Towards an interpretation of the medical expert system CADIAG2

David Picado Muiño
European Centre for Soft Computing
Edificio Científico Tecnológico
Gonzalo Gutiérrez Quirós s/n. 33600 Mieres, Spain
david.picado@softcomputing.es

Thomas Vetterlein
Johannes Kepler University Linz - JKU
Department of Knowledge-Based Mathematical Systems
Altenbergerstrasse 69. A-4040 Linz, Austria
thomas.vetterlein@jku.at

Agata Ciabattoni
Technische Universität Wien
Institut für Computersprachen
Favoritenstrasse 9-11. A-1040 Wien, Austria
agata@logic.at

August 22, 2011

Abstract

The present paper responds to an attempt to *interpret* the inference process and output of the medical expert system CADIAG2. In order to do so, we first provide a logical formalization of the inference process by means of a set of rules aimed at describing the steps along the inference and later attempt to provide a sound semantics for them. Two semantics are taken as reference for our purpose: probabilistic semantics and fuzzy (*t-norm*-based) semantics.

1 Introduction

CADIAG2 (Computer Assisted DIAGnosis) is a well known rule-based expert system that aims 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 University of Vienna Medical School (now Medical University of Vienna)

by Klaus-Peter Adlassnig –see [20] for a historical perspective on the origins and motivations of the system–.

CADIAG2 consists basically of two pieces: a knowledge base and an inference engine. CADIAG2's knowledge base is formed by a set of *IF-THEN* rules, also known as *production* rules in the literature, intended to represent relationships between distinct medical entities: symptoms, findings, signs and test results on the one hand (to which we will commonly refer as *symptoms*) and diseases and therapies on the other (to which we will commonly refer as *diagnoses*). The rules in CADIAG2 are defined along with a certain *degree of confirmation* which is intended to express the *degree* to which the antecedent *confirms* the consequent. For example,

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

The inference engine in CADIAG2 takes as input (possibly) *imprecise* medical information about the patient, normally in the form of a set of symptoms present *to some degree* in the patient, and yields as output a set of possible diagnoses, each along with a value intended to represent some degree of *certainty* about its presence in the patient. The inference rules in the knowledge base of the system are brought into play along the inference process, which is based on methodology from *fuzzy set theory*, in the sense of [23] and [24].

The main aim of the present paper responds to an attempt to *interpret* the inference process in CADIAG2, and ultimately the output of the system, on the grounds of a *sound* semantics. Two semantics will be taken as reference in our attempt: probabilistic semantics and fuzzy (*t-norm*-based) semantics –see for example [10] or [12] for more on *t-norms* and some other concepts mentioned below that are related to them–. The use of probabilistic semantics is motivated by the *natural* identification of the degrees of confirmation in the rules of the system with probabilities (in principle with frequencies, as suggested in [3], *estimated* from medical databases and patient records, although not all degrees of confirmation were obtained in this way) and the rules themselves with probabilistic conditional statements. The use of a fuzzy semantics is mostly motivated by the *natural* identification of the *degrees of presence* of symptoms in the patient with *membership degrees* in fuzzy set theory (i.e., as *truth degrees*). It is common practice in the field to choose a *t-norm* as the interpretation of the conjunction and its *residuum* as the interpretation of the implication with which we will characterize the rules of the system: rules in this context will be formalized as *graded* implications in the context of many-valued logics, in a sense that will be made clear later.

As is probably expected, the outcome of such an attempt can be at least partially anticipated. The inference mechanism in CADIAG2 is built on methodology from fuzzy set theory and thus it is bound to be unsound with respect to probabilistic semantics. Some aspects of the probabilistic unsoundness of the inference mechanism in CADIAG2 were soon observed in earlier studies concerning

the celebrated rule-based expert system MYCIN (that shares some background methodology with CADIAG2) –see [4] or [21] for a description of MYCIN, [8] for a comparison of CADIAG2 and MYCIN-like systems and [11], [13], [14], [22] for probabilistic approaches to it–. However, it remains to be seen *how far* the system CADIAG2 is from probabilistic soundness and thus *how much* of its inference process could be interpreted probabilistically. A better match of the inference process may be expected with respect to some *t-norm*-based fuzzy semantics (in particular, as will be clear later, with the semantics based on the indentification of the *t-norm* with the *minimum* operator) yet, as with probabilistic semantics, it needs to be seen *how much* of the inference process can be interpreted on the grounds of this semantics.

It is worth mentioning here that, although the interest among theoretical AI researchers in rule-based expert systems seems to be lesser today than some years ago, rule-based expert systems are very popular among AI engineers. Many CADIAG2-like systems are in use and more are being built for future implementation. It is mainly for this reason that we believe that further analysis and understanding of CADIAG2-like systems is of relevance (CADIAG2 is presented in some monographs as an example of a *fuzzy expert system* –for example in [15] or [25]– and thus is used as a reference for some newly developed knowledge-based systems).

The paper is structured as follows: in Section 2 we introduce some notation and give some preliminary definitions necessary for the description of the inference mechanism of the medical system CADIAG2, which is done in Section 3, where we also describe the knowledge base of the system. In Section 4 we introduce the logical system **CadL** that consists of a collection of rules that formalize the steps made along the inference process. Such a formalization will facilitate the semantic analysis of the system carried out in Section 5. Section 6 summarizes results.

2 Preliminary definitions

In this section we define a pair of concepts that we will need to describe the inference process in CADIAG2.

First we define a partial ordering relation.

Definition 1 *Let $\preceq$ be the partial ordering relation on $[0,1]$ defined as follows: for $a, b \in [0,1]$, $a \preceq b$ if and only if $0 < a \leq b$ or $0 \leq a < 1$ and $b = 0$.*

We define the strict partial ordering $\prec$ from $\preceq$ in the conventional way.

As we will see later, the definition of the ordering $\preceq$ responds to the use of both 0 and 1 as maximal values in CADIAG2 for the interval $[0,1]$. The value 0 denotes certainty in the non-occurrence of an event or falsity of a statement and the value 1 denotes certainty in its occurrence or its truth.

For the next definition let

$$\mathbb{D} = [0,1] \times [0,1] - \{(0,1),(1,0)\}.$$

Definition 2 *The function* $\max^* : \mathbb{D} \longrightarrow \mathbb{R}$ *is defined as follows, for* $(a,b) \in \mathbb{D}$*:*

$$\max^*(a,b) = \begin{cases} a & \text{if } b \prec a \\ b & \text{otherwise} \end{cases}$$

In words, $\max^*(a,b)$ returns the biggest value among a and b with respect to the ordering $\preceq$ just defined.

3 The medical expert system CADIAG2

In this section we briefly introduce the medical expert system CADIAG2 –for more details on its design one can look at [1], [2] or [3]–.

As already mentioned in the introduction, CADIAG2 consists of two fundamental pieces: the knowledge base and the inference engine. We first describe the different types of rules in the system, mostly in relation to the role they play along the inference process, and later describe the essentials of the inference mechanism. Before getting started some notation is needed.

Let $p_1, \ldots, p_l$ denote the *basic* medical entities that occur in CADIAG2 (i.e., the symptoms and diseases in the system), for some $l \in \mathbb{N}$. CADIAG2 deals also with *compound* entities, build from basic ones by means of *conjunction* ($\wedge$), *disjunction* ($\vee$) and *negation* ($\sim$) (i.e., built as Boolean combinations of basic ones).

Strictly speaking, the system regards two additional types of connectives called *at least* n *out of* m and *at most* n *out of* m, with $n, m \in \mathbb{N}$ and $n \leq m$. However, these can be expressed in terms of conjunction and disjunction and thus are not taken into account in this paper. As an example,

$$\textit{at least } 2 \textit{ of } (\phi_1, \phi_2, \phi_3)$$

can be rewritten as

$$(\phi_1 \wedge \phi_2) \vee (\phi_1 \wedge \phi_3) \vee (\phi_2 \wedge \phi_3),$$

for ϕ_1, ϕ_2, ϕ_3 some arbitrary medical entities in CADIAG2.

The knowledge base. The knowledge base of CADIAG2, to which we will commonly refer as KB, consists of a collection of approximately 40.000 *IF-THEN* rules that express possibly uncertain relationships among distinct medical entities.

Rules in CADIAG2 can be characterized by triples of the form $\langle \theta, \phi, \eta \rangle$, where θ is the *antecedent*, ϕ the *consequent* and η is the degree to which θ confirms ϕ (i.e., the degree of confirmation), for θ, ϕ medical entities and $\eta \in [0,1]$.

In some literature about CADIAG2 –see for example [1] or [7]– rules are defined as 4-tuples of the form $\langle \theta, \phi, \eta, \zeta \rangle$, for θ, ϕ medical entities and $\eta, \zeta \in [0, 1]$, where η stands for the degree to which θ (the antecedent) confirms ϕ (the consequent) and ζ for the degree to which ϕ confirms θ, sometimes referred to in the corresponding literature as the *strength of confirmation* and the *frequency of occurrence* of the rule respectively. The 4-tuple $\langle \theta, \phi, \eta, \zeta \rangle$ corresponds in our notation to the pair of triples $\langle \theta, \phi, \eta \rangle$ and $\langle \phi, \theta, \zeta \rangle$.

We can distinguish among three different types of rules by considering their form and *how* they are used along the inference process in relation to the values assigned to their antecedent or consequent.

- Type ***confirming to the degree*** η (c_η). A rule of this type is of the general form $\langle \theta, \phi, \eta \rangle$, for θ, ϕ medical entities (ϕ a basic medical entity) and $\eta \in (0, 1]$. It is triggered in a run of the inference mechanism in CADIAG2 for strictly positive values or grades of its antecedent (in a way that will be made clear below, where we describe the inference engine of CADIAG2). We will generally refer to rules of this kind as rules of type **c**.

A rule of type **c** formalizes (possibly) uncertain interrelations among medical entities, the bigger the degree of confirmation η the more certain the presence of the consequent given the antecedent of the rule.

What follows is an example of a rule of type **c**, taken from [1]:[1]

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

- Type ***mutually exclusive*** (**me**). A rule of type **me** is of the form $\langle \theta, \phi, 0 \rangle$, for θ, ϕ medical entities (ϕ a basic medical entity). It is only triggered in a run of the inference engine of the system when there is certainty about the truth or occurrence of θ.

A rule of type **me** expresses mutual exclusiveness between antecedent and consequent (i.e., the presence of one of them excludes the other).

The one that follows is an example of a rule of this type:

> IF *positive rheumatoid factor*
> THEN NOT *seronegative rheumatoid arthritis*

- Type ***always occurring*** (**ao**). A rule of type **ao** is of the form $\langle \theta, \phi, 1 \rangle$, for θ, ϕ medical entities. It can be triggered by the system only when there is certainty about the falsity or non-occurrence of ϕ.

[1] The subsequent examples in this subsection are also taken from [1].

A rule of type **ao** expresses the fact that the antecedent implies the consequent. It follows that, if the consequent is excluded, the presence of the antecedent is also excluded.

Notice that a rule $\langle \theta, \phi, 1 \rangle$ of type **ao** can be alternatively formalized by the triple $\langle \sim \phi, \sim \theta, 1 \rangle$ and that it is not a special case of a rule of type **c** due to the fact that $\sim \theta$ is not a basic medical entity.

Next we give an example of a rule of this type:

> IF NOT (*rheumatoid arthritis* AND *splenomegaly* AND *leukopenia* $\leq 4000/\mu l$)
> THEN NOT *Felty's syndrome*

There are other typologies for the rules in KB that will prove useful in further sections in this paper. A very general typology is the one that follows:

- ***Binary* rules**. Rules of the form $\langle \theta, \phi, \eta \rangle$ where θ, ϕ are basic entities.
- ***Compound* rules**. Rules of the form $\langle \theta, \phi, \eta \rangle$ where θ is a compound medical entity and ϕ is a basic entity.

The vast majority of rules in KR are binary. There are less than one hundred compound rules in KR yet, despite the number, they are important for the functioning of the system.

We have a further distinction among binary rules of use in further sections:

- ***Symptom-symptom* rules**. Rules of the form $\langle \theta, \phi, \eta \rangle$ where both θ, ϕ are symptoms and $\eta \in \{0, 1\}$.
- ***Disease-disease* rules**. Rules of the form $\langle \theta, \phi, \eta \rangle$ where both θ, ϕ are diseases and $\eta \in \{0, 1\}$.
- ***Symptom-disease* rules**. Rules of the form $\langle \theta, \phi, \eta \rangle$ where θ is a symptom, ϕ a disease and $\eta \in [0, 1]$.
- ***Disease-symptom* rules**. Rules of the form $\langle \theta, \phi, \eta \rangle$ where θ is a disease, ϕ is a symptom and $\eta \in [0, 1]$.

Most rules of type *disease-symptom* in KB are not used by the inference engine, only those of type **ao** are used by it.

The inference engine. CADIAG2 gets started with medical information about the patient. Such information is formally given by a set of basic medical entities present in the patient, each one together with a number in the interval $[0, 1]$ which, in principle, is intended to represent the degree to which such entity is present (i.e., its degree of presence). These values are, in most of the literature on CADIAG2, interpreted as membership degrees in the context of fuzzy set theory and respond to the possibly *vague* nature of medical entities in CADIAG2.

Values assigned to compound medical entities in the system (in principle only for those that are relevant for the inference) are generated according to the following rules, for θ, ϕ any medical entities:

- The assignment to $\theta \wedge \phi$ is obtained as the *minimum* between the corresponding assignments to θ and ϕ.
- The assignment to $\theta \vee \phi$ is obtained as the *maximum* between the corresponding assignments to θ and ϕ.
- The assignment to $\sim \theta$ is obtained as the difference between 1 and the assignment to θ.

After the initial medical information about the patient is obtained and entered into the system the rules in the knowledge base come into play. All the rules triggered by the initial information about the patient are used during the inference process. At each step in the inference process a rule of type **c**, **me** or **ao** is applied (that is done, in principle, in no particular order). Rules of these types are triggered as follows, for general entities θ, ϕ:

- A rule $\langle \theta, \phi, \eta \rangle$ of type **c** can be triggered at some step in the inference process if a strictly positive value has been previously assigned to θ. The use of the rule $\langle \theta, \phi, \eta \rangle$ will generate a new assignment for ϕ, calculated as the minimum between the value assigned to θ that triggers the rule and η.
- A rule $\langle \theta, \phi, 0 \rangle$ of type **me** can be triggered during the inference process if certainty about the presence of θ in the patient (i.e., the assignment 1) has been previously concluded. The application of $\langle \theta, \phi, 0 \rangle$ allows us to conclude certainty about the absence of ϕ in the patient (i.e., the assignment 0).
- A rule $\langle \theta, \phi, 1 \rangle$ of type **ao** can be triggered if certainty about the absence of ϕ has been previously concluded. The application of the rule $\langle \theta, \phi, 1 \rangle$ will allow us to conclude certainty about the absence of θ in the patient.

The inference process goes on until the system comes to the stage where neither new medical entities nor new assignments for those already generated can be inferred. CADIAG2 yields as outcome of the inference the set of diseases generated during the inference process along with the *maximal value* (with respect to the ordering $\preceq$ defined above) assigned to them during the inference.

It has to be mentioned that, according to part of the literature on CADIAG2 –for example [1]–, the original inference process in CADIAG2 works in a slightly different way. The update in the value of the distinct sentences involved in the inference is done as soon as two different values for the same sentence are produced by the system. The value chosen in the update for atomic sentences in L is the *maximal* one (with respect to the ordering $\preceq$). Notice though that this feature has a highly undesirable result (unless further restrictions on the

rules or on the order in which the rules are used are imposed), which is that the outcome of a run of the inference mechanism can depend on the order in which the rules are applied. Such a drawback is easily avoided by assuming (as we do for this paper) that the chosen value, the maximal among all those produced along the inference with respect to the partial ordering $\preceq$, is only computed at the end of the process.

Notice that the system can generate what is called a *runtime inconsistency* given the ordering $\preceq$, produced when both values 0 and 1 are assigned to a medical entity along the inference process. In such case the system stops and produces an error message.

4 A formalization of the inference process

In this section we provide a logical formalization of the inference process in CADIAG2 by means of a *complete* set of rules aimed at describing the possible steps along the inference.

Let L be a finite propositional language and SL the set of sentences obtained from L as its closure under *conjunction* ($\wedge$), *disjunction* ($\vee$) and *negation* ($\sim$). In the context of CADIAG2, the set $\{p_1, \ldots, p_l\}$ of basic medical entities will be a subset of L and the compound medical entities that can be obtained from L will be a subset of SL.

Let $\Gamma = \{\phi_1, ..., \phi_n\} \subset SL$, for some $n \in \mathbb{N}$. We will denote the sentence $\phi_1 \wedge ... \wedge \phi_n$ by $\bigwedge \Gamma$.

Definition 3 *A graded statement in L is a pair of the form (ϕ, η), with $\phi \in SL$ and $\eta \in [0, 1]$.*

In the context of CADIAG2 a graded statement of the form (ϕ, η) represents the medical entity ϕ together with the value assigned to it, η, either at the outset (i.e., if ϕ is part of the initial information with which CADIAG2 gets started) or during the inference process.

4.1 The calculus CadL

In this subsection we summarize results in [6] and present, in a slightly simplified version, the calculus **CadL** aimed at formalizing the inference process in CADIAG2.

First we define the notion of theory of **CadL**:

Definition 4 *A theory $\mathcal{T}$ of **CadL** is a pair of the form (Φ, R) characterized as follows:*

- *Φ is a finite set of graded statements in L.*
- *$R = R^c \cup R^{me} \cup R^{ao}$, with R^c, R^{me} and R^{ao} finite collections of rules of type **c**, **me** and **ao** respectively.*

In the context of CADIAG2 Φ would be given by the input of the system (i.e., the initial information about the patient) and R would be given by KR.

Let $\mathcal{T} = (\Phi, R)$ be a theory of **CadL**. We have the following rules:

- **Reflexivity rule**

$$(REF) \quad \frac{(\phi, \eta) \in \Phi}{\mathcal{T} \vdash (\phi, \eta)}$$

- **Evaluation rules**

$$(AND) \quad \frac{\mathcal{T} \vdash (\phi, \eta) \quad \mathcal{T} \vdash (\theta, \zeta)}{\mathcal{T} \vdash (\phi \wedge \theta, \min(\eta, \zeta))}$$

$$(OR) \quad \frac{\mathcal{T} \vdash (\phi, \eta) \quad \mathcal{T} \vdash (\theta, \zeta)}{\mathcal{T} \vdash (\phi \vee \theta, \max(\eta, \zeta))}$$

$$(NOT) \quad \frac{\mathcal{T} \vdash (\phi, \eta)}{\mathcal{T} \vdash (\sim \phi, 1 - \eta)}$$

- **Manipulation rules**

$$(C) \quad \frac{\langle \theta, \phi, \eta \rangle \in R^c \quad \mathcal{T} \vdash (\theta, \zeta)}{\mathcal{T} \vdash (\phi, \min(\eta, \zeta))} \quad \text{for } \zeta > 0$$

$$(ME) \quad \frac{\langle \theta, \phi, 0 \rangle \in R^{me} \quad \mathcal{T} \vdash (\theta, 1)}{\mathcal{T} \vdash (\phi, 0)}$$

$$(AO) \quad \frac{\langle \theta, \phi, 1 \rangle \in R^{ao} \quad \mathcal{T} \vdash (\phi, 0)}{\mathcal{T} \vdash (\theta, 0)}$$

REF simply aims at formalizing for general theories of the form (Φ, R) that a graded statement that belongs to Φ (i.e., to the input in CADIAG2) is itself inferred as a consequence (i.e., as part of the output in CADIAG2). The evaluation rules AND, OR and NOT aim at formalizing the assignments to compound medical entities along the inference process in a run of the inference mechanism of CADIAG2 and the rules C, ME and AO correspond to the use of rules of type **c**, **me** and **ao** respectively during the inference process, as explained in the previous section.

Given a theory $\mathcal{T}$ of **CadL** and a graded statement (ϕ, η), a *proof* of (ϕ, η) from $\mathcal{T}$ in **CadL** is defined as a finite sequence of *sequents* of the form

$$\mathcal{T} \vdash (\phi_1, \eta_1), ..., \mathcal{T} \vdash (\phi_n, \eta_n)$$

with $(\phi_n, \eta_n) = (\phi, \eta)$ and where, for $i \in \{1, ..., n\}$, each (ϕ_i, η_i) in $\mathcal{T} \vdash (\phi_i, \eta_i)$ follows from $\mathcal{T}$ by the application of one of the rules above, from graded statements in previous sequents.

We say that there exists a *maximal* proof of (ϕ, η) from $\mathcal{T}$ in **CadL** if there exists a proof of (ϕ, η) from $\mathcal{T}$ and there is no proof from $\mathcal{T}$ of (ϕ, ζ) with $\eta \prec \zeta$.

Let us now consider the theory $\mathcal{T} = (\Phi, R)$, with $R = KR$, ϕ a disease in L and $\eta \in [0, 1]$. As should be clear from the description of the inference mechanism of CADIAG2 given in Section 3, the medical entity ϕ along with the value η would be given as an outcome in a run of the inference process of CADIAG2 on input Φ only if there exists a maximal proof of (ϕ, η) from $\mathcal{T}$ in **CadL** –for more details on this point see [6]–. In **CadL** a runtime inconsistency generated by the system would imply the existence of maximal proofs of $(\theta, 0)$ and $(\theta, 1)$ from $\mathcal{T}$, for some medical entity θ.

5 Towards a semantics for CadL

In this section we look at the interpretation of the inference process in CADIAG2. We consider two possible alternatives in our attempt: probabilistic semantics and fuzzy semantics.

5.1 Probabilistic semantics

The motivation for a probabilistic interpretation of the inference in CADIAG2 comes from the identification of the degrees of confirmation of rules in KR with *frequencies* or, more generally, probabilities and the rules in KR themselves with probabilistic conditional statements.

In this subsection we will assume that rules of the form $\langle \theta, \phi, \eta \rangle \in KR$ represent probabilistic conditional statements, where θ is the conditioning event or evidence, ϕ the uncertain event and η the probability of ϕ given that θ is true or that it occurs.

In order to set the inference process on probabilistic grounds and analyze its adequacy with probability theory we need also a *suitable* probabilistic interpretation of the graded propositions taken as input and generated along the process by the system. Recall that the value η in a statement of the form (ϕ, η) in the input of CADIAG2 is intended to represent the degree of presence of ϕ in the patient, normally identified with a membership degree in the context of fuzzy set theory (i.e., with a *degree of truth*). Here though we will adopt a probabilistic interpretation for these values.

We will focus our analysis on the binary fragment of KR (i.e., on the binary rules in KR), which we will denote by KR^{bin}. The vast majority of rules in KR are, as mentioned earlier, binary and they constitute the most characteristic fragment of CADIAG2 when seen as a representative example of a certain type of expert system. This restriction means leaving the *evaluation* rules in **CadL** aside. We will focus our analysis of the inference engine and thus of **CadL** on the *manipulation* rules.

Before going any further we need to introduce some preliminary notation and definitions.

Definition 5 *Let $\omega : SL \longrightarrow [0,1]$. We say that ω is a probability function on L if the following two conditions hold, for all $\theta, \phi \in SL$:*

- *If $\models \theta$ then $\omega(\theta) = 1$.*
- *If $\models \sim (\theta \wedge \phi)$ then $\omega(\theta \vee \phi) = \omega(\theta) + \omega(\phi)$.*[2]

The first clause of the definition simply states that if θ is *always* true (or if it always occurs) then its probability must be 1 whereas the second one states that if ϕ and θ are *never* true at once (or that they never occur together) then the probability of $\theta \vee \phi$ is equal to the sum of the probabilities of θ and ϕ.

From Definition 5 the standard properties of probability functions on propositional languages follow. We give some without proof –for a proof and more details on probability functions see for example [17]–. For ω a probability function on L and $\theta, \phi \in SL$,

- $\omega(\theta \vee \phi) = \omega(\theta) + \omega(\phi) - \omega(\theta \wedge \phi)$,
- $\omega(\sim \theta) = 1 - \omega(\theta)$,
- If $\theta \models \phi$ then $\omega(\theta) \leq \omega(\phi)$.

For the next definition let us consider $\langle \theta, \phi, \eta \rangle$ to be a conditional probabilistic statement, for $\theta, \phi \in SL$ and $\eta \in [0,1]$.

Definition 6 *We say that a probability function ω on L satisfies $\langle \theta, \phi, \eta \rangle$ if*

$$\frac{\omega(\theta \wedge \phi)}{\omega(\theta)} = \eta.$$

If there exists such a probability function we then say that $\langle \theta, \phi, \eta \rangle$ is *satisfiable*.

As seen in the previous section, the inference mechanism in CADIAG2 gets started with a set of graded statements of the form (q, η), with $q \in L$ a basic medical entity present in the patient. Let us consider as an example the medical entity '*reduced glucose in serum*'. Let us assume that the value assigned at the outset in a run of the inference engine by the evaluation system in CADIAG2 to the statement '*Patient A has reduced glucose in serum*' out of the evidence given by the corresponding measurement of the amount of glucose in Patient A is η, for some $\eta \in [0,1]$. As an example, we could interpret such value as the *degree of belief* that a medical doctor has in the truth of the statement given the evidence. As such η could be interpreted as a probability. The

[2]Here and throughout $\models$ represents classical entailment.

probabilistic interpretation is certainly favoured by the *discretization* applied to medical concepts in CADIAG2 (for example, the concept '*glucose in serum*' generates five distinct medical entities in CADIAG2: '*highly reduced glucose in serum*', '*reduced glucose in serum*', '*normal glucose in serum*', '*elevated glucose in serum*' and '*highly elevated glucose in serum*'). Notice that such an interpretation places us within the subjective probabilistic frame and thus, for the sake of coherence, the knowledge base KR should also be interpreted subjectively. Other interpretations are also possible though. For example, one could regard such values as the ratio given by the number of doctors that agree on the truth of the statement out of all the doctors involved in the assessment. In order to accommodate such values into a coherent probabilistic frame along with the statements in KR one could justify them as being *subjective* probabilities assessed by a *group* of experts –see [9] or [16] for an analysis and justification of such concept–.

Formally, let $q \in L$ represent a basic medical entity present in the patient and assume that $\eta \in [0,1]$ is the initial value assigned to it by the evaluation system of CADIAG2. We can identify the graded statement (q, η) with a probabilistic conditional statement of the form $\langle \kappa, q, \eta \rangle$, where $\kappa \in SL$ is the evidence that supports the presence of q in the patient.

Let us assume that the input of the system consists of

$$\langle \kappa_1, q_1, \eta_1 \rangle, ..., \langle \kappa_n, q_n, \eta_n \rangle,$$

for some $q_1, ..., q_n$ basic medical entities and $\eta_1, ..., \eta_n \in [0,1]$. Under this view, the set $\Omega = \{\kappa_1, ..., \kappa_n\} \subset SL$ constitutes the initial *evidence* about the patient, which is then propagated along the inference process by the application of the rules in KR^{bin}.

Within our probabilistic interpretation the reflexivity and manipulation rules in **CadL** adopt the following form, for input Φ in $\mathcal{T}$ now formally given by the above conditional statements:

$$(REF^*) \quad \frac{\langle \kappa, \phi, \eta \rangle \in \Phi}{\mathcal{T} \vdash \langle \kappa, \phi, \eta \rangle}$$

$$(C^*) \quad \frac{\langle \theta, \phi, \eta \rangle \in R^c \quad \mathcal{T} \vdash \langle \kappa, \theta, \zeta \rangle}{\mathcal{T} \vdash \langle \kappa, \phi, \min(\eta, \zeta) \rangle} \quad \text{for } \zeta > 0$$

$$(ME^*) \quad \frac{\langle \theta, \phi, 0 \rangle \in R^{me} \quad \mathcal{T} \vdash \langle \kappa, \theta, 1 \rangle}{\mathcal{T} \vdash \langle \kappa, \phi, 0 \rangle}$$

$$(AO^*) \quad \frac{\langle \theta, \phi, 1 \rangle \in R^{ao} \quad \mathcal{T} \vdash \langle \kappa, \phi, 0 \rangle}{\mathcal{T} \vdash \langle \kappa, \theta, 0 \rangle}$$

Within this frame, final outputs of the form (ϕ, η) produced by the inference engine shall be interpreted as conditionals of the form $\langle \bigwedge \Omega, \phi, \eta \rangle$ (i.e., as the probability of ϕ given all the medical evidence available about the patient). In order to make such interpretation operative and formalize it we need to

extend **CadL** by introducing two new inference rules (the extended system will be denoted by **CadL***). The first of these rules formalizes the maximization process done by the system in order to yield as output the set of medical entities (diseases) along with the maximal value generated by it, with respect to the ordering $\preceq$:

$$(MAX) \quad \frac{\mathcal{T} \vdash \langle \bigwedge \Delta_1, \phi, \eta \rangle \quad \mathcal{T} \vdash \langle \bigwedge \Delta_2, \phi, \zeta \rangle}{\mathcal{T} \vdash \langle \bigwedge (\Delta_1 \cup \Delta_2), \phi, \max^*(\eta, \zeta) \rangle}$$

for $\Delta_1, \Delta_2 \subseteq \Omega$.

An additional rule is necessary to produce the desired outcome:

$$(EX) \quad \frac{\mathcal{T} \vdash \langle \bigwedge \Delta, \phi, \eta \rangle \quad \mathcal{T} \nvdash \langle \kappa, \phi, \zeta \rangle \text{ for all } \zeta \in [0,1]}{\mathcal{T} \vdash \langle \kappa \wedge \bigwedge \Delta, \phi, \eta \rangle}$$

for $\Delta \subset \Omega$ and $\kappa \in \Omega$.

This last rule, which we call EX as abbreviation of '*exhaustive*', simply states that, if κ is a piece of evidence that says nothing about the presence of ϕ in the patient (i.e., that κ and ϕ are *independent*) then the probability of ϕ given Δ should stay the same if in addition we consider the piece of evidence κ (i.e., $\Delta \cup \{\kappa\}$).

Consider now the theory $\mathcal{T} = (\Phi, R)$, with $R = KR^{bin}$ and Φ the input of the system which, as mentioned earlier, under our probabilistic interpretation takes the form of a collection of conditional probabilistic statements

$$\langle \kappa_1, q_1, \eta_1 \rangle, ..., \langle \kappa_n, q_n, \eta_n \rangle,$$

for some $q_1, ..., q_n$ basic medical entities, $\eta_1, ..., \eta_n \in [0,1]$ and $\Omega = \{\kappa_1, ..., \kappa_n\} \subset SL$ the initial evidence about the patient. The disease ϕ along with the value η would be given as an output in a run of the inference process of CADIAG2 on input Φ only if there exists a maximal proof (defined for **CadL*** essentially as for **CadL**) of $\langle \bigwedge \Omega, \phi, \eta \rangle$ from $\mathcal{T}$ in **CadL***. In our probabilistic interpretation, a runtime inconsistency in CADIAG2 can be manifested by the existence of maximal proofs of $\langle \bigwedge \Omega, \phi, 0 \rangle$ and $\langle \bigwedge \Omega, \phi, 1 \rangle$ from $\mathcal{T}$, for some medical entity ϕ, or by the non-existence of a proof of $\langle \bigwedge \Omega, \phi, \eta \rangle$ together with the existence of a proof of a statement of the form $\langle \kappa, \phi, \zeta \rangle$ from $\mathcal{T}$ (due to the fact that max^* is not defined for $(0,1)$) –for more details on all these issues see [19]–.

5.1.1 CadL* and probabilistic soundness

Among the manipulation rules in **CadL***, probabilistic soundness of ME^* is clear (i.e., that any probability function on L that satisfies $\langle \kappa, \theta, 1 \rangle$ and $\langle \theta, \phi, 0 \rangle$ also satisfies $\langle \kappa, \phi, 0 \rangle$). So is soundness of AO^*. However, C^* is certainly not sound with respect to probabilistic semantics. Among the two new additional rules in **CadL*** introduced to provide a probabilistic interpretation of the inference, MAX is clearly not sound and EX assumes some probabilistic *independence* among entities that may not actually be independent. Overall, **CadL***

does not score well within probability theory. This is no surprise. The computation of conditional probabilistic statements in a compositional way, as done by CADIAG2 primarily by means of the *min* and max^* operators, is clearly bound to be probabilistically unsound. One may wonder though what could be done in order to improve the inference on probabilistic grounds from a knowledge base like KR^{bin}. The answer seems to be 'not much'. Certainly a KR^{bin}-like knowledge base (i.e., a knowledge base given by some binary probabilistic conditional statements) is not the most convenient for inferential purposes in probability theory for medical applications like CADIAG2. As is well known, there are other knowledge-base structures better suited for that purpose, Bayesian networks being the most celebrated among them –see for example [5] or [18]–.

It is worth noting that $\mathbf{CadL}^*$ satisfies what we can call *weak consistency* –called *weak soundness* in [11]–, defined as follows: if there is a maximal proof in $\mathbf{CadL}^*$ of a statement of the form $\langle \bigwedge \Delta, \phi, 1\rangle$ (or $\langle \bigwedge \Delta, \phi, 0\rangle$) from some theory $\mathcal{T}$, with $\phi \in SL$ and $\Delta \subset SL$ then, if there is a maximal proof in $\mathbf{CadL}^*$ of a statement of the form $\langle \bigwedge \Delta^*, \phi, \eta\rangle$, with $\Delta \subset \Delta^*$, then $\eta = 1$ (or $\eta = 0$ respectively). That is to say, if $\mathbf{CadL}^*$ concludes certainty about the occurrence of some event or about the truth or falsity of some sentence then adding new evidence does not alter this certainty. Weak consistency is provided in $\mathbf{CadL}^*$ and so in the inference mechanism of CADIAG2 by the operator $\max^*$ defined over the ordering $\preceq$.

5.2 Fuzzy semantics

The motivation for an interpretation of the inference in CADIAG2 on the grounds of a fuzzy semantics is mostly motivated by the interpretation of the degree of presence η in a graded statement of the form (ϕ, η) in the input of CADIAG2 in the natural, most intuitive way: as a membership degree (i.e., truth degree) in the context of fuzzy set theory. However, in our attempt to provide a fuzzy interpretation of the inference in CADIAG2, we also need an interpretation of the rules in the system in those same terms. As mentioned in previous sections, degrees of confirmation are intended to represent degrees of certainty about the presence of the corresponding diseases in the patient and are better characterized by means of uncertainty measures such as probability functions. Even though we acknowledge the possibility of using fuzzy semantics to model uncertainty we would rather consider a characterization of the rules of the system and the corresponding degrees of confirmation in terms of truth degrees, arguably more suitable from the point of view of the intended fuzzy semantics.

Graded statements in our settings become in this context what have been called *graded formulas* in [10] or [12]. Truth degrees in them will now be interpreted as lower-bound thresholds (i.e., η in a graded statement of the form (ϕ, η) on L will now be regarded as a lower-bound threshold for the degree of truth of ϕ). Such an interpretation is not only motivated by the fact that it constitutes the common one to fuzzy logics but also by the inference in CADIAG2

itself when interpreted on fuzzy grounds: the choice of the maximal value with respect to the ordering $\preceq$ generated in relation to a certain disease as the output value for it goes well with the characterization of any values generated at each step in the inference as lower-bound thresholds.[3]

For our fuzzy semantics, the interpretation for conjunction ($\wedge$), disjunction ($\vee$) and negation ($\sim$) suggests itself by the values (degrees of truth in this context) that the system assigns to compound medical entities in SL along the inference process. Therefore, for $v : L \longrightarrow [0,1]$ a fuzzy valuation on L, we will have the following constraints, for $\phi, \theta \in SL$:

- $v(\phi \wedge \theta) = \min(v(\phi), v(\theta))$.
- $v(\phi \vee \theta) = \max(v(\phi), v(\theta))$.
- $v(\sim \phi) = 1 - v(\phi)$.

It is common in the field of fuzzy logic to identify the interpretation of the conjunction to a *t-norm* (based on some natural, desirable properties that such an interpretation should satisfy) and the interpretation of the implication ($\rightarrow$) to its *residuum* –for more details on these notions see [12]–. Such identification places a further constraint on v:

$$v(\theta \rightarrow \phi) = \sup\{v(p) \,|\, v(\theta \wedge p) \leq v(\phi)\},$$

for $\theta, \phi \in SL$.

The identification of the interpretation of the conjunction ($\wedge$) with the Gödel *t-norm* (i.e., with the *minimum* operator) leads to the following interpretation of the implication ($\rightarrow$), for $\phi, \theta \in SL$:

$$v(\theta \rightarrow \phi) = \begin{cases} 1 & \text{if } v(\theta) \leq v(\phi) \\ v(\phi) & \text{otherwise} \end{cases}$$

In this framework we can identify an inference rule of the form $\langle \theta, \phi, \eta \rangle$ in the knowledge base of CADIAG2 –for $\theta, \phi \in SL$ and $\eta \in [0,1]$– with the graded statement $(\theta \rightarrow \phi, \eta)$.

Satisfiability of rules in KR and, in general, of any graded statements is defined in our framework as expected.

Definition 7 *The fuzzy valuation v on L is said to satisfy (ϕ, η), for some $\phi \in SL$ and $\eta \in [0,1]$, if $v(\phi) \geq \eta$.*

[3]This is not so in our probabilistic interpretation of the rules and graded statements involved in the inference process, as seen in the previous subsection. Recall that, in our probabilistic characterization, distinct values generated along the inference for the same disease (or, in general, medical entity) were intended to represent distinct degrees of certainty about the presence of such disease in the patient due mostly to differing amounts of evidence (i.e, subsets of what we denoted by Ω).

If such a valuation exists we say that (ϕ, η) is *satisfiable*.

It has to be mentioned that the residuum is not strictly necessary in order to characterize the inference rules in CADIAG2 by means of the intended semantics. Notice that, in particular for v a fuzzy valuation as defined above by the given restrictions, $v(\theta \rightarrow \phi) \geq \eta$ is equivalent to $v(\theta) \geq \min(\eta, v(\phi))$ and thus the residuum does not need to be defined. However, it makes our characterization clearer and less cumbersome.

5.2.1 CadL and fuzzy soundness.

Among the manipulation rules in **CadL**, soundness of the rule C under the intended interpretation is clear (i.e., any fuzzy valuation v that satisfies (θ, ζ) and $(\theta \rightarrow \phi, \eta)$ also satisfies $(\phi, \min(\eta, \zeta))$) and so is soundness of ME and AO. The rule C responds to what is basically called *fuzzy modus ponens*, see for example [10]. As for the evaluation rules in **CadL**, AND and OR are sound with respect to the intended semantics but NOT is not sound (due to the interpretation of the truth values in graded statements as lower-bound thresholds). As shown in [6], soundness of **CadL** can be basically provided by restricting the use of the rule NOT along the inference and by reinterpreting the intended meaning of the truth degrees in some graded statements: truth degrees in the graded statements that constitute the input in a run of the inference engine can be regarded as point values and also those in graded statements that are obtained from them by the application of any rules in **CadL** other than C. The rule NOT would only be applied to these statements (i.e., to graded statements where the truth degree is known to represent a point value). Thus, graded statements obtained as a result of the application of the rule C would not be used by the rule NOT –for more details on this point and, in general, on the content of this section see [6]–.

6 Conclusion

Two semantics have been taken as reference in our attempt to provide an interpretation of the inference process and output of the medical expert system CADIAG2: probabilistic semantics and fuzzy (*t-norm*-based) semantics. The choice of probabilistic semantics was mostly motivated by the natural identification of the *degrees of confirmation* in the rules of the system with probabilities whereas the choice of a fuzzy semantics was mostly motivated by the natural identification of the input values of the input symptoms in a run of the inference engine with *membership degrees* (i.e., truth degrees) in fuzzy set theory. In order to set the inference process on probabilistic grounds a probabilistic interpretation of the input values was needed and thus its natural interpretation, (arguably) more in keeping with a fuzzy semantics, had to be overlooked. On the other hand, in order to set the inference process fully on the grounds of a fuzzy semantics the degrees of confirmation in the rules of the system needed to

be interpreted accordingly, despite the fact that such degrees are better represented by uncertainty measures such as probabilities. This granted, we showed that both semantics could account well for several steps along the inference process, in particular the attempted *t-norm*-based fuzzy semantics –based on the identification of the *t-norm* with the *minimum* operator– yet, overall, none of them proved fully suitable as the intended interpretation of the system.

References

[1] Adlassnig, K., Kolarz, G., Scheithauer, W. & Grabner, H. (1986). *Approach to a hospital based application of a medical expert system.* Medical Informatics, 11(3), 205-223.

[2] Adlassnig, K., Kolarz, G., Effenberger, W. & Grabner, H. (1985). *CADIAG: Approaches to computer-assisted medical diagnosis.* Computers in Biology and Medicine, 15, 315-335.

[3] Adlassnig, K. (1986). *Fuzzy set theory in medical diagnosis.* IEEE Transactions on Systems, Man and Cybernetics, 16(2), 260-265.

[4] Buchanan, B.G. & Shortliffe, E.H. (eds). (1984). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project.* Reading, MA, Addison-Wesley.

[5] Castillo, E., Gutiérrez, J. & Hadi, A. (1997). *Expert Systems and Probabilistic Network Models.* Springer-Verlag, New York.

[6] Ciabattoni, A. & Vetterlein, T. (2009). *On the fuzzy (logical) content of CADIAG2.* Fuzzy Sets and Systems, 161(14), 1941-1958.

[7] Ciabattoni, A. & Rusnok, P. (2010). *On the classical content of monadic $G^{\sim}$ and its application to a fuzzy medical expert system.* Proceedings of the 12^{th} International Conference on the Principles of Knowledge Representation and Reasoning (KR2010).

[8] Daniel, M., Hájek, P. & Nguyen P.H. (1997). *CADIAG2 and MYCIN-like systems.* Artificial Intelligence in Medicine, 9, 241-259.

[9] Genest, C. & Zidek, J. (1986). *Combining probability distributions: A critique and an annotated bibliography.* Statistical Science, 1/1, 114-135.

[10] Gerla, G. (2000). *Fuzzy Logic.* Kluwer Academic Press.

[11] Hájek, P. (1988). *Towards a probabilistic analysis of mycin-like expert systems.* COMPSTAT 88 proceedings. Physica-Verlag Heidelberg.

[12] Hájek, P. (1998). *Metamathematics of Fuzzy Logic.* Kluwer Academic Press.

[13] Hájek, P. & Valdes, J. (1994). *An analysis of MYCIN-like expert systems.* Mathware and Soft Computing 1, 45-68.

[14] Heckerman, D. (1986). *Probabilistic interpretations for MYCIN's certainty factors.* In *Uncertainty in Artificial Intelligence.* North-Holland, Amsterdam, 167-196.

[15] Klir, G. & Folger, T. (1988). *Fuzzy Sets, Uncertainty and Information.* Prentice-Hall International.

[16] Osherson, D. & Vardi, M. (2006). *Aggregating disparate estimates of chance.* Games and Economic Behavior, 56, 148-173.

[17] Paris, J. B. (1994). *The Uncertain Reasoner's Companion.* Cambridge University Press.

[18] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman, San Mateo, California.

[19] Picado Muiño, D. (2011). *A probabilistic interpretation of the medical expert system CADIAG2.* Soft Computing (to appear).

[20] Seising. R. (2007). *The Fuzzification of Systems.* Springer-Verlag Berlin Heidelberg.

[21] Shortliffe, E. H. (1976). *Computer-Based Medical Consultations: MYCIN.* Elsevier, New York.

[22] Valdes, J. (1992). *An integrated environment for uncertainty processing using graphical models.* IPMU 92 proceedings.

[23] Zadeh, L. (1965). *Fuzzy sets.* Information and Control, 8, 338-353.

[24] Zadeh, L. (1975). *Fuzzy logic and approximate reasoning.* Synthese, 30, 407-428.

[25] Zimmermann, H. (1991). *Fuzzy Set Theory and its Applications.* Kluwer Academic Press.