	Expert Weight Value
G1	Before the menstrual cycle	-0.8
G2	There is no apetite	0.8
G3	Tend to be a loner	0.8
G4	Insomnia	0.8
G5	Feeling restless	0.8
G6	Wringing hands	0.4
G7	Talk inconsequential	-0.6
G8	Can not feel the emotions of grief	0.8
G9	Feeling a deep guilt	0.8
G10	Want to commit suicide	1
G11	Loss of the menstrual cycle	0.4
G12	Personality changes	0.4
G13	Looked grim	0.8

G14	Passive	0.8
G15	Feeling suspicious	0.6
G16	Like to criticize	-0.4
G17	Filled with negative thoughts	0.8
G18	Excessive fear will go crazy	0.8
G19	Experiencing delusions	0.4
G20	Experiencing hallucinations	0.4
G21	Feeling insecure	0.4
G22	Feel like being watched	0.4
G23	Feeling worthless	0.8

The input of the system is the symptoms that are felt by system users accompanied by how much the level of user confidence in the symptoms felt. The data is used by the system to determine the type of depression suffered by the user. In this case two types of rules are used: symptoms in men and symptoms in women. The formation of rules for depressive symptoms is shown in Table 3.

Table 3. The Symptom Rule Of Men.

Rule Code	Rule
R1	IF [G2] AND [G3] AND [G4] THEN **D1**
R2	IF [G2] AND [G5] AND [G6] AND [G7] AND [G8] AND [G9] AND [G10] THEN **D2**
R3	IF [G12] AND [G13] AND [G14] AND [G15] AND [G16] AND [G17] AND [G18] THEN **D3**
R4	IF [G19] AND [G20] AND [G21] AND [G22] AND [G23] THEN **D4**

For the rules of symptoms in women, a special rule is formed because according to experts, the tendency of women to be depressed is likely one of them is caused by changes in hormone levels that occur in the menstrual cycle. The rules of symptoms in women are shown in Table 4.

Table 4. The Symptom Rule Of Women.

Rule Code	Rule
R1	IF [G1] AND [G2] AND [G3] AND [G4] THEN **D1**
R2	IF [G2] AND [G5] AND [G6] AND [G7] AND [G8] AND [G9] AND [G10] AND [G11] THEN **D2**
R3	IF [G1] AND [G11] AND [G12] AND [G13] AND [G14] AND [G15] AND [G16] AND [G17] AND [G18] THEN **D3**
R4	IF [G19] AND [G20] AND [G21] AND [G22] AND [G23] THEN **D4**

The types of depression are shown in Table 5. These types of depression are related to the rules of symptoms in Table 4.

Table 5. The Types Of Depression.

Depression Code	Name of Depression
D1	Vegetative Depression
D2	Agitation Depression
D3	Dysrhytmic Depression
D4	Psychotic Depression

Based on the symptom rules that have been shown in Table 3 and Table 4, the decision tree from this expert system is shown in Figure 2. The decision tree search is carried out using the Forward Chaining method.

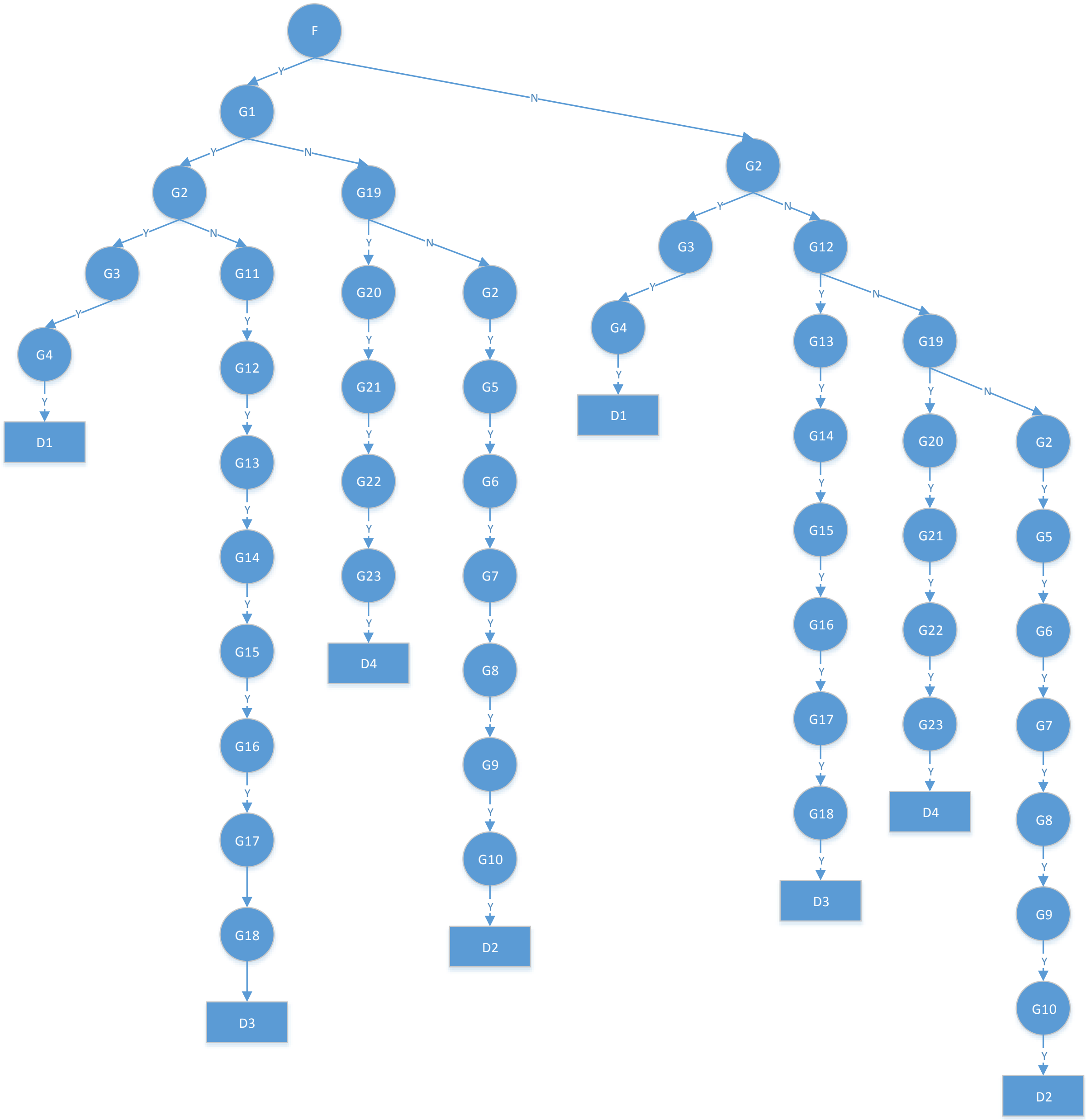

Figure 2. The Decision Tree Of Expert System.

RESULT AND DISCUSSION

System Interface

This expert system was built based on web technology, this was chosen to make it easier for users to access the system because it only requires an internet connection and a browser that is definitely available on a various devices and operating systems. The system access rights are divided into two consisting of administrator and non-expert user. The administrator is responsible for managing the data symptoms, weights, rules, and types of depression. The users who are non-expert users input symptoms that are felt through a collection of questions that are formed into a closed questionnaire. The emergence of questions related to symptoms felt by the user is based on tracing the symptoms in the decision tree shown by Figure 3. The disease consultation page in the expert system is shown in Figure 3.

Figure 3. The disease consultation page interface of non-expert user.

The consultation result page display is shown in Figure 4. The consultation results page displays the symptoms felt by the user accompanied by the certainty value of the symptoms. After symptoms, the next diagnosis is displayed with the value of certainty in accordance with the calculations performed with the Certainty Factor method. The results of the diagnosis are complemented by an explanation of the type of depression that is inferred.

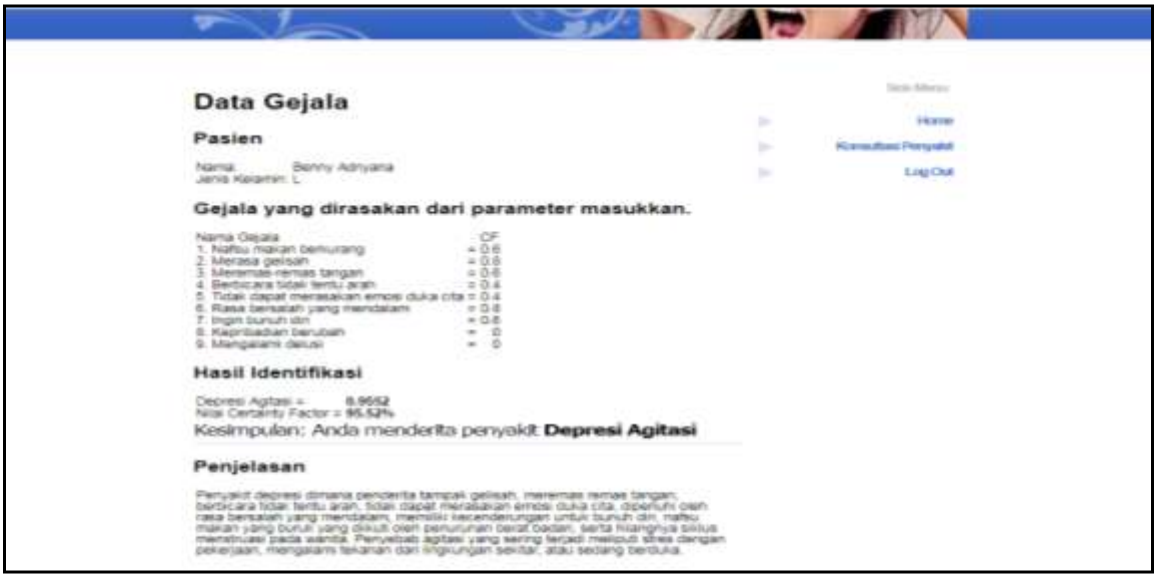

Figure 4. The Consultation Result Page.

Result

The case used to test the resulting expert system was a male user. The user feels symptoms like there is no apetite, feeling restless, wringing hands, talk inconsequential, can not feel emotional of grief, feeling a deep guilt, and want to commit suicide. The certainty value of each symptom felt by the user is shown in Table 6.

Table 6. The Certainty Value Of Symptoms Felt By The User.

No.	Name of The Symptom	Certainty Value from The User	Certainty Value from The Expert
1.	There is no apetite [G2]	0.6	0.8
2.	Feeling restless [G5]	0.8	0.8
3.	Wringing hands [G6]	0.6	0.4
4.	Talk inconsequential [G7]	0.4	-0.6
5.	Can not feel emotional of grief [G8]	0.4	0.8
6.	Feeling a deep guilt [G9]	0.8	0.8
7.	Want to commit suicide [G10]	0.6	1

The Forward Chaining search flow in the decision tree for the symptoms felt by the user is shown in Figure 5. Based on Figure 5 it is known that the user is diagnosed with agitation depression [D2]. To find out the level of certainty of a user suffering from agitation depression, the search results with Forward Chaining continued by calculating the level of certainty using Certainty Factor. The calculation of Certainty Factor is carried out in several steps.

The first step is to break the rule with multiple premise rules [R2] into single premise rules:

IF [G2] THEN D2
IF [G5] THEN D2
IF [G6] THEN D2
IF [G7] THEN D2
IF [G8] THEN D2
IF [G9] THEN D2
IF [G10] THEN D2

The second step is to calculate the value of CF_{expert} with CF_{user} using equation (2):

$$CF[H, e_2] = 0.6 * 0.8 = 0.48$$
$$CF[H, e_5] = 0.8 * 0.8 = 0.64$$

$CF[H,e_6] = 0.6 * 0.4 = 0.24$
$CF[H,e_7] = 0.4 * -0.6 = -0.24$
$CF[H,e_8] = 0.4 * 0.8 = 0.32$
$CF[H,e_9] = 0.8 * 0.8 = 0.64$
$CF[H,e_{10}]10 = 0.6 * 1 = 0.6$

The next step is to combine the CF value of each rule:

1. Combine $CF[H,e_2]$ dengan $CF[H,e_5]$ with equation (4) because the values of $CF[H,e_2]$ and $CF[H,e_5]$ are greater than 0 :

$CF_{COMBINE}(CF_2,CF_5) = CF_2 + CF_5 * (1 - CF_2)$, thus becoming
$CF_{COMBINE}(CF_2,CF_5) = 0.48 + 0.64 * (1 - 0.48)$
$= 0.48 + 0.3328$
$= 0.8128\ CF_{old}$

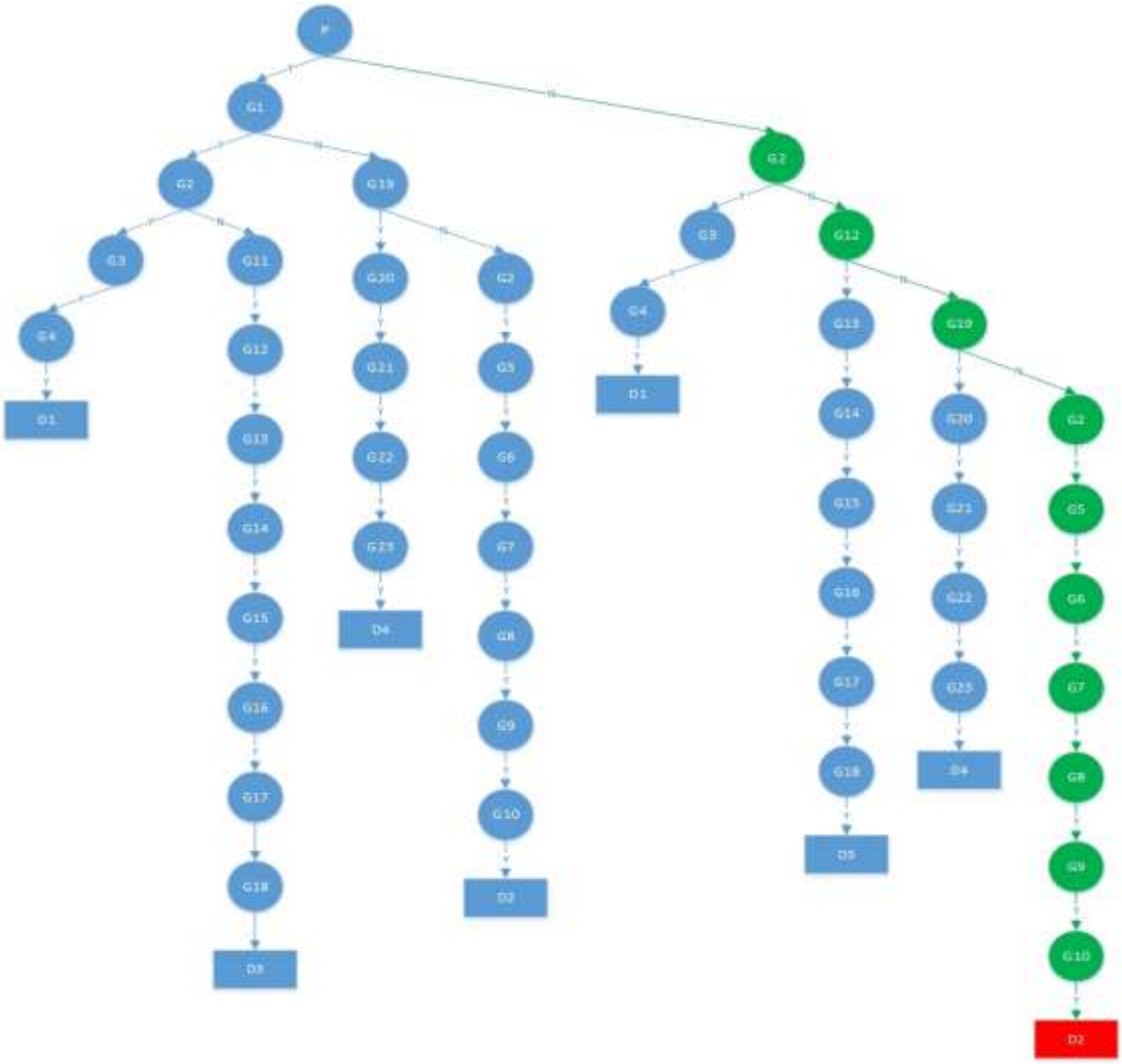

Figure 5. The Forward Chaining inferencing.

2. Combine CF_{old} dengan $CF[H,e_6]$ with equation (4) because the values of CF_{old} and $CF[H,e_6]$ are greater than 0 :

$CF_{COMBINE}(CF_{old},CF_6) = CF_{old} + CF_6 * (1 - CF_{old})$, thus becoming
$CF_{COMBINE}(CF_{old},CF_6) = 0.8128 + 0.24 * (1 - 0.8128)$
$= 0.8128 + 0.0449$

$= 0.8577\ CF_{old}$

3. Combine CF_{old} dengan $CF[H, e_7]$ with equation (4) because the values of CF_{old} and $CF[H, e_7]$ are less than 0 :

$$CF_{COMBINE}(CF_{old}, CF_7) = CF_{old} + CF_7/(1 - min(|CF_{old}|, |CF_7|)),\text{ thus becoming}$$
$$CF_{COMBINE}(CF_{old}, CF_7) = 0.8577 + (-0.24)/(1 - min(|0.8577|, |-0.24|))$$
$$= 0.8577 + (-0.24)/(1 - 0.24)$$
$$= 0.8577 + (-0.3158)$$
$$= 0.5419\ CF_{old}$$

4. Combine CF_{old} dengan $CF[H, e_8]$ with equation (4) because the values of CF_{old} and $CF[H, e_8]$ are greater than 0 :

$$CF_{COMBINE}(CF_{old}, CF_8) = CF_{old} + CF_8 * (1 - CF_{old}),\text{ thus becoming}$$
$$CF_{COMBINE}(CF_{old}, CF_8) = 0.5419 + 0.32 * (1 - 0.5419)$$
$$= 0.5419 + 0.1466$$
$$= 0.6885\ CF_{old}$$

5. Combine CF_{old} dengan $CF[H, e_9]$ with equation (4) because the values of CF_{old} and $CF[H, e_9]$ are greater than 0 :

$$CF_{COMBINE}(CF_{old}, CF_9) = CF_{old} + CF_9 * (1 - CF_{old}),\text{ thus becoming}$$
$$CF_{COMBINE}(CF_{old}, CF_9) = 0.6885 + 0.64 * (1 - 0.6885)$$
$$= 0.6885 + 0.1944$$
$$= 0.8879\ CF_{old}$$

6. Combine CF_{old} dengan $CF[H, e_{10}]$ with equation (4) because the values of CF_{old} and $CF[H, e_{10}]$ are greater than 0 :

$$CF_{COMBINE}(CF_{old}, CF_{10}) = CF_{old} + CF_{10} * (1 - CF_{old}),\text{ thus becoming}$$
$$CF_{COMBINE}(CF_{old}, CF_{10}) = 0.8879 + 0.6 * (1 - 0.8879)$$
$$= 0.8879 + 0.0673$$
$$= 0.9552\ CF_{old}$$

From the calculation above, it can conclude that percentage of confidence is 95.52%. The percentage means that certainty factor calculation for agitation depression has a system confidence level of 95.52%.

CONCLUSION AND FUTURE WORK

Based on the description above, it can be concluded that a prototype of expert system has successfully built. This prototype aim to diagnose type of depression by using the Forward Chaining and Certainty Factor method. The result of using both methods in expert system is a percentage of a user's confidence suffering from depression based on the weight of the symptoms given by the user. The prototype system will be improved by adjusting the weight for each symptom based on the expert of a psychiatrist because it will adjust to the treatment that will be given by the patient. The development of the system will continue to the stage of testing to find out whether the system is able to provide diagnoses in accordance with diagnosis from experts.

References

[1] https://www.who.int/news-room/detail/30-03-2017--depression-let-s-talk-says-who-as-depression-tops-list-of-causes-of-ill-health.

[2] Kusumadewi, Sri. 2003. Artificial Intelligence (Teknik dan Aplikasinya). Yogyakarta: Graha Ilmu.

[3] Yanto, Bagus Fery, Indah Werdiningsih, Endah Purwanti, April 2017. "Aplikasi Sistem Pakar Diagnosa Penyakit Pada Anak Bawah Lima Tahun Menggunakan Metode Forward Chaining". Journal of Information Systems Engineering and Business Intelligence 3 1:61-67.
[4] Sihotang, Hengki Tamando, Juni 2014. "Sistem Pakar Mendiagnosa Penyakit Kolesterol pada Remaja dengan Metode Certainty Factor (CF) Berbasis Web". Jurnal Manajemen dan Informatika Komputer Pelita Nusantara 15 1:16-23.
[5] Giarratano & Riley, 2004, "Expert System Principle and Programming (Third Edition)", ISBN 7-111-10844-2/TP.2586.
[6] Supartini, Windah & Hindarto, November 2016. "Sistem Pakar Berbasis Web dengan Metode Forward Chaining dalam Mendiagnosis Dini Penyakit Tuberkulosis di Jawa Timur". KINETIK 1 3:147-154.