ARTIFICIAL AND HUMAN INTELLIGENCE (A. Elithorn and R. Banerji, editors)
Elsevier Science Publishers, B.V.

Chapter 24

EXPERT EVALUATION IN LOGIC OF ENVIRONMENTAL RESOURCES THROUGH NATURAL LANGUAGE

LUIZ MONIZ PEREIRA, EUGENIO OLIVEIRA and PAUL SABATIER

Dept. de Informatica, Universidade Nova de Lisboa,
Quinta da Torre, 2825 Monte de Caparica, Portugal

1. Introduction

Sophisticated artificial intelligence techniques have been increasingly used in knowledge engineering to build useful understandable and highly performant systems called Expert Systems.

There are several formalisms to represent knowledge such as production rules, semantic networks and frames, among others. But for the sake of transparency declarativity and easiness we choose to represent and compute all kinds of knowledge in an unique formalism. This homogeneous formalism is a subset of predicate first order logic named Horn Clause Logic, on which the programming language PROLOG (Warren, Pereira and Pereira, 1977) is based. We use it to represent domain knowledge (inference rules and facts), meta-knowledge (knowledge about that knowledge), natural language syntax and semantics.

Knowledge is expressed not only explicitly (facts encoded by unit clauses) but also by means of axioms which are manipulated deductively in order to answer complex queries by deriving new facts from old ones.

We call this kind of system a Logic Data Base, a query is regarded as a theorem to be proved from the assertions (see Kowalski, 1981, for comparison with work in the Data Base field, namely Relational Data Bases).

2. ORBI

ORBI is an expert system for environmental biophysical resource evaluation, commissioned by the Portuguese Government and developed with the collaboration of the Department of Environment. What makes ORBI of an enormous utility is that it has not only the knowledge of one expert (or a set of them) in a certain restricted domain, but expertise on several disciplines like geology, hydrology, fauna, flora, microclimate and so on, to obtain results

about the interdisciplinary domain of biophysical resource evaluation. ORBI makes judgements about the aptitudes of certain regions for industry, agriculture, recreation, and others. Related, observed basic data are digitized into quantified descriptors (e.g. risk of soil erosion). By means of inference rules established with the help of experts, we obtain quantified values for aptitudes which reveal the environmental assets available relative to human needs (e.g. intensive agriculture) The descriptors' values at representative points on a grid over a map are codified as facts in unit clauses. The values range from 0 to 5 and to each is attached a representativeness (from 1 to 5), which measures the extent to which the region represented by the point is homogeneous for that value. ORBI can communicate with the user in two modes: using natural language interface and letting him have the initiative, or by taking the initiative and conducting the user through a menu of fixed questions. The system can be switched from one mode to the other at any time. We want our system to have in the near future more meta-knowledge as well as the ability of acquiring new knowledge directly from the expert, the whole being represented in first order predicate logic.

3. Natural language consultation

In the natural language mode of consultation each question goes through linguistic processors providing for different analysis (morphological, syntactic and semantic). As a result of natural language query the output of the linguistic parsing is a logical sentence (a list of ordered goals) directly executable by Prolog, as in CHAT80 (Warren and Pereira, 1981) and MICROSIAL (Pique and Sabatier, 1981) systems.

The syntactic and semantic analysis is realized by means of grammar containing rules and controls. Rules handle the fundamental structures of Portuguese, in particular:

- yes-no and "why" questions
- affirmative, negative, relative prepositional, coordinate, extraposed and elliptic clauses;
- complex noun complementation and adjunction;
- universal, existential, numeral, definite and indefinite determiners;
- anaphora and pronouns; and nouns, verbs and adjectives specific to the application and to the linguistic terminology.

Controls verify number and gender agreements, designation of complex entities, and noun and verb complementations. unknown words, incorrect queries, grammatical faults, and erroneous designations and complementation are pointed out to the user, inviting him to correct them.

At any time, the user may ask ORBI on how he can formulate his question,

calling its structures, vocabulary and examples of queries. The linguistic knowledge necessary for such abilities – which dispense the user from consulting a (complex and difficult) reference manual – is not entirely grouped in a specific database, but also expressed and computable from the different natural language processors' clauses (morphology, dictionary, grammar). Future developments will consider allowing ORBI's grammar to be consulted as a lingustic database for making the accepted structures implicit and for giving examples of queries. Its lingustic abilities might be extended to point out parts of a sentence ORBI cannot understand, to force analysis of incorrect strings, and to allow the user to define new structures and words.

4. Implementation

The PROLOG interpreter, itself a consequent driven production system, has fixed depth-first, right-to-left strategy with backtracking on failure. Nevertheless, we wrote a small interpreter in PROLOG itself embedding a problem-dependent strategy to execute deterministic computations first, and to look at the code to generate explanations. The whole system is implemented on a small 64K DEC LSI 11/23 with associated video terminal, a digitizer and graphic terminal. The programming effort has been of about two man-years.

5. Remark

This paper describes the state of our system in 1981. Subsequent improvements are reported in Pareira, Oliviera and Sabatier (1982), and Pareira and Oliviera (1983).