Chapter 1

The Consistency of the Medical Expert System CADIAG–2:
A Probabilistic Approach

Pavel Klinov
The University of Manchester, UK

Bijan Parsia
The University of Manchester, UK

David Picado Muiño
Institut für Diskrete Mathematik und Geometrie, Austria

ABSTRACT

CADIAG-2 is a well known rule-based medical expert system aimed at providing support in medical diagnose in the field of internal medicine. Its knowledge base consists of a large collection of IF-THEN rules that represent uncertain relationships between distinct medical entities. Given this uncertainty and the size of the system, it has been challenging to validate its consistency. Recent attempts to partially formalize CADIAG-2’s knowledge base into decidable Gödel logics have shown that, on formalization, the system is inconsistent. In this paper, the authors use an alternative, more expressive formalization of CADIAG-2’s knowledge base as a set of probabilistic conditional statements and apply their probabilistic logic solver (Pronto) to confirm its inconsistency and compute its conflicting sets of rules under a slightly relaxed interpretation. Once this is achieved, the authors define a measure to evaluate inconsistency and discuss suitable repair strategies for CADIAG-2 and similar systems.

INTRODUCTION

CADIAG-2 (Computer Assisted DIAGnosis) is a well known rule-based expert system aimed at providing support in diagnostic decision making in the field of internal medicine. Its design and construction was initiated in the early 80’s at the Medical University of Vienna by K.P. Adlassnig (Adlassnig et al., 1985, 1986; Adlassnig, 1986; Leitich et al., 2002) for more on the origins and design of CADIAG-2.
CADIAG-2 consists of two fundamental pieces: the inference engine and the knowledge base. The inference engine (for alternative formalizations and analyses of CADIAG-2’s inference see Ciabattoni et al., 2010; Picado Muiño, 2010) is based on methods of approximate reasoning in fuzzy set theory, in the sense of (Zadeh, 1965, 1975). In fact CADIAG-2 is presented in some monographs as an example of a fuzzy expert system (see for example Klir et al., 1988; Zimmermann, 1991).

The knowledge base consists of a set of IF-THEN rules –also known in the literature as production rules– intended to represent relationships between distinct medical entities: symptoms, findings, signs and test results (S) on the one hand and diseases and therapies (D) on the other. The vast majority of them are binary (i.e., they relate single medical entities) and only such rules are considered in this paper. The one that follows is an example of a binary rule of CADIAG-2 (Adlassnig et al., 1986):

IF suspicion of liver metastases by liver palpation
THEN pancreatic cancer
with degree of confirmation 0.3

The degree of confirmation refers, intuitively, to the degree to which the conditioning event (i.e., ‘suspicion of liver metastases by liver palpation’ in the example above) confirms the uncertain event (i.e., ‘pancreatic cancer’ above). How these degrees of confirmation are to be formally interpreted will be discussed later.

In this paper we present a formalization of a coded version of the binary fragment of CADIAG-2’s knowledge base (i.e., that contains only codes for the identification of the distinct medical entities) as a probabilistic logic theory. We then check the satisfiability of that formalization with Pronto, our probabilistic description logic solver, which we briefly introduce. We find that CADIAG-2 is highly unsatisfiable (confirming the results of an alternative, weaker formalization, Ciabattoni et al., 2010) and analyze the sources of unsatisfiability.

To our knowledge, the probabilistic version of CADIAG-2 is the largest PSAT (Probabilistic SATisfiability) problem to be solved by an automated reasoner and is certainly the largest non-artificial one. This is, perhaps, a bit misleading as it is comparatively easy to detect unsatisfiability by first heuristically detecting small but likely unsatisfiable fragments and then performing a satisfiability check on each fragment. While this might suffice to validate that CADIAG-2 is unsatisfiable it is not sufficient, without further qualification, to detect all conflicting sets of rules, nor can it ensure that a satisfiable fragment is so in the context of the entire knowledge base.

As CADIAG-2 is too large (the number of rules in the binary fragment we are concerned with is over 18000) we describe an approach to split the knowledge base into comparatively large fragments that can be tested independently and prove that such methodology is complete, i.e., is guaranteed to find all conflict sets. With this methodology we are able to determine that CADIAG-2 contains numerous sets of conflicting rules and compute all of them for a slightly relaxed interpretation of the knowledge base.

We complete the paper with the introduction of an inconsistency measure aimed at evaluating CADIAG-2-like databases and a brief account of suitable repair strategies for CADIAG-2 and similar systems. The measure presented attempts to quantify how far the knowledge base is from consistency and its computation, in as much as it yields an adjustment in the degree of confirmation or uncertainty of each conditional statement, provides the modeler with a possible repair of the database.