A Symbolic Computation-Based Expert System for Alzheimer’s Disease Diagnosis*

Begoña Herrero¹, Luis M. Laita¹, Eugenio Roanes-Lozano², Víctor Maojo¹, Luis de Ledesma¹, José Crespo¹, and Laura Laita³

¹ Depto. de Inteligencia Artificial, Facultad de Informática, Universidad Politécnica de Madrid, Boadilla del Monte, 28660-Madrid, Spain
² Depto. de Álgebra, Universidad Complutense de Madrid, c/ Rector Royo Villanova s/n, 28040-Madrid, Spain
³ Escuela de Enfermería, Universidad Complutense de Madrid
{laita,eroanes}@fi.upm.es

Abstract. In this paper we summarize a method of construction of a rule-based expert system (denoted RBES) for Alzheimer’s disease diagnosis. Once the RBES is constructed, Symbolic Computation techniques are applied to automatically both verify (that is, check for consistency) and extract new knowledge to produce a diagnosis.

Keywords: Rule-based expert systems, diagnosis of Alzheimer’s disease, ideal membership problem, Gröbner bases

1 Introduction

In this paper we summarize a method of construction of a rule-based expert system (denoted as RBES) for Alzheimer’s disease diagnosis. The full RBES is described in [6]. After building the RBES, symbolic computation techniques are applied to both automatically verify (that is, check for consistency) and extract new knowledge to output a diagnosis.

RBES usually have two components: a “knowledge base” and an “inference engine”. In this article, the knowledge base and the inference engine are, respectively, symbolically expressed and symbolically implemented using the computer algebra language CoCoA [4]. We use the 3.0b MSDOS version of CoCoA because, even though the Windows version [14] is more user friendly, it appears to be less efficient.

* Partially supported by projects TIC2000-1368-C03-01 and TIC2000-1368-C03-03 (Ministry of Science and Technology, Spain).

This Master thesis (in Spanish) contains a large amount of information about the illness, an explanation of the 343 rules entered, a glossary and a diskette with the entire program.

Our particular knowledge base consists of production rules, ADDI (additional information given by the experts) and facts. Classic bi-valued logic will be used in this system.

The “production rules” are, in this article, logical implications of the following forms:

\[ \circ x[i] \land \circ x[j] \land \ldots \land \circ x[k] \rightarrow \circ y[l] \]

or

\[ \circ x[i] \lor \circ x[j] \lor \ldots \lor \circ x[k] \rightarrow \circ y[l] \]

(or, possibly, a combination of \( \lor \) and \( \land \) in the antecedent), where “\( \circ \)” represents the symbol “\( \neg \)” or no symbol at all.

For example the first rule of our system, \( R1 \), refers to memory disorder. It is:


The most relevant “facts” are what are known as “potential facts”, which are the literals (that is, variables or variables preceded by “\( \neg \)”). These are on the left-hand side of the production rules and never on the right-hand side.

For example the first fact of our system, \( x[1] \), is “difficulty in remembering what day it is” (the list of the potential facts can be found in section 4.2).

Potential facts allow “fire” the rules. A rule is fired when, given the literals of its left-hand side, its right-hand side is obtained by the logical rule of modus ponens (from \( \alpha \) and \( \alpha \rightarrow \beta \), then \( \beta \)).

It is important to note that, in this paper, we consider the set of both potential facts and their negations. But, in each case study (that is, each class of patients w.r.t. the equivalence relation “his/her relatives have answered the questions about the symptoms the same way and the results of the neurological evaluation and laboratory tests are equal”), we consider one maximal consistent subset of such a set of facts. “Consistent”, because a fact and its negation are never included in any case study; “maximal”, because either each fact or its negation must be included. Note that, depending on which maximal consistent set of facts is used, different rules are fired, giving different results.

Regarding the additional auxiliary information, denoted “ADDI”, these are formulae that the experts judge should or must be taken into account.

Let us say that CoCoA requires all logical formulae of the RBES to be written in prefix form. For instance, the prefix form of the first formula above is:

\[ \text{IMP(AND1(AND1(AND1\ldots(AND1(\circ x[i], \circ x[j], \ldots, \circ x[k])\ldots), \circ y[1])))}. \]

AND1 (also OR1 in other formulae) are used instead of AND and OR, because these last two words are reserved words in CoCoA.

Our inference engine proceeds as follows: first, logical formulae are automatically translated into polynomials and, second, “Gröbner bases” and “normal forms” are applied to these polynomials (using CoCoA). A substantial part of the background theory is the original work of a group of researchers to which the authors belong [25][9][15].