Modification of Certainty Factor Method In Solving Expert System Problems

Natalia Silalahi¹, Nurmala Sridewi², Nelly Astuti Hasibuan¹, Guidio Leonarde Ginting¹ {natalia.silalahi@gmail.com}

¹Department Of Computer Science, AMIK STIEKOM, Sumatera Utara, Indonesia
 ²Department Of Computer Science, STMIK Logika, Sumatera Utara, Indonesia
 ³Department of Computer Science, STMIK Budi Darma, Medan, Indonesia

Abstract. The Certainty Factor method is used as a method of resolving the problem of certainty by calculating every possible value. The Certainty Factor method calculates the probable value of all the factors involved in a problem. The result of a certainty factor calculation is the result of a combination of the calculation of each factor. The combination stage of the certainty method is standard so that the calculation results will be greatly affected by the calculation of the first-factor value calculated. The certainty factor method is also known as the expert system solution method that has been used in many studies in diagnosing a problem. The influence of probability value of the first factor greatly influence the diagnosis result, therefore needed a research which discusses modifies at combination stage of method certainty factor.

Keywords. Modification, Certainty Factor, Combination, Premise

1 Introduction

Calculation of the value of certainty has been raised quite a lot in various studies in various fields of science. There are many techniques that can be used in solving uncertainty problems such as probability techniques, vague logic techniques, smoothing techniques for prediction and so on. The problem-solving technique of certainty is used to solve different assurance problems. The probability technique is used to calculate the value of certainty based on all possible possibilities [1]. The fuzzy logic technique solves the problem of certainty by explaining a value that is considered vague. While smoothing technique is usually used to solve prediction problems or estimates in the future will come based on facts that have today.

The expert system is one of the problems that the solution uses the technique of certainty. An expert system is used to replace experts in solving problems related to the science that they are working on. The expert system works by harnessing the facts of the event or event experienced by an object. An expert to draw the conclusions of an event or problem can be done by comparing the knowledge it possesses with the facts that exist in an event. The conclusion begins with the knowledge of a person adopted by the computer will be processed by building the inference engine, so that the formation of knowledge is called the rule and will be stored in a container called Knowledge Base. The next knowledge base will be processed by looking at the facts obtained from a problem or event resulting in various possibilities. Because of the many possibilities that arise when it is necessary to draw conclusions by

performing calculations to determine the value of the possibility, where every possibility that appears must have interrelationship or influence each other.

One of the certainty techniques that can be applied to solving expert system problems is the certainty factor method. The certainty factor method solves the problem of the expert system by performing the calculations by forming the premise based on the facts experienced by the object in a particular event, then the fact will be compared with the expert knowledge. Next, each premise will be combined with each other to get the final value or final conclusion value. Combining the premise is crucial to the results of diagnosis. The certainty factor method performs a continuous combination from the first premise to the final premise. However, the combining technique [2] does not take into account the value of each press. Therefore, in this research, the combination of the premise will be done by considering the value of each premise. Combining the premise with regard to the value of each premise will, of course, result in a different final or final conclusion.

2 Theory

2.1 Expert System

Expert systems are defined as a computer application built to make decisions such as decisions taken by one or more experts. Formation of rules in expert systems is done by combining inference rules with certain knowledge base given by one or more experts in a particular field [3] The combination of inference rules with the knowledge base provided by the expert is stored in the computer, which is then used in the decision-making process for specific problem-solving. The main purpose of the expert system is not to replace the position of a member or an expert, but only to promote the knowledge and experience of experts. Along with the growth of the human population, then in the future, the expert system is expected to be very helpful in decision making [4]. Edward Feigenbaum, a professor at Stanford University, described the expert system as an Intelligent Computer Program that utilizes knowledge and procedure inference to solve difficult problems that require special skills from humans. Based on the description above-described expert system is artificial intelligence [5] software that adopts the knowledge of experts in order to solve problems of expertise by generating decisions based on conclusions obtained. In expert systems, expert knowledge is stored in a knowledge base with a rule form. The expert knowledge rule is then called the hypothesis (H) where each hypothesis will have a certain value known as expert value. Next rule or hypothesis will be formed into questions that must be answered by the user [6]. The expert system will be used by the user to know the answer to the problem faced. Users are required to provide input into the system of facts experienced by the object of the problem. Facts experienced by the object of further problems will be referred to as evidence (E). The evidence is derived from the user's answer to questions built on a stored rule or hypothesis. Each user's answer will have a certain value or commonly called the user value [7]

Therefore, it can be concluded that in an expert system there are two (2) values that are considered to interfere with the problem-solving process to produce a conclusion or decision, namely hypothesis (expert value) and evidence (user value) [8] The conclusions of the expert system are formulated as follows:

 $K[H, E]_i = K[H]_i * K[E]_i$

Description: $K[H, E]_i$: Possible hypotheses affected evidence $K[H]_i$: Possible hypothesis $K[E]_i$: Possible Evidence

Hypothesis (expert value) and evidence (value of user) as described previously is a value of certainty so it requires a problem-solving technique of certainty or probability theory [6]

2.2 Certainty Factor

There is a method in the expert system application to solve the problem of data uncertainty. Factor certainty (certainty factor) is one method that can be used to solve the problem [9]. The belief factor was introduced by Buchanan Shortliffe in the making of MYCIN. Certainty factor (CF) is the value of clinical parameters given by MYCIN to indicate the magnitude of trust. There are 2 kinds of certainty factors used, namely hypothesis and evidence. A hypothesis is a certainty factor that is filled by the expert along with the rules. Evidence is a certainty factor provided by the user [10]. Theories are found to solve uncertainty such as classical probability, Bayesian probability, Hartley theory based on classical sets, Shannon theory based on probability, Dempster-Shafer theory, Zadeh's fuzzy theory and certainty factor (CF) is clinical parameter value given by MYCIN to show confidence level. Certainty factor is defined in equation 1 as follows [11]:

CF(H, E) = MB(H, E) - MD(H, E)

Description:

CF(H,E) : certainty factor from hypothesis H influenced by evidence E. CF value is from -1 to 1. Value of -1 shows absolute uncertainty whereas value of 1 shows absolute certainty.

MB(H,E): measure of increased belief to hypothesis H influenced by evidence E.

MD(H,E): measure of increased disbelief on hypothesis H influenced by evidence E McAllister

The hypothesis or factor of certainty filled by the expert describes the expert's belief in the relationship between antecedent and consequent. While the evidence or the certainty factor of the user shows the amount of trust in the existence of each element in the antecedent [12]. Each factor of good certainty obtained from the expert or user-given certainty factor will be translated or converted in the form of a certain value that is valued between 0 to 1. Factor certainty is converted into numbers by using the following table terminology [13]:

Uncertanty Term	CF	
Definitely Not	- 1.0	
Almost Definitely Not	- 0.8	
Most likely Not	- 0.6	
Probably not	- 0.4	
Do not know	- 0.2 to 0.2	
Maybe	0.4	
Most likely	0.6	

Fable 1.	Uncertainty	Term

Almost certainly	0.8
Certainly	1.0

Each factor is certainty of the form in a premise where each premise will be calculated by the following formula [14] :

1. For single premises If there are 2 (two) events p and q then:

 $CF[H, E]_{(i)} = CF[H]_{(i)} * CF[E]_{(i)}$

- 2. For a compound premise
 - a. If two conditions occur simultaneously then: $CF[A \land B] = Min(CF[a], CF[b]) * CF[rule]$
 - b. If 2 conditions do not occur simultaneously then: $CF[A \lor B] = Max(CF[a], CF[b]) * CF[rule]$
- 3. For Combined Premises

$$CF_{(Cmb)}[H,E]_{(i,j)} = CF[H,E]_{(i)} + CF[H,E]_{(j)} * (1-CF[H,E]_{(i)})$$

3 Result & Discussion

This research will modify certainty factor method where modification is done in the step of combining premise. For that the settlement will begin by completing a single premise that will be combined in accordance with the certainty factor method, then it will be recombined but with the modify used. Expert systems can be applied to calculate the value of disease diagnoses. Here's an example of a settlement step using an expert system:

 $\begin{array}{c} \mathrm{G1} = 0,2 \quad \clubsuit \quad H_1 \\ \mathrm{G2} = 0,6 \quad \clubsuit \quad H_2 \\ \mathrm{G3} = 0,4 \quad \clubsuit \quad H_3 \\ \mathrm{G4} = 0,6 \quad \clubsuit \quad H_4 \\ \mathrm{G5} = 0,8 \quad \clubsuit \quad H_5 \\ \mathrm{G6} = 0,4 \quad \clubsuit \quad H_6 \end{array}$

In a diagnosis a farmer gives answers to the hypothesis being asked as follows:

 $\begin{array}{c} \mathrm{G1} = 0,4 \quad \not \rightarrow E_1 \\ \mathrm{G2} = 0,2 \quad \not \rightarrow E_2 \\ \mathrm{G3} = 0,8 \quad \not \rightarrow E_3 \\ \mathrm{G4} = 0,6 \quad \not \rightarrow E_4 \\ \mathrm{G5} = 0,2 \quad \not \rightarrow E_5 \\ \mathrm{G6} = 0,4 \quad \not \rightarrow E_6 \end{array}$

By using certainty factor method, the diagnosis result is described as follows: 1. Calculation of the probable value of each premise

$$CF[H, E]_{(1)} = CF[H]_{(1)} * CF[E]_{(1)}$$

= 0,2 * 0,4
=0,08

 $CF[H, E]_{(2)} = CF[H]_{(2)} * CF[E]_{(2)}$

$$CF[H, E]_{(3)} = 0,6 * 0,2$$

=0,12
$$CF[H, E]_{(3)} = CF[H]_{(3)} * CF[E]_{(3)}$$

= 0,4 * 0,8
= 0,32

$$CF[H, E]_{(4)} = CF[H]_{(4)} * CF[E]_{(4)}$$

= 0,6 * 0,6
=0,36

$$CF[H, E]_{(5)} = CF[H]_{(5)} * CF[E]_{(5)}$$

= 0,8 * 0,2
= 0,16

$$CF[H, E]_{(6)} = CF[H]_{(6)} * CF[E]_{(6)}$$

= 0,4 * 0,4
= 0,16

- 2. Calculation of the premise value of the premise with the combination premise, the following steps to perform a combination of a premise:
 - a. The combined premise is calculated starting from premise 1 combined with premise 2 and the result is called the Old1 premise
 - b. The old1 premise will be recombined with premise3 using the same combination formula where the old1 premise replaces the 1st premise and the combination result is called the old2 premise
 - c. If there is still a 4th premise and so on, then re-combine the old premise with the next single premise.

Here is the scheme of the combination of premises:

$$CF[H, E]_{(1)}$$

$$CF[H, E]_{(2)}$$

$$CF[H, E]_{(3)}$$

$$CF[H, E]_{(4)}$$

$$CF[H, E]_{(5)}$$

$$CF[H, E]_{(6)}$$

$$CFold 4$$

The calculation of the combination premise is described as follows : $CF_{cmb}[H, E]_{(1,2)} = CF[H, E]_{(1)} + CF[H, E]_{(2)} * (1 - CF[H, E]_{(1)})$

5

= 0,08 + 0,12 * (1 - 0,08)= 0,08 + 0,12 * 0,92 = 0,08 + 0,11 = 0,19 => CFold 1 $CF_{cmb}[H,E]_{(old1,3)} = CF[H,E]_{(old1)} + CF[H,E]_{(3)} * (1 - CF[H,E]_{(old1)})$ = 0,19 + 0,32 * (1 - 0,19) = 0,19 + 0,32 * 0,81 = 0,19 + 0,26

=0,45 => CFold 2 $CF_{cmb}[H, E]_{(old2,4)} = CF[H, E]_{(0ld2)} + CF[H, E]_{(4)} * (1 - CF[H, E]_{(old2)})$ =0.45+0.36*(1-0.45)= 0.45 + 0.36 * 0.55= 0.45 + 0.19=0.64 => CFold 3 $CF_{cmb}[H,E]_{(old3,5)} = CF[H,E]_{(old3)} + CF[H,E]_{(5)} * (1 - CF[H,E]_{(old3)})$ = 0,64 + 0,16 * (1 - 0,64)= 0,64 + 0,16 * 0,36= 0,64 + 0,05=0,69 => CFold 4 $CF_{cmb}[H, E]_{(old4,6)} = CF[H, E]_{(old4)} + CF[H, E]_{(6)} * (1 - CF[H, E]_{(old4)})$ = 0,69 + 0,16 * (1 - 0,69)= 0,69 + 0,16 * 0,31= 0,69 + 0,049= 0,739 => CFold 5Diagnosis = CFold-n * 100%= 0,739 * 100%

3. Modification of certainty factor method

The modification to be performed on the certainty factor method is at the premisecombining stage. The combination of the premise of the certainty factor method is carried out continuously so that the premise of each premise will be interconnected or interrelated. But the combination technique of this method does not take into account or consider the high and low value of the premise. For that modification performed for the combination stage is divided into several types, among others:

a. Carry out premise based on premise value starting from the premise with the highest value to the premise with the lowest value. Here are the modifications.

Single premise value in the above case example:

 $CF[H, E]_{(1)} = 0,08$ $CF[H, E]_{(2)} = 0,12$ $CF[H, E]_{(3)} = 0,32$ $CF[H, E]_{(4)} = 0,36$ $CF[H, E]_{(5)} = 0,16$ $CF[H, E]_{(6)} = 0,16$

The premise will be sorted based on the value of the premise from the premise with the highest value to the premise with the lowest value and then followed by combining the premise

 $CF[H, E]_{(4)} = 0.36 \Rightarrow CF 1$ $CF[H, E]_{(3)} = 0.32 \Rightarrow CF 2$ $CF[H, E]_{(5)} = 0.16 \Rightarrow CF 3$ $CF[H, E]_{(6)} = 0.16 \Rightarrow CF 4$

 $CF[H, E]_{(2)} = 0.12 \implies CF 5$ $CF[H, E]_{(1)} = 0.08 => CF 6$ $CF_{cmb}[H, E]_{(1,2)} = CF1 + CF2 * (1 - CF1)$ =0,36+0,32*(1-0,36)= 0,36 + 0,32 * 0,64 =0,36+0,20=0,56 => CFold 1 $CF_{cmb}[H,E]_{(old1,3)} = CFold1+CF3*(1-CFold1)$ =0,56+0,16*(1-0,56)= 0,56 + 0,16 * 0,44=0,56+0,07=0,63 => CFold 2 $CF_{cmb}[H,E]_{(old2,4)} = CFold2+CF4*(1-CFold2)$ = 0,63 + 0,16 * (1 - 0,63)= 0,63 + 0,16 * 0,37 = 0,63 + 0,059=0,689 => CFold $CF_{cmb}[H,E]_{(old3.5)} = CFold3+CF5*(1-CFold3)$ = 0,689 + 0,12 * (1 - 0,689)= 0,689 + 0,12 * 0,311= 0,689 + 0,037=0,726 => CFold 4 $CF_{cmb}[H,E]_{(old4,6)} = CFold4+CF6*(1-CFold4)$ =0,726+0,08*(1-0,726)= 0,726 + 0,08 * 0,274= 0,726 + 0,021=0,747 => CFold 5

Diagnosis = CFold-n * 100% = 0.747 * 100%

$$= 0,747 \cdot 10$$

= 74,7%

b. Carry out premise based on premise value Starting from the premise with the lowest value to the highest rated premise. Here are the modifications.

Single premise value in the above case example:

 $CF[H, E]_{(1)} = 0,08$ $CF[H, E]_{(2)} = 0,12$ $CF[H, E]_{(3)} = 0,32$ $CF[H, E]_{(4)} = 0,36$ $CF[H, E]_{(5)} = 0,16$ $CF[H, E]_{(6)} = 0,16$

The premise will be sorted based on the value of the premise starting from the premise with the lowest value to the premise with the highest value and then followed by combining premise

 $CF[H, E]_{(1)} = 0.08 \implies CF 1$

 $CF[H, E]_{(2)} = 0.12 \implies CF 2$ $CF[H, E]_{(5)} = 0.16 \implies CF 3$ $CF[H, E]_{(6)} = 0.16 \implies CF 4$ $CF[H, E]_{(3)} = 0.32 \implies CF 5$ $CF[H, E]_{(4)} = 0.36 \implies CF 6$ $CF_{cmb}[H, E]_{(1,2)} = CF1 + CF2 * (1 - CF1)$ = 0.08 + 0.12 * (1 - 0.08)= 0.08 + 0.12 * 0.92= 0.08 + 0.11=0,19 => CFold 1 $CF_{cmb}[H,E]_{(old1,3)} = CFold1+CF3*(1-CFold1)$ = 0,19 + 0,16 * (1 - 0,19)= 0,19 + 0,16 * 0,81 = 0,19 + 0,129=0,319 => CFold 2 $CF_{cmb}[H,E]_{(old2,4)} = CFold2+CF4*(1-CFold2)$ = 0,319 + 0,16 * (1 - 0,319)= 0,319 + 0,16 * 0,681=0,319+0,108= 0,427 => CFold 3 $CF_{cmb}[H,E]_{(old3.5)} = CFold3+CF5*(1-CFold3)$ = 0,427 + 0,32 * (1 - 0,427)=0,427+0,32*0,573= 0,427 + 0,183= 0,61 => CFold 4 $CF_{cmb}[H,E]_{(old4,6)} = CFold4+CF6*(1-CFold4)$ = 0,61 + 0,36 * (1 - 0,61)= 0,61 + 0,36 * 0,39=0,61+0,14=0,75 => CFold 5Diagnosis = CFold-n * 100% = 0,75 * 100%

Doing paired combinations by pairing each premise then each pair of premises will be combined and the results combine again

$$\begin{array}{cccc} CF[H, E]_{(1)} \\ CF[H, E]_{(2)} \\ CF[H, E]_{(3)} \\ CF[H, E]_{(4)} \end{array} \end{array} \xrightarrow{CFold} CFold \\ CF[H, E]_{(5)} \\ CF[H, E]_{(6)} \end{array} \xrightarrow{CFold} CFold \end{array}$$

= 75%

$$\begin{aligned} CF_{cmb}[H, E]_{(1,2)} &= CF1 + CF2 * (1 - CF1) \\ &= 0,08 + 0,12 * (1 - 0,08) \\ &= 0,08 + 0,11 \\ &= 0,19 => CFold 1 \\ CF_{cmb}[H, E]_{(3,4)} &= CF3 + CF4 * (1 - CF3) \\ &= 0,32 + 0,36 * (1 - 0,32) \\ &= 0,32 + 0,36 * 0,68 \\ &= 0,32 + 0,24 \\ &= 0,56 => CFold 2 \\ CF_{cmb}[H, E]_{(5,6)} &= CF5 + CF6 * (1 - CF5) \\ &= 0,16 + 0,16 * (1 - 0,16) \\ &= 0,16 + 0,16 * (1 - 0,16) \\ &= 0,16 + 0,13 \\ &= 0,29 => CFold 3 \\ CF_{cmb}[H,E]_{(old1,old2)} &= CFold1 + CFold2*(1 - CFold1) \\ &= 0,19 + 0,56 * (1 - 0,19) \\ &= 0,19 + 0,56 * (0,81) \\ &= 0,19 + 0,56 * (0,81) \\ &= 0,64 => CFold 4 \\ CF_{cmb}[H,E]_{(old4,old3)} &= CFold4 + CFold3*(1 - CFold4) \\ &= 0,64 + 0,29 * (0,36) \\ &= 0,64 + 0,29 * (0,36) \\ &= 0,64 + 0,104 \\ &= 0,744 => CFold 7 \\ Diagnosis &= CFold-n * 100\% \\ &= 0,744 * 100\% \\ &= 74,4\% \end{aligned}$$

4 Conclusion

From the results of the above analysis, it can be concluded that:

- 1. The combination of premises used in the certainty factor method is continuous between one premise and another premises but does not pay attention to the value of premises
- 2. Modification of the combination steps on the certainty factor method by doing the ordering of the premise value from the premise with the highest value up to the premise with the lowest value and vice versa.
- 3. Modification is also done by forming a paired premise where each pair of premises will be combined until the al premise is obtained

4. The modification of a combination step on a certainty factor method results in a value that is not very different, or it can be said that the value of the modified result tends to be equal to the value obtained by the original combination step

Reference

- R. Rahim, I. Zulkarnain, and H. Jaya, "Double hashing technique in closed hashing search process," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 237, p. 012027, Sep. 2017.
- [2] R. Rahim, I. Zulkarnain, and H. Jaya, "A review: search visualization with Knuth Morris Pratt algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 237, p. 012026, Sep. 2017.
- [3] Roni Pambudi & Sumarno, "Aplikasi Sistem Pakar Diagnosa Penyakit Kanker Menggunakan Metode Certainty Factor," J. Ilm. Tek. Inf., no. Sistem Pakar, p. IF-64-IF-65, 2015.
- [4] L. A. Latumakulita, "Sistem Pakar Pendiagnosa Penyakit Anak Menggunakan Certainty Factor (Cf)," J. Ilm. Sains, vol. 12, no. 2, p. 120, 2012.
- [5] D. Siregar, D. Arisandi, A. Usman, D. Irwan, and R. Rahim, "Research of Simple Multi-Attribute Rating Technique for Decision Support," J. Phys. Conf. Ser., vol. 930, no. 1, p. 012015, Dec. 2017.
- [6] N. A. Hasibuan, K. Yusmiarti, F. T. Waruwu, and R. Rahim, "Expert systems with genetics probability," *Int. J. Res. Sci. Eng.*, vol. 3, no. 2, pp. 112–116, 2017.
- [7] I. Rasal, "Rule-Based Expert System for Diagnosing Toddler," *Rule-Based Expert Syst.*, pp. 215–220, 2011.
- [8] D. T. Yuwono, A. Fadlil, M. T. Informatika, U. Ahmad, and D. Yogyakarta, "PENERAPAN METODE FORWARD CHAINING DAN CERTAINTY FACTOR PADA SISTEM PAKAR," vol. 04, no. 02, pp. 136–145, 2017.
- P. Lucas, "Certainty-factor-like structures in Bayesian networks," *Lect. Notes Comput. Sci.* (*including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics*), vol. 1792 LNAI, pp. 25–36, 2000.
- [10] A. Niederli, "A Modification of the Stanford Certainty Factor Algebra for Uncertain Expert Systems A. Niederli ń ski," *AI-METH*, pp. 1–5, 2004.
- [11] P. T. Sharma, P. N. Tiwari, and P. K. Shah, "Student Counseling System : A Rule-Based Expert System based on Certainty Factor and," *Int. J. Appl. or Innov. Eng. Manag.*, *ISSN 2319 - 4847*, vol. 2, no. 1, pp. 371–375, 2013.
- [12] R. R. Fanny, N. A. Hasibuan, and E. Buulolo, "Perancangan Sistem Pakar Diagnosa Penyakit Asidosis Tubulus Renalis Menggunakan Metode Certainty Factor Dengan Penulusuran Forward Chaining," *Media Inform. Budidarma*, vol. 1, no. 1, pp. 13–16, 2017.
- [13] S. D. Taksiah, N. A. Hasibuan, I. Saputra, and Fadlina, "Sistem pakar diagnosa bayi terlahir cacat dengan menggunakan teorema bayes," *Media Inform. Budidarma*, vol. 1, no. 2, pp. 38–41, 2017.
- [14] I. K. G. D. Putra and P. M. Prihatini, "Fuzzy Expert System for Tropical Infectious Disease by Certainty Factor," *Telkomnika*, vol. 10, no. 4, pp. 825–836, 2012.