CLIPS: A Tool for Corn Disease Diagnostic System and An Aid to Neural Network for Automated Knowledge Acquisition

Cathy Wu, Pam Taylor, George Whitson and Cathy Smith

Department of Mathematics and Computer Science
The University of Texas at Tyler
Tyler, Texas 75701-6699

ABSTRACT

This paper describes the building of a corn disease diagnostic expert system using CLIPS, and the development of a neural expert system using the fact representation method of CLIPS for automated knowledge acquisition. The CLIPS corn expert system diagnoses 21 diseases from 52 symptoms and signs with certainty factors. CLIPS has several unique features. It allows the facts in rules to be broken down to <object-attribute-value> (OAV) triples, allows rule-grouping, and fires rules based on pattern-matching. These features combined with the chained inference engine result to a natural user query system and speedy execution.

In order to develop a method for automated knowledge acquisition, an Artificial Neural Expert System (ANES) is developed by a direct mapping from the CLIPS system. The ANES corn expert system uses the same OAV triples in the CLIPS system for its facts. The LHS and RHS facts of the CLIPS rules are mapped into the input and output layers of the ANES, respectively; and the inference engine of the rules is imbedded in the hidden layer. The fact representation by OAV triples gives a natural grouping of the rules. These features allow the ANES system to automate rule-generation, and make it efficient to execute and easy to expand for a large and complex domain.

INTRODUCTION

Many criteria can be used to evaluate an expert system: the accuracy and efficiency, the ease of use, the ease of initial building and later expansion, and extra features such as the explanation facility and certainty representation. Diagnostic rule-based expert systems are among the most important and successful expert systems. However, the implicit nature of the domain knowledge has made it difficult to develop new expert systems on different domains even with the available rule-based expert system shell, because it requires explicit rules. Artificial neural system [4] has been used as an alternative approach to build diagnostic expert systems to overcome the knowledge acquisition bottleneck [1,3,5,7]. On the other hand, the neural systems lack a built-in explanation facility and a natural query system. Furthermore, the representation of the domain knowledge in a large single set of values makes the neural expert systems not suitable for a large and complex domain.

This paper descibes and compares the two corn disease diagnostic systems, one rule-based using CLIPS [2] and one neural network using ANES [6]. The paper also shows the automated knowledge acquisition scheme used in the ANES corn system with a direct mapping to CLIPS system. The fact representation method in both systems allows the rule-grouping and result to speedy execution, natural query system, and easy system expansion.

CLIPS CORN EXPERT SYSTEM

CLIPS, a rule-based expert system tool developed at NASA, is used to build the corn disease diagnostic system that identifies 21 diseases from 52 symptoms and signs. The facts are broken down to <object-attribute-value> (OAV) triples. Each object in the OAV triples has two components: <plant_part> and <pathogen_type>. There are five plant_parts, namely, seedling, whole_plant, leaf, stalk_or_root, and tassel_or_ear; and three pathogen_type, fungus, bacterium, and virus. The attribute is the <descriptor>, which can be a symptom or sign or a disease. The facts for the 52 symptoms and signs are grouped into ten fact lists (ie., ten deffacts), five for symptoms on five plant_parts each and five for signs on five plant_parts each. The fact template for symptom or sign has the form of: (<plant> <pathogen>

<symptom/sign> <value>). Figure 1A shows the fact lists for signs on seedlings.

The fact template for disease, however, needs two additional fields, a Certainty Factor (CF) field, and a tag field. The certainty factor ranges from 0 to 1, to indicate the degree of confidence for the firing of certain disease(s) (RHS of the rule) from the observed symptom or sign (LHS of the rule). To tag each fact uniquely, a unique tag is generated for each disease fact (OAV triple) using the gensym function [2]. Thus, the final fact template for disease has the form of: (<plant> <pathogen> <disease> <value> <CF> <tag>). There are 52 IF-THEN rules (ie., defrules) that associate each one of the 52 symptoms or signs to its related disease(s) (Figure 1B). The same OAV triples that are derived by separate rules are combined to produce a single OAV triple with a combined certainty factor (Figure 2A).

CLIPS fires rules based on a pattern-matching mechanism. The fact representation method combined with the pattern-matching mechanism creates a natural rule-grouping. The priority of the firing of each rule group can be further controlled by the use of salience. In the corn expert system there are 15 rule groups, each corresponding to an object (ie., a plant_part and pathogen_type combination). The rule-grouping mechanism and the chained inference engine result to a speedy execution. Furthermore, the rule-grouping provides a natural user interface to query only a subset of symptoms or signs in order to reach a conclusion (Figure 2B). This makes it unnecessary to emulate the backward chaining inference engine commonly used for goal satisfaction. However, the emulation of backward chaining in CLIPS is fairly straightforward. For example, one can simply add a fact list that relates all diseases with their associated symptoms and signs for the back tracking and let the rule fire in the normal forward fashion.

ARTIFICIAL NEURAL CORN EXPERT SYSTEM

ANES is an artificial neural expert system tool developed at the University of Texas at Tyler that uses back-propagation network [6]. With ANES, it is possible to build a diagnostic expert system by mapping the symptoms/signs directly to diseases without knowing the exact contribution (with certainty factor) of individual symptom/sign to a particular disease. The former is the implicit knowledge of a domain expert. The latter is an explicit if-then rule derived from the implicit knowledge by the domain expert through a time-consuming process.

The facts are represented by the same OAV triples that are used in the CLIPS system. Each input fact of a rule, a triple, is converted to an input vector of 32 neurons (Figure 3) by a preprocessor, while each output fact of a rule is obtained from the output vector by a postprocessor. Thus, the input (LHS) and output (RHS) facts are mapped into the input and output layers of the ANES, respectively; and the inference engine for the rules are imbedded in the hidden layer (Figure 4). The fact representation by OAV triples gives a natural grouping of the rules. There are two rule groups in the corn ANES: rule group 1 to connect symptoms and signs to diseases, and rule group 2 to determine a disease from all possible candidates (Figure 5). Rule group 1 consists of 4 subgroups, each of which corresponds to symptoms, fungal signs, bacterial signs, and viral signs, respectively. The subgroup in turn consists of five rules, each associates symptoms or signs on a particular plan_part to certain diseases.

Because of the rule-grouping mechanism of the ANES, the system can be implemented onto a parallel architecture to break down one large neural networks to many small parallel networks [6]. This would speed up execution and make the expansion of the knowledge base much easier. The direct mapping of the CLIPS corn expert system to ANES using the same rule-grouping mechanism allows the development of an automated knowledge acquisition scheme. The ANES inference engine is capable of extracting the implicit knowledge embedded in the neural network [6].

CONCLUSION

Both CLIPS and ANES expert system tools produced corn diagnostic systems that diagnose accurately and are easy to use. The representation of facts using OAV triples in both systems allows the grouping of rules, which speeds up the execution, provides a natural way to break down a complex system to subsystems, and allows a chained inference and natural query system. Building an expert system using ANES is easier, however, because of the automated knowledge acquisition.

SYMBOLS AND ABBREVIATIONS

ANES (Artificial Neural Expert System), OAV (Object-Attribute-Value), RHS (Right-Hand Side), LHS (Left-Hand Side).

```
A.      (deffacts seedling-sign
                (seedling fungus sign mycelia-and-spores)
                (seedling bacterium sign bacterial-droplets))
```

```
B.      (defrule myccelia-and-spores
                ?has <- (it has seedling fungus sign myccelia-and-spores)
        =>
                (assert (seedling fungus disease seedling-blight .7 =(gensym)))
                (assert (seedling fungus disease root-rot .7 =(gensym)))
                (retract ?has))
```

Figure 1. Fact (A) and rule (B) representation in CLIPS

```
A.      (defrule combine-CF
                ?fact1 <- (?plant ?pathogen disease ?name ?CF1 ?)
                ?fact2 <- (?plant ?pathogen disease ?name ?CF2 ?)
                (test (neq ?fact1 ?fact2))
        =>
                (retract ?fact1 ?fact2)
                (bind ?CF3 (+ ?CF1 ?CF2))
                (assert (?plant ?pathogen disease ?name ?CF3 =(gensym))))
```

```
B.      (defrule goal
                (?plant ?pathogen disease ?name ?CF ?)
                ?sym-sign <- (?plant ?pathogen ? ?value)
        =>
                (if (> ?CF .8)
                then
                (fprintout t "The disease may be a(n)" ?name crlf)
                (fprintout t "Certainty Factor is  " ?CF crlf)
                else
                (fprintout t "Has it " ?value "on " ?plant crlf)))
```

Figure 2. Rules for (A) combining certainty factors, and (B) goal satisfaction in CLIPS.

32 Input/output neurons for each OAV triple

8 Objects		3 Attributes	21 Values
5 Plant_Parts +	3 Pathogen_Types	3 Descriptors	21 Diseases
Seedling	Fungus	Symptom	or
Whole_plant	Bacterium	Sign	Symptoms/Signs
Leaf	Virus	Disease	on Plant_Part
Stalk_or_root			
Tassel_or_ear			

Figure 3. Mapping of facts (OAV triples) to input and output vectors in ANES.

Input Layer (LHS Facts)		Hidden Layer (Rules)		Output Layer (RHS Facts)
Fact in-1	->	Rules	->	Fact out-1
8 Objects				8 Objects
3 Descriptors				3 Descriptors
21 Values				21 Values
Fact in-2				Fact out-2
.				.
.				.
Fact in-M				Fact out-N
Input Layer		Hidden Layer		Output Layer
(LHS Facts)		(Rules)		(RHS Facts)

Figure 4. Fact and rule representation in ANES.

Rule Group 1

Rule Subgroup 1

	r1		
31 Symptoms	r2		21
on 5 Parts ->	r3	->	Diseases
	r4		
	r5		

Rule Subgroup 2

	r6		
Fungal Signs	r7		Fungal
on 5 Parts ->	r8	->	Diseases
	r9		
	r10		

Rule Subgroup 3

Rule Group 2

Diseases		r21		
Identified		r22		
From	->	r23	->	Disease
Group 1		r24		
		r25		

Figure 5. Rule groupings in ANES

REFERENCES

1. Gallant, S.I. 1988. Connectionist expert systems. Communications of the ACM 31(2): 152-169.
2. Giarratano, J. and G. Riley. 1989. Expert Systems: Principles and Programming. PWS-KENT Publishing, Boston.
3. Kulkarni, Arun D., George Whitson, Jim Bolin and Cathy Wu. 1989. Some applications of the parallel distributed processing models. Proceedings of Workshop on Applied Computing '89: 185-192.
4. Rumelhart, D.E. and J.L. McClelland. (eds). 1986. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1: Foundations. A Bradford Book, MIT Press.
5. Saito, k and R. Nakano. 1988. Medical diagnostic expert system based on PDP model. Proceedings of the IEEE Second Annual International Conference on Neural Networks I: 255-262.
6. Whitson, George, Cathy Wu and Pam Taylor. 1990. Using a neural system to determine the knowledge base of an expert system. Proceedings of the 1990 ACM Sigsmall/PC Symposium on Small Computers: 268-270.
7. Yoon, Y.O., R.W. Brobst, P.R. Bergstresser and L.L. Peterson. 1989. A desktop neural network for dermatology diagnosis. Journal of Neural Network Computing, Summer: 43-52.