RESEARCH OF UNCERTAINTY REASONING IN PINEAPPLE DISEASE IDENTIFICATION SYSTEM

Liqun Liu [1,*], Haifeng Fan [2]

[1] *Ministry of Personnel, Guangdong Ocean University, Zhanjiang, Guangdong Province, P.R. China 524088*

[2] *Navy south China sea fleet command post, Zhanjiang, Guangdong Province, P.R. China 524055*

[*] *Corresponding author, Address: Ministry of Personnel, Guangdong Ocean University, Zhanjiang 524088, Guangdong Province, P. R. China, Tel: +86-759-2383115, Fax: +86-759-2383115, Email: gdoullq@126.com*

Abstract: In order to deal with the uncertainty of evidences mostly existing in pineapple disease identification system, a reasoning model based on evidence credibility factor was established. The uncertainty reasoning method is discussed, including: uncertain representation of knowledge, uncertain representation of rules, uncertain representation of multi-evidences and update of reasoning rules. The reasoning can fully reflect the uncertainty in disease identification and reduce the influence of subjective factors on the accuracy of the system.

Keywords: pineapple, identification system, certainty factor, uncertainty reasoning

1. INTRODUCTION

Pineapple is one of the important fruits in the tropical and subtropical regions of China. There are large areas of cultivation, mainly concentrated in Guangdong, Guangxi, Fujian and Taiwan. The cultivation of pineapple is attacked by diseases frequently, causing a decline in output and quality. It is particularly important to create a disease identification system for pineapple. However, in the process of establishing the system, the pineapple disease lacks a clear description, because experts often use some ambiguous language to describe the symptom of a particular disease, such as: "shorter" and "smaller" etc., which leads to the uncertainty of knowledge reasoning in the system (Zhang Wen et al., 2008). It is difficult to solve such problems if certainty theory and method is used. Therefore, it will adopt uncertainty reasoning method to solve such problems.

2. REASONING MODEL

2.1 Analysis of reasoning problem

Uncertainly reasoning based on conclusions, in practice, cannot avoid some problems as the followings:(1) the premise of application demands the sub conditions in condition set are relatively independent; (2) if the certain degree of general assumption condition is 1 or 0, when a disease performance is atypical or when users' subjective understanding is uncertain, human errors increase, which will increase subjective influence on reasoning results (Zhang Xuenong, 2005).

2.2 Uncertain representation of knowledge

The reasoning that is based on the uncertainty knowledge is defined as uncertainty reasoning. In this system, the production rule with a certainty factor (*CF*) is adopted for the uncertain representation of knowledge (Yang Jian, 2003):

IF *E* THEN *H(CF(H, E))*

E stands for the rules preconditions which can be a simple single precondition, or a complex precondition, i.e. a complex precondition derived from conjunction and extraction. *H* stands for the rules conclusions which can be a single conclusion, or conclusions.

CF (H, E) stands for the rules credibility, also known as the rules strength which reflects the level of reliability of conclusion *H* when precondition *E* is true, $-1 \leq CF(H, E) \leq 1$. The greater the value of *CF (H, E)* is, the higher reliability of the truth of conclusion *H* becomes when precondition *E* is true. On the contrary, the smaller the value of *CF (H, E)* is, the lower reliability of the truth of conclusion *H* becomes when precondition *E* is true.

2.3 Reasoning strategy

The reasoning strategy in the system is mixed reasoning (Chen Zhuangjian et al., 2007). In the process of reasoning, the system conducts the forward reasoning firstly, that is, reasoning the assumed conclusion from the initial symptoms. Lastly it conducts backward reasoning, that reasoning the try-out of the assumed conclusion to get the final identification summary (Lin Peiguang et al., 2008). The detailed reasoning process is shown in Fig.1.

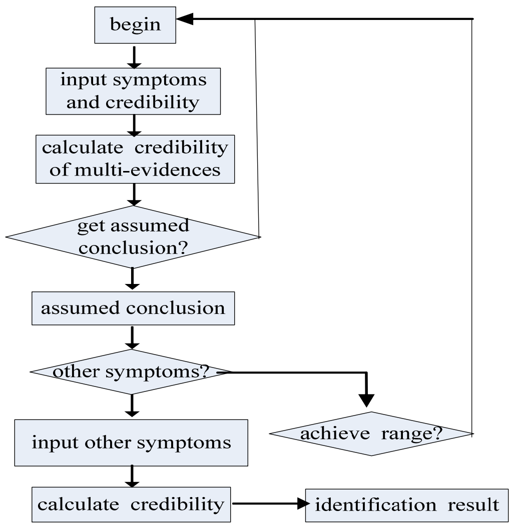

Fig.1: Reasoning process

3. UNCERTAINTY OF REASONING PROCESS

3.1 Uncertain representation of rules

In this system, a rule is known: $E \rightarrow H$. But how to define its credibility: *CF (H, E)*. First, define a prior probability *P (H)*. It expresses the believable degree of conclusion *H* when there is no evidence. Then, conditional probability *P (H / E)*. It expresses the believable degree of conclusion *H* when there is evidence *E* (Bai Yong et al., 2006).

(1)*P (H / E)*: it means that the credibility of the conclusion *H* is increased by the appearance of evidence *E*. We use *MB (H, E)* to express that the evidence *E* works on the growth of trust of conclusions *H*, then:

$$MB(H,E)=\{\frac{\max[P(H/E)^{1},P(H)]-P(H)}{1-P(H)} \quad (1)$$

(2)*P (H / E)*: it means that the credibility of the conclusion *H* is reduced by the appearance of evidence E. We use *MD (H, E)* to express that the evidence E works on the growth of distrust of conclusions *H*, then:

$$MD(H,E)=\{\begin{matrix}\min[P(H/E)^{1},P(H)]-P(H)\\ -P(H)\end{matrix} \quad (2)$$

The definition of certainty factor *CF (H, E)* can be got when based on the definition of *MB (H, E)* and *MD (H, E)*:

$$CF(H,E)=MB(H,E)-MD(H,E) \quad (3)$$

Thereinto, their value ranges are: $0 \le MB (H, E) \le 1$, $0 \le MD (H, E) \le 1$, $-1 \le CF (H, E) \le 1$. Certainty factor *CF (H, E)* is given by pineapple experts.

3.2 Uncertain representation of multi-evidences

The certainty factor of evidence is expressed with *CF (E)*, $-1 \le CF (E) \le 1$. *CF (E)* = 1 means evidence *E* is true with assurance; *CF (E)* =-1 means evidence *E* is false with assurance; *CF (E)* = 0 means evidence *E* is beyond our knowledge; $0 < CF (E) < 1$ means the extent of truth of *E*; $-1 < CF (E) < 0$ means the extent of falsehood extent of *E.*

(1)When the evidence is the conjunction of evidences, that is, *E = E1* and *E2* and… and E_n, and the corresponding certainty factor are: $CF (E_1)$, $CF (E_2)$… $CF (E_n)$, it is resulted: $CF (E) = min (CF (E_i))$, (i = 1, 2…n) (Shi Minghui et al., 2006);

(2)When the evidence is the extraction of evidences, that is, *E = E1* or *E2* or… or E_n, and the corresponding certainty factor are: $CF (E_1)$, $CF (E_2)$… $CF (E_n)$, it is resulted: $CF (E) = max (CF (E))$, (i = 1, 2… n);

(3)If the following two rules exist:

R_1: IF E_1 THEN E_2 $(CF (E_2, E_1))$

R_2: IF E_2 THEN H $(CF (H, E_2))$

That is, when the middle result deduced from R_1 acts as the evidence of R_2, certainty factor of E_2 $CF(E_2)$ is the certainty factor deduced from the use of R_1 (Shi Minghui et al., 2007).

3.3 Update of reasoning rules

(1)When the certainty factor of conclusions is calculated with the certainty factor of preconditions and rules, the formula for certainty factor of the conclusions *CF (H)* is as follows:

$$CF(H)=\max\{0,CF(E)\}\times CF(H,E) \quad (4)$$

(2)When two separate evidences and two different rules are used to deduce the certainty factor of a same conclusion, that is,

E1 → *H*, certainty factor of evidence and conclusion are respectively: *CF (E1)*, *CF (H, E1)*;

E2 → *H* certainty factor of evidence and conclusions are respectively: *CF (E2)*, *CF (H, E2)*;

CF1 (H), *CF2 (H)* are used to express the certainty factor of the two rules as:

$$CF_1(H) = \max(0, CF(E_1)) \times CF(H, E_1)$$
$$CF_2(H) = \max(0, CF(E_2)) \times CF(H, E_2) \qquad (5)$$

CF (H) is defined as follows:

$$CF(H) = \{ \begin{array}{l} CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H) \\ CF_1(H)+CF_2(H)+CF_1(H)\times CF_2(H) \\ CF_1(H)+CF_2(H)/(1-\min(CF_1(H),CF_2(H)) \end{array} \qquad (6)$$

4. EXPERIMENT

The above reasoning method has been widely applied to pineapple disease identification system. A specific disease, Heart Rot Disease (short as *H1*), a common disease of pineapple cultivation, is quoted as an example to illustrate its experimental reasoning process.

A rule is assumed as:

IF E_1 *and (* E_5 *or* E_4*) and* E_5 *and* E_6 THEN H_1 (0.96)

The certainty factor of the rules is given by experts as 0.96. Contrastingly, and the certainty factor of the evidences is given by users, including: $CF(E_1)$ = 0. 75, $CF(E_2)$ = 0. 8, $CF(E_4)$ = 0. 85, $CF(E_5)$ = 0. 9, $CF(E_6)$ = 0. 9. The meanings of evidences are shown in Table 1.

Table 1. Description of evidence

Evidence	Description
E_1	N,P,K in pineapple are lower than the standard
E_2	the pineapple abdominal shallowness
E_4	the pineapple abdominal height
E_5	the pineapple root slight rot
E_6	the pineapple dingy skin

According to the calculation method of certainty factor of conclusions discussed previously:

$$CF(H) = \max\{0, CF(E)\} \times CF(H, E) \qquad (7)$$

A formula can be got:

$$
\begin{aligned}
CF(H1) &= max\{0, CF(E1\ and\ (E2\ or\ E4)\ and\ E5\ and\ E6)\} \\
&= max\{0, min\{CF(E1), max\{CF(E2), CF(E4)\}, \\
&\quad CF(E5), CF(E6)\}\} \times 0.96 \\
&= max\{0, min\{0.75, max\{0.8, 0.85\}, 0.9, 0.9\}\} \times 0.96 \\
&= max\{0, 0.75\} \times 0.96 \\
&= 0.72
\end{aligned} \tag{8}
$$

That is: the credibility factor of Heart Rot Disease is 0.72.

5. CONCLUSION

The reasoning based on the evidence of the credibility fully considers the complexity of the objective world and the inaccuracy and incompleteness when human beings come to know the objective things. At the same time, it reduces the impact of subjective factors on the judgment and reasoning process. The reasoning method introduced by this paper has been used in the pineapple disease identification system. Experiments prove that when the value of CF (E) is set reasonably by users, reasoning result is closer to experts' practical experience and psychological instinct.

REFERENCES

Bai Yong, LanZhangli. Research of individual algorithm based on uncertainty inference, Computer Science, 2006, 33(3): 205-207

Chen Zhuangjian, Xie Wei. Implementation of import and export ontology and its reasoning frame, Computer Engineering, 2007, 33(21): 186-188(in Chinese)

Lin Peiguang. Indefinite knowledge representation based on ontology and cloud theory, Computer Engineering and Applications, 2008, 44(5): 51-54(in Chinese)

Shi Minghui, Zhou Changle, Wu Qingfeng. Uncertain inference based on certainty degree and its implementation by artificial neural network, Application Research of Computers, 2007, 24(1): 241-243(in Chinese)

Shi Minghui, Zhou Changle. Approach to uncertain inference for syndrome differentiation in TCM based on artificial neural network, Computer Engineering and Applications, 2006, 43(7):10-13(in Chinese)

Yang Jian. Research and application of CBR's progression, Computer Engineering and Design, 2003, 29(3): 710-712(in Chinese)

Zhang Wen, Zhang Guixu. Improved rule-based backward fuzzy reasoning algorithm, Journal of Communications, 2008, 29(2): 101-105(in Chinese)

Zhang Xuenong. A gradual approach for model-based diagnosis, Journal of Software, 2005, 19(3): 584-593(in Chinese)