A PHP-CLIPS Based Intelligent System for Diabetic Self-Diagnosis

Huiqing H. Yang and Sharnei Miller
Department of Math & Computer Science, Virginia State University
Petersburg, VA, 23806, United State

Abstract-*Diabetes is one of the major risk diseases for health care in our lives. In this paper, an online self-diagnosis for diabetes is developed, which can assist people in checking their risk for diabetes without consulting the specialists directly. In order to deal with some concepts expressed by verbal language in diagnosis, a fuzzy inference approach is adopted. The major endeavor of the research is to explore the potential of human intelligence and web-based technology in diabetic self-diagnosis. For the purpose of implementation, the paper integrates CLIPS tool with PHP programming technique for web-based Medicare applications. The online diagnosis system provides individual visitors with the ability to perform self-diagnosis under uncertainty. By means of GD graph feature in PHP, the system also provides graphic results for diagnosis.*

Keywords: Fuzzy Inference, Uncertainty, Medicare, Artificial Intelligent, PHP and CLIPS.

1 Introduction

Diabetes is a serious health problem today. Most of people are unaware that they are in risk of or may even have type 2 diabetes. Type 2 diabetes is the most common form of diabetes. There are about 90-95 percent of case studies. Diabetes is when the body does not produce enough insulin or the cells ignore the insulin. This work is aimed for people who do not like to take the time out to see a doctor/specialist regularly or even do not have time to see a doctor. If people were aware of the factors of diabetes and know how much of risks they are of getting diabetes, diabetes may be prevented early. In this paper, an online self-diagnosis system is developed, which is simple, easy to understand and user-friendly.

Due to the complexity of medical issues today, high technology is needed. The development of computer technology and tools has provided a great assistance for Medicare in a plethora of ways. Artificial Intelligence was primarily concerned in Medicine from the very earliest moments in the modern history of computer [1-2]. It is true that the medical field is a crucial and beneficial aspect of artificial intelligence.

The intention of our research is to provide an online self-diagnosis for people to see their risk for having diabetes. Since health diagnosis results expressed by verbal language often involve a mixture of uncertainties in the outcomes that are governed by the meaning of linguistic terms, inference under uncertainty is always a major issue. In our previous work [3], a fuzzy inference approach was developed, which will be used for diabetic diagnosis in this paper.

We also present an implementation based on the C Language Integrated Production System (CLIPS) and PHP web programming language. CLIPS is a productive development and delivery expert system tool which provides an environment for the construction of rule and/or object-based expert systems [4-5]. Because of its portability, extensibility, capabilities, low cost and powerful pattern matching called the Rete Algorithm, CLIPS has received widespread acceptance in AI research fields. Recently, many web applications have shown that PHP is flexible and fast enough to fulfill the requirements of designing an online information system. Thus, PHP is utilized as the web server-side scripting environment in this research.

The paper is organized as follows. Section 2 presents major risk factors for diabetic diagnosis and their measurements. Section 3 describes fuzziness consideration, and fuzzy inference is exampled in section 4. Section 5 shows details about implementation, and online demonstration results are provided in section 6. Our conclusion is given in the last section.

2 Major Risk Factors & Measurements

Since measuring heart rate and blood pressure are two important ways of assessing people health. The heart rate and blood pressure are mainly used as risk factors in this research. In addition, the Body Mass Index (BMI) is another major factor, which will be used to measure of person’s weight relative to the height.

Combining this information with additional risk factors such as cigarette smoking, individual can perform basic self-diagnosis regularly.

Measurement of heart rate

Normal resting heart rates range anywhere from 40 beats per minute up to 100 beats per minute. Ideally the normal heart rate is between 60-90 beats per minute. From a generally used formula, the maximum heart rate is about 220 minus person's age.

.

Measurement of blood pressure

The blood pressure is an essential and normal part of the way the body works. As blood is pumped around the body, it carries oxygen and nutrients that are essential for life. High blood pressure can enlarge and weaken the heart. It also damages the blood vessels. If the blood vessels become narrow or blocked, it may result in a heart attack or stroke.

General expression for the blood pressure is two numbers, such as 120/80 mmHg. The top number is the systolic blood pressure that is a measure of the pressure when the heart muscle is contracted and pumping blood. The bottom number is the diastolic blood pressure that is the pressure when the heart is relaxed and filling with blood.

Measurement of Body Mass Index (BMI)

Body mass index is measure of body fat based on height and weight that applies to both adult men and women. The BMI categories are listed as follows:

- BMI≤18.5 underweight
- 18.5≤BMI≤25 normal weight
- 25≤BMI≤30 overweight
- 30≤BMI Obesity

3 Fuzziness of Risk Factors

Since medical decision-making deals with some concepts expressed by verbal language, such as, fast heart rate, normal heart rate, and slow heart rate, fuzziness is frequently involved. In order to support fuzzy decision-making, the quantitative measures (Heart Rates and Blood pressure) are mapped to a range that corresponds to a term known as linguistic qualifiers. In this paper, we define certain qualifiers for different symptoms in order to perform a fuzzy inference. For example:

Heart Rate Qualifier:

HQ = {fast, normal, slow}.

Systolic Blood Pressure (SBP) Qualifier:

SBPQ={high, normal, low}.

Diastolic Blood Pressure (DBP) Qualifier:

DBPQ={high, normal, low}.

Each qualifier can be represented as a fuzzy subset of the linguistic words. Based on the fuzzy set theory, the membership function of qualifiers can be defined as a smooth transition. More formally a fuzzy set in a universe is characterized by a membership function $\mu: U \rightarrow [0,1]$.

Figure 1 illustrates the primary terms of fuzzy variable Heart Rate. Each term represents a specific fuzzy set.

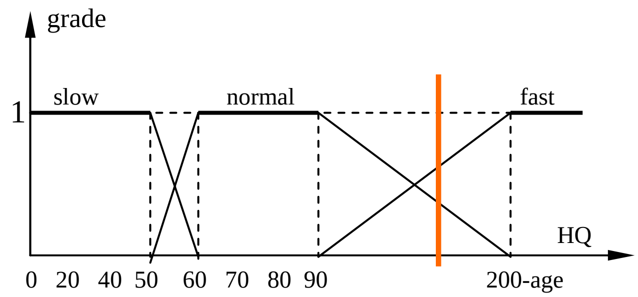

Figure 1. Membership function for Heart Qualifier

The fuzzy set functions for heart qualifiers can be described as:

$$\mu_{slow}(HQ) = \begin{cases} 1.0 & \text{if } 0 \le HQ \le 50 \\ (60 - HQ)/10 & \text{if } 50 \le HQ \le 60 \end{cases}$$

$$\mu_{normal}(HQ) = \begin{cases} (HQ - 50)/10 & \text{if } 50 \le HQ \le 60 \\ 1.0 & \text{if } 60 \le HQ \le 90 \\ (v - HQ)/(v-90) & \text{if } 90 \le HQ \le v \end{cases}$$

$$\mu_{fast}(HQ) = \begin{cases} (HQ-90)/(v-90) & \text{if } 90 \le HQ \le v \\ 1.0 & \text{if } v \le HQ \end{cases}$$

where $v = 200\text{-age}$.

In the same way, the fuzzy set functions for blood pressure can be described below:

$$\mu_{low}(SBPQ) = \begin{cases} 1.0 & \text{if } 0 \le SBPQ \le 100 \\ (110-SBPQ)/10 & \text{if } 100 \le SBPQ \le 110 \end{cases}$$

$$\mu_{normal}(SBPQ) = \begin{cases} (SBPQ - 100)/10 & \text{if } 100 \le SBPQ \le 110 \\ 1.0 & \text{if } 110 \le SBPQ \le 130 \\ (140-SBPQ)/10 & \text{if } 130 \le SBPQ \le 140 \end{cases}$$

$$\mu_{high}(SBPQ) = \begin{cases} (SBPQ - 130)/10 & \text{if } 130 \le SBPQ \le 140 \\ 1.0 & \text{if } 140 \le SBPQ \end{cases}$$

$$\mu_{low}(DBPQ) = \begin{cases} 1.0 & \text{if } 0 \le DBPQ \le 60 \\ (70\text{-}DBPQ)/10 & \text{if } 60 \le DBPQ \le 70 \end{cases}$$

$$\mu_{normal}(DBPQ) = \begin{cases} (DBPQ\text{-}60)/10 & \text{if } 60 \le DBPQ \le 70 \\ 1.0 & \text{if } 70 \le DBPQ \le 90 \\ (100\text{-}DBPQ)/10 & \text{if } 90 \le DBPQ \le 100 \end{cases}$$

$$\mu_{high}(DBPQ) = \begin{cases} (DBPQ - 90)/10 & \text{if } 90 \le DBPQ \le 100 \\ 1.0 & \text{if } 100 \le DBPQ \end{cases}$$

We define BMIQ as a fuzzy qualifier for BMI. The fuzzy set function for BMI can be written as:

$$\mu_{under}(BMIQ) = \begin{cases} 1.0 & \text{if } 0 \le BMIQ \le 18 \\ 19\text{-}BMIQ & \text{if } 18 \le BMIQ \le 19 \end{cases}$$

$$\mu_{normal}(BMIQ) = \begin{cases} BMIQ\text{-}18 & \text{if } 18 \le BMIQ \le 19 \\ 1.0 & \text{if } 19 \le BMIQ \le 24 \\ (26\text{-}BMIQ)/2 & \text{if } 24 \le BMIQ \le 26 \end{cases}$$

$$\mu_{over}(BMIQ) = \begin{cases} (BMI - 24)/2 & \text{if } 24 \le BMIQ \le 26 \\ 1.0 & \text{if } 26 \le BMIQ \end{cases}$$

Since a common feature of the fuzzy set is overlapping, that is, the qualifier may be associated with two different terms at the intersect intervals. For instance, the Heart Rate qualifier HQ may take '*normal*' and '*fast*' at the same time. How do we make the decision under this situation? This reveals uncertainty – the lack of adequate and correct information to make a decision.

4 A Fuzzy Inference

We propose an inferencing approach for health diagnosis, in which we construct a certainty factor (**CF**) to evaluate the degree of certainty [3]. The key idea relevant to the determination of the certainty factors is described as follows.

Case 1. Individual fact with a qualifier

In this case, there is only one qualifier associated with a fuzzy fact.

- If the qualifier only takes one term at given interval, the grade of membership $\mu(\ )$ can be used as a CF.

- If the given fact is in the overlapping area, two fuzzy terms will be related. For example, given age=60 and HQ=120, the fuzzy inference may give the following results:

Heart Rate is ***normal*** *with* $CF_h = \mu_{normal}(HQ) = 0.4$

Heart Rate is ***fast*** *with* $CF_d = \mu_{fast}(HQ) = 0.6$

CF_h is used for common health diagnoses. CF_d can be used to diagnose a disease.

In this paper, we mainly concern about health diagnoses. Thus, the final diagnosis results might look like these:

Heart Rate is ***normal*** *with* $CF = CF_h = 0.4$.

Case 2. Individual fact with multi-qualifiers

This is a case in which there are multiple qualifiers associated with a fact. Many pieces of fuzzy terms are conjoined (AND) or disjoined (OR). The blood pressure is a typical example for this case.

Normal blood pressure depends on both systolic and diastolic blood pressures. For example,

Systolic blood pressure is ***normal***
with $CF_{sb} = \mu_{normal}(SBPQ=132) = 0.8$

Diastolic blood pressure is ***normal***
with $CF_{db} = \mu_{normal}(DBPQ=97) = 0.3$

Hence, to perform these kinds of diagnosis, we have to handle multiple fuzzy qualifiers. According to the fuzzy set theory, the conjunction can be defined as the minimum of the involved qualifiers. Therefore, the CF can be determined by the following formula:

$$CF=min\{ CF_{sb}, CF_{db} \}.$$

Case 3. Multiple Facts

In general there are many facts that affect our health. A general formula can be written as:

Health ={`f1, f2, … …, fi`}.

The certainty can be calculated as follows:

$$CF = w_{f1} \times CF_{f1} + w_{f2} \times CF_{f2} + \dots \dots + w_{fi} \times CF_{fi}.$$

where

fi is a fact, i represents the number of the facts;
w_{fi} is a weight associated with the fact fi;
CF_{fi} is the certainty factor associated with fi.

and a set of weights satisfies:

$$\sum w_{fi} = 1.$$

5 Implementation

In this paper, an intelligent system has been developed. The structure of the system is shown in the Figure 2.

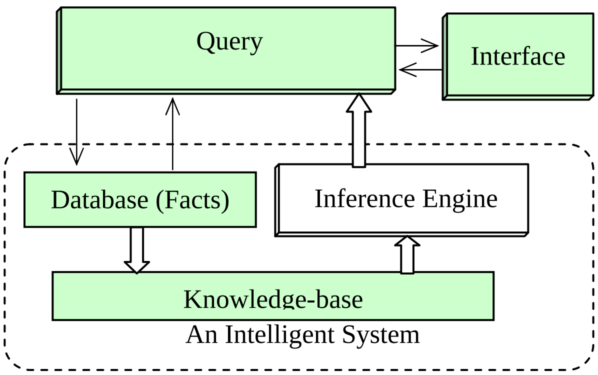

Figure 2. System architecture

The system is implemented by means of HTML, PHP and CLIPS. CLIPS is a tool for building expert systems. PHP is a server-side web programming language. The components are described as follows:

Interface: The user can access the system through a friendly web interface, which has been implemented by using HTML and PHP scripts.

Query Module: The query processing uses PHP scripts to pass the queries to an intelligent system, which performs inference based on the user queries.

An Intelligent System: Internally, the system consists of three main components. The database includes all the facts on which inferences are derived. The knowledge-base contains rules with which the inference engine draws conclusions. These conclusions are the system's responses to the user's queries. The inference engine makes reasoning. The system is encoded using CLIPS and PHP

5.1 Store Facts in CLIPS

The facts are the critical resources querying in suggestion self-test in this research. Before facts can be created, CLIPS are required to declare the structure which can define the list of valid slots for a given relation name. This can be done by using the **deftemplate** construct. Samples of facts used in this paper are listed in Table 1.

Table 1. Define templates

```
; Define templates
(deftemplate Survey “A Self-test”
        (slot weight)
        (slot pressure)
        (slot smoke)
        . . . . . .
)
```

The type of information stored in the database includes all of 'Yes' or 'No' answers. In the online system, facts representing information are dynamically added to the fact list according to the user's queries.

5.2 Knowledge Representing

In order to accomplish work, an intelligent system must have facts as well as rules. Based on the system model, we define the rules to support the processing of user's queries, and perform intelligent diagnosis. In CLIPS environment, rules can be defined by using the **defrule** construct. Table 2 shows a sample of rules. The forward chain inference is applied here.

Table 2. Rules in CLIPS

```
(defrule Yes-Weight-Rule
    (Survey (weight Yes))
 =>
   (assert Response (action “You should
           try to change your diet and
           exercise to loose weight.”)) )
)

(defrule No-Weight-Rule
    (Survey (weight No))
 =>
   (assert Response (action “You are not
                     overweight.”)) )
)
```

Rules used to calculate a Certainty Factor (CF) in fuzzy inference are implemented as functions by PHP scripts. Sample code is given in Table 3.

Table 3. Fuzzy Inference Rules

```
<?php
  //A function used to calculate
  //Certainty factor for normal heart rate
  function normalHR($hr, $age)
  {
     $a=200-$age;
     if($hr<60) return 0;
     if($hr>=60 && $hr<90) return 1;
     if($hr>=90 && $hr<=$a)
         return (($a-$hr)/($a-90));
  }

  //A function used to calculate
  //Certainty factor for fast heart rate
  function fastHR($hr, $age)
  {
     $a=200-$age;
     if($hr<90) return 0;
     if($a<$hr) return 1;
     if($hr>=90&&$hr<=$a)
         return (($hr-90)/($a-90));
   }
?>
```

5.3 Query and Inference

All of the CLIPS code must be processed through the CLIPS interpreter. Table 4 presents samples of PHP source code that accesses CLIPS commands or functions such as clear, reset, load and run, and retrieves the results stored in the fact list.

Table 4. PHLIPS Source Code for Query

```
<?php
  //get user’s answers
  $weight=$_REQUEST['weight'];
  $history=$_REQUEST['history'];
  $pressure=$_REQUEST['pressure'];
  $age=$_REQUEST['age'];
   . . . . . . .

  //Load all of clips files
  clips_clear();
  clips_load("practice.clp");
  clips_load("Survey.clp");
  clips_reset();

  //Fire the rule and get the results
  $firedRule=clips_run();
  $fact=clips_get_fact_list();

  //Display the results
  . . . . . .

?>
```

Table 5 shows PHP source code for fuzzy inference.

Table 5. PHP Code for Fuzzy Inference

```
<?php
   . . . . . .
  //Calculate CF for each risk factor
  $BMI=calcBMI($height, $weight);
  $overBMI=overBMI($BMI);
  $normalBMI=normalBMI($BMI);

  if($overBMI>$normalBMI)
  {  $bmiCF=$overBMI;
     $bmiInfo='Over Weight';
  }
  else
  {     $bmiCF=$normalBMI;
        $bmiInfo='Normal weight';
  }
   . . . . . . .

  //Calculate the final CF
  $CF=0.1*$hbCF + 0.3*$dbfCF + 0.3*$sbpCF +
      0.2*$bmiCF + 0.1*$smokingCF;
  . . . . . .
?>
```

6 Online System Demonstration

The system has been evaluated by sample data. The main page is shown in Figure 3. The system provides two types of diagnosis. Version 1 is advice/suggestion self-test, which is based on Yes/No questions. Sample data is listed in Figure 4 and 5. Figure 6-9 provide an example for Version 2, which is a fuzzy-based diagnosis.

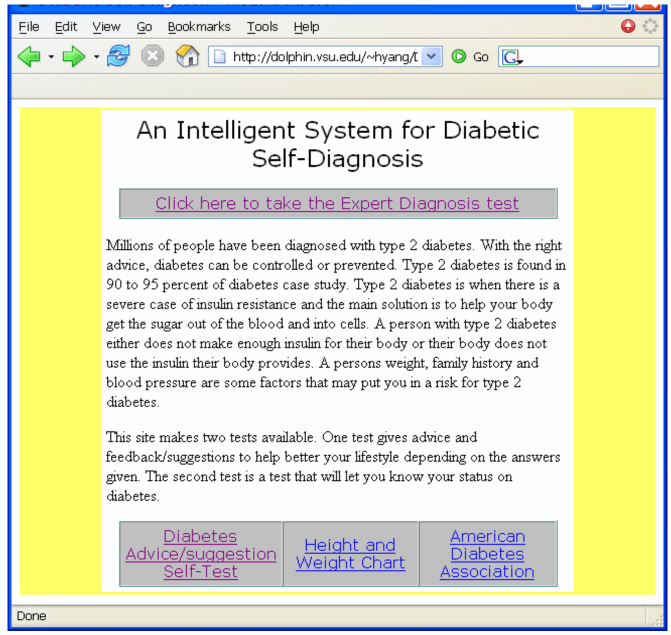

Figure 3. Main Page of Diagnosis System

Version 1: Advice/Suggestion Self-Test

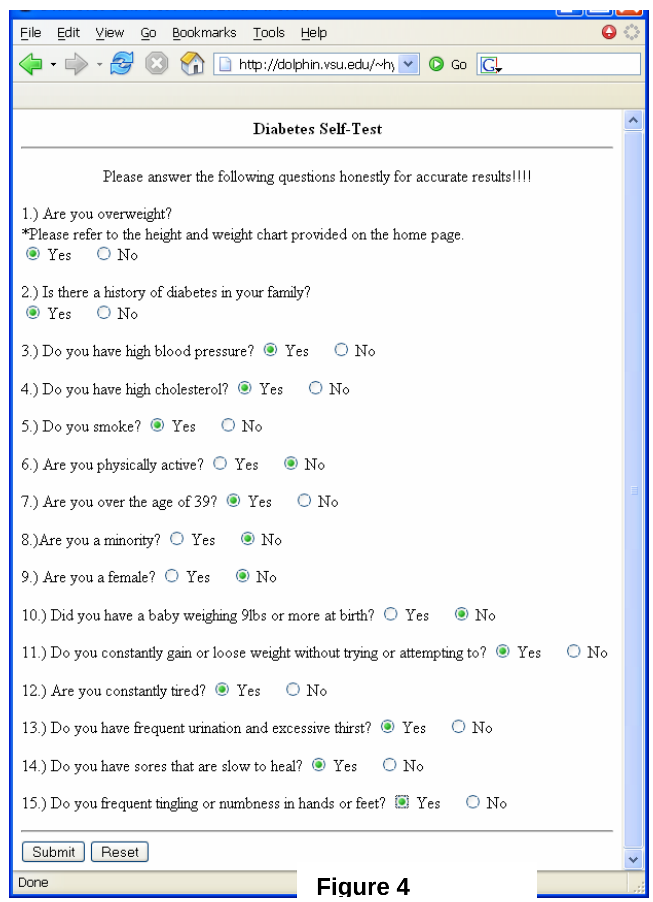

Figure 4

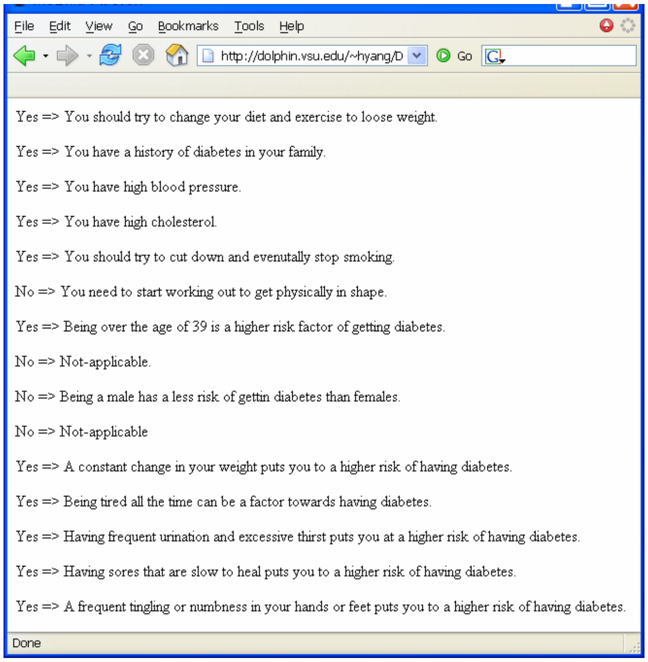

Figure 5. A Self-test Result

Version 2: Fuzzy-based Diagnosis

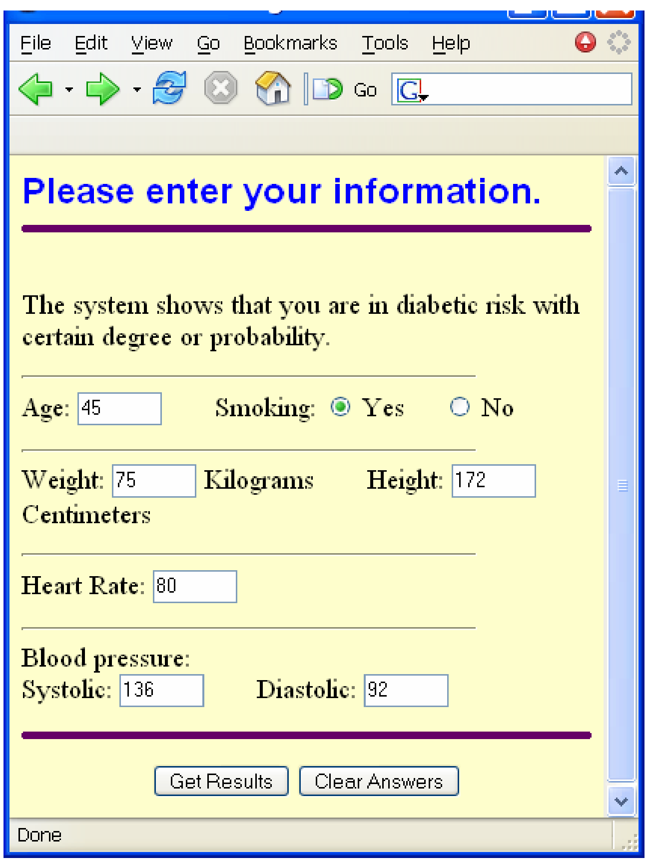

Figure 6. Fuzzy-based Diagnosis

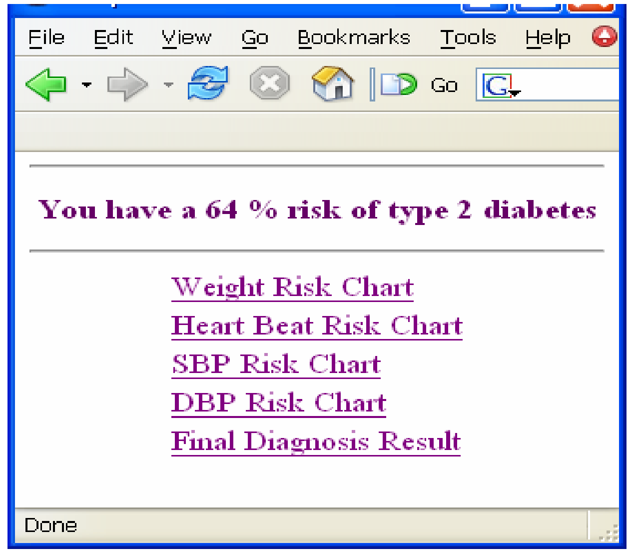

Figure 7. A Result for Fuzzy-based Diagnosis

The diagnosis results in version 2 can be represented by using graphics. Figure 8 shows the certainty factor (CF) for SBP risk factor. The final diagnosis CF is displayed in Figure 9.

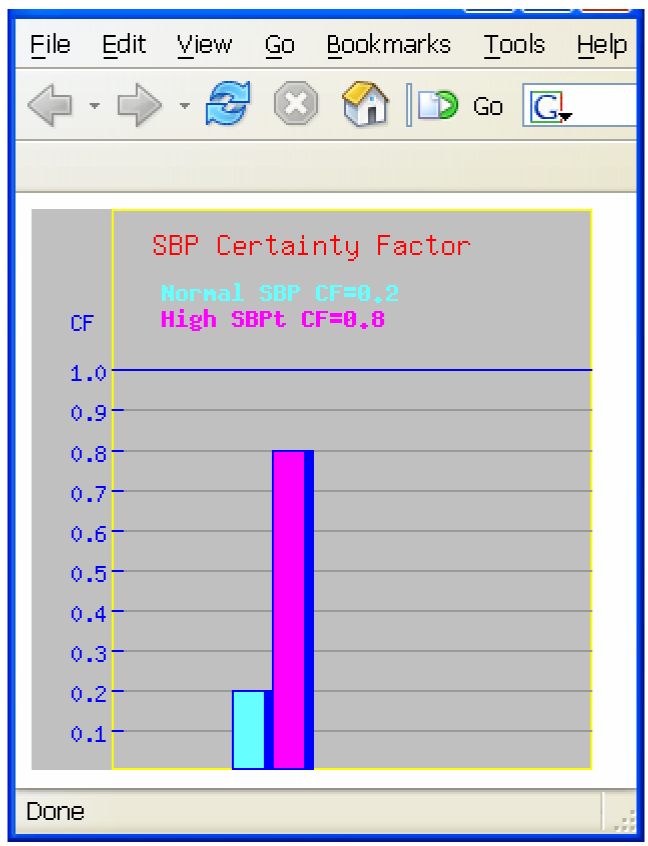

Figure 8. Certainty Factor (CF) for SBP

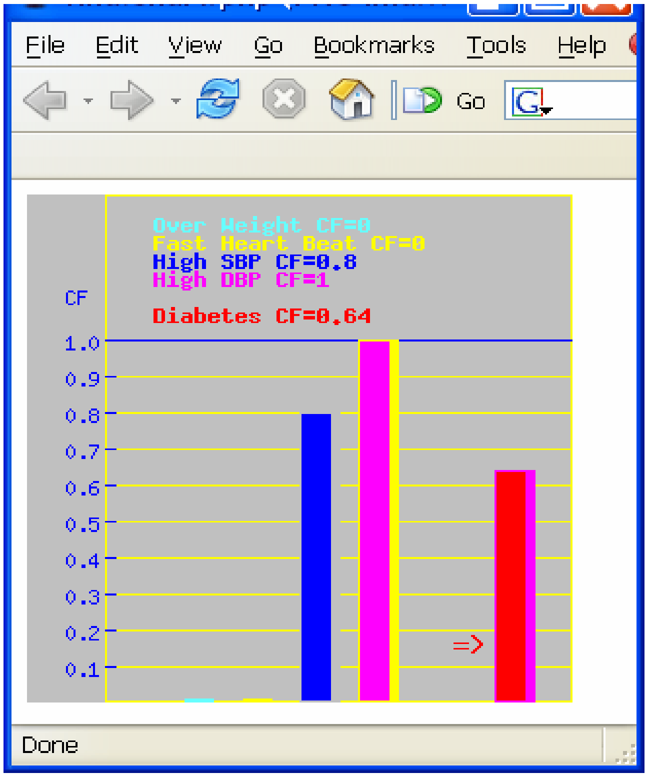

Figure 9. Certainty Factor (CF) for Diabetes

7 Conclusion

The modern health diagnosis requires sophisticated theory and computer technologies to plan diagnostic procedures, interpret the results, and conduct research. Based on the general knowledge, and the current recommendation of the United State Centers for Disease Control and Prevention, an online diabetic self-diagnosis system is built with PHP and CLIPS in this paper. The system provides an instant online analysis based on important medical symptoms, and gives suggestions that you are "at risk" for developing a diabetes disease with certain degree. The purpose of this research is to help individuals to reduce or control risk of diabetic disease and improve health condition.

8 References

[1] Enrico Coiera, Guide to Health Informatics, the Internet and Telemedicine, 1st Edition, Chapman & Hall, 1997.

[2] Edward H. Shortliffe, Leslie E. Perreault, et al, Medical Informatics: Computer Applications in Health Care and Biomedicine, Springer-Verlag New York, Inc. 2001.

[3] H.Yang, "A Fuzzy Inference Approach for Health Self-Diagnosis Under Uncertainty", The 9th World Multi-Conference on Systematic Cybernetics and Informatics, IIIS and IFSR, Orlando, Florida, 2005.

[4] CLIPS Reference Manual, Version 6.10, 1998.

[5] R. M. Wygant, "CLIPS – A Powerful Development and Delivery Expert System Tool", Computers in Engineering, Vaol. 17, No.14, 1989, pp546-549.

[6] Larry Bielawski and Robert Lewand, Intelligent Systems Design: Integrating Expert Systems, Hypermedia, and Database Technologies, John Willey and Sons, 1991.

[7] Joseph C. Giarratano and Gary D. Riley, "Expert Systems: Principle and Programming", 4th Edition, Thompson Learning Inc., 2005.

[8] Stuart Russell and Peter Norvig, "Artificial Intelligence and A Modern Approach", 2nd Edition, Prentice Hall, 2003.

[9] Leon Atkinston, PHP Programming, Prentice Hall, 2004.