

Research Article

Abstract Computation in Schizophrenia Detection through Artificial Neural Network Based Systems

**L. Cardoso,¹ F. Marins,¹ R. Magalhães,^{2,3} N. Marins,⁴ T. Oliveira,¹ H. Vicente,⁵
A. Abelha,¹ J. Machado,¹ and J. Neves¹**

¹University of Minho, CCTC, 4710-057 Braga, Portugal

²Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, 4710-057 Braga, Portugal

³ICVS/3B's, PT Government Associate Laboratory, Guimarães, 4710-057 Braga, Portugal

⁴Centro Hospitalar Conde de Ferreira, 4200-227 Porto, Portugal

⁵Department of Chemistry and Évora Chemistry Centre, School of Science and Technology, University of Évora, 7004-516 Évora, Portugal

Correspondence should be addressed to J. Neves; jneves@di.uminho.pt

Received 22 July 2014; Revised 17 September 2014; Accepted 28 September 2014

Academic Editor: Xiangfeng Luo

Copyright © 2015 L. Cardoso et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Schizophrenia stands for a long-lasting state of mental uncertainty that may bring to an end the relation among behavior, thought, and emotion; that is, it may lead to unreliable perception, not suitable actions and feelings, and a sense of mental fragmentation. Indeed, its diagnosis is done over a large period of time; continuous signs of the disturbance persist for at least 6 (six) months. Once detected, the psychiatrist diagnosis is made through the clinical interview and a series of psychic tests, addressed mainly to avoid the diagnosis of other mental states or diseases. Undeniably, the main problem with identifying schizophrenia is the difficulty to distinguish its symptoms from those associated to different untidiness or roles. Therefore, this work will focus on the development of a diagnostic support system, in terms of its knowledge representation and reasoning procedures, based on a blended of Logic Programming and Artificial Neural Networks approaches to computing, taking advantage of a novel approach to knowledge representation and reasoning, which aims to solve the problems associated in the handling (i.e., to stand for and reason) of defective information.

1. Introduction

Schizophrenia is a brain disorder that strikes people as they are entering the prime of their life and, in many cases, may run a recurrent and ultimately chronic course that will lead to substantial disability. Indeed, this highly destructive mental illness affects the essence of what makes people human, that is, their personality and intellect [1].

The incidence of this disease is roughly one new patient per year per ten thousand human beings. In the one hand it occurs in all cultures, in spite of more schizophrenic ones appearing in the lower socioeconomic classes of the industrialized countries; on the other hand, the admission rates for schizophrenia are higher in the urban areas than in

the rural ones. The gender distribution is approximately equal in both sexes, and the peak of incidence is situated between 15 and 25 years in males, and between 25 and 35 years in females [1, 2].

There are also several risk factors that must be considered, namely, the genetic ones like predisposing, pregnancy, and birth complications. Indeed, and to some extent, some precipitating circumstances, such as family interactions (e.g., dysfunctional families and expressed emotion, a kind of behavior with overt criticism and hostility), life events (e.g., happenings that provokes a high level of stress), or drug abuse, must also be object of attention [2].

When working on the diagnosis of schizophrenia, the challenging question with which practitioners are met relies

on the similarity between the signs and symptoms among psychic diseases and states. Typically, the earliest act for the diagnosis of schizophrenia proceeds from the detection of a specific outbreak, ordinarily evinced by hallucinations, delusions, disorganized speech, catatonic behavior, abolition, and social separation or even additional personality disorders. The psychiatric diagnosis is made through the clinical interview based on psychopathology, but the problem is that knowledge of psychopathology of anthology depends on the training and experience of the psychiatrist. These limitations “open the door” to mistake and doubt. From here, the practitioners may perform psychic exams intended to detect cognitive flaws, delusional perception, and abnormal changes or thoughts, among others [1, 2].

Above any kind of dispute, the main problem with the diagnosis of schizophrenia comes from the large number of different mental diseases and states that may mimic its signs, such as epilepsy, drug-induced psychosis, affective psychoses (e.g., bipolar, major depressive disorders), Asperger syndrome, schizoaffective disorder, and Wilson’s disease. To face this situation the practitioner is forced to order extra exams, namely, toxicological tests, electroencephalography (EEG) and brain computed tomography (CT)’s [1]. Also this pinpointing is usually done over a large period of time (at least 6 months), therefore generating a huge amount of data, which has to be treated and interpreted by the practitioner(s).

Facing such a large amount of facts, even experienced psychiatrists have difficulties to make a precise diagnosis and distinguishing between this and other diseases. Unfortunately, there is not too much work done in the area of Medical Informatics to help in this process, although Strous et al. [3] worked on a system that uses the differences on writing to diagnose schizophrenia.

With this paper we make a start on the development of an unusual or original diagnosis assistance system for schizophrenia. We will present a logic programming based approach in order to represent the knowledge and reasoning, with a focus on the Degree of Confidence (DoC) of the attributes set, that makes a function or a predicate [4].

2. Related Work

Many studies presenting the concept of uncertainty and/or “imperfect data” like [5–8] show that there is an emergent interest in the problem of uncertainty as compared to accuracy or error in data [9, 10]. The notion of uncertainty is broader than error or accuracy and includes these more restricted concepts. While accuracy is the closeness of measurements or computations to their “true” value or some value agreed to be the “truth”, uncertainty can be considered any aspect of the data that results in less than perfect knowledge about the phenomena being studied [11–13].

The consequences of the uncertainty in the data and in the data quality to user exploring, modelling, visualizing, and

querying have been also referred to in many studies [6, 8, 14–18]. On the one hand, it is consensual that when the data are uncertain, a different representation and uncertainty which can be reduced by “acquiring additional information or improving the quality of the information available” [6] is needed; that is, in almost all decisions that one may take, the information is not always exact, but indeed imperfect, in the sense that we handle estimated values, probabilistic measures, or degrees of uncertainty [19, 20]. On the other hand, knowledge and belief are generally incomplete, contradictory, or even error sensitive. It is advisable the use of formal tools to deal with the problems that arise with the use of partial, contradictory, ambiguous, imperfect, nebulous, or missing information [13, 21–23]. Some general models have been presented [8, 16, 24], where uncertainty is associated to the application of Probability Theory [16, 17], Fuzzy Set Theory [25], Similarities [26, 27]. Other approaches for knowledge representation and reasoning have been proposed using the Logic Programming (LP) paradigm, namely, in the area of Model Theory [28–30] and Proof Theory [29, 31, 32]. Qualitative models and qualitative reasoning have been around in Artificial Intelligence research for some time [14, 33], in particular due to the growing need to offer support in decision-making processes. The evaluation of knowledge that stems out from logic programs becomes a point of research. In this sense, the evaluation of knowledge that stems out from logic programs becomes a point of research. Shi et al. [15] and Schneider [25] work is a good example of quality evaluation using logic. They used abduction [14] and temporal logic for quality-checking of medical guidelines, proposing a method to diagnose potential problems in a guideline, regarding the fulfillment of general medical quality criteria at a metalevel characterization. They explored an approach which uses a relational translation to map the temporal logic formulas to first-order logic and a resolution-based theorem prover [25]. In another research line, the Quality-of-Information concept (QoI) [6, 7, 26, 34] demonstrated their applicability in many dynamic environments and for decision making purposes. The objective is to build a quantification process of the QoI and an assessment of the argument values of a given predicate with relation to their domains (here understood as Degree-of-Confidence (DoC)), which stems from a logic program or theory during the evolutive process of searching solutions in order to solve a problem in environments with default data. Our main contribution relies on the fact that at the end, the initial extensions of the predicates that make the universe of discourse are given in terms of DoCs predicates that stand for one’s confidence that the initial predicates arguments values of the predicates that make the universe of discourse fit into their respective domains. This approach potentiates the use of diverse computational paradigms, like Case Based Reasoning [35], Artificial Neural Networks (ANN) [36, 37], and Particle Swarm [38], just to name a few. It also encapsulates, in itself a new vision of Multivalued Logics, once a proof of a theorem in a conventional way is evaluated to the interval [0, 1]. Following these studies good results were achieved for different purposes, namely, in Healthcare [27], Civil Law [39],

Multiagent Systems [40], Virtual Entities [34], Ambient Assisted Living [41], and Decision Making Environments [30]. Indeed, some interesting results have been obtained, namely, in the fields of Coronary Risk Evaluation [42], Hyperactivity Disorder [43], and Length of Hospital Stay [44] among others, which fit the case studies insufficiency that are common in this area of research and applications.

3. Knowledge Representation and Reasoning

We follow the proof theoretical approach and an extension to the LP language [31], to knowledge representation and reasoning. An Extended Logic Program (ELP) stands for a finite set of clauses in the form:

$$q \leftarrow p_1 \wedge p_n \wedge \text{not}q_1 \wedge \dots \wedge \text{not}q_m \quad (1)$$

$$?(p_1 \wedge \dots \wedge p_n \wedge \text{not}q_1 \wedge \dots \wedge \text{not}q_m) \quad (n, m \geq 0),$$

where ? is a domain atom denoting falsity and the p_i, q_j , and p are classical ground literals, that is, either positive atoms or atoms preceded by the classical negation sign \neg [32]. Under this representation formalism, every program is associated with a set of abducibles [28, 30], given here in the form of exceptions to the extensions of the predicates that make the program. Once again, LP has emerged as an attractive formalism for knowledge representation and reasoning tasks, introducing an efficient search mechanism for problem solving. Due to the growing need to offer user support in decision-making processes some studies have been presented [45, 46], related to the qualitative models and qualitative reasoning in Database Theory and in Artificial Intelligence (AI) research. With respect to the problem of knowledge representation and reasoning mechanisms in LP, a measure of the Quality of Information (QoI) of such programs has been object of some work with promising results [47–49], that is, the QoI denotes one’s confidence that a particular term of a predicate that makes the universe of discourse belongs to its extension. Indeed the QoI indeed [31] with respect to the extension of a predicate i will be given by a truth-value in the interval $[0, 1]$; that is, if the information is known (positive) or false (negative) the QoI for the extension of predicate i is 1. For situations where the information is unknown, the QoI is given by

$$\text{QoI}_i = \lim_{N \rightarrow \infty} \frac{1}{N} = 0 \quad (N \gg 0), \quad (2)$$

where N denotes the cardinality of the set of terms or clauses of the extension of predicate i that stand for the incompleteness under consideration. For situations where

the extension of predicate i is unknown but can be taken from a set of values, the QoI is given by

$$\text{QoI}_i = \frac{1}{\text{Card}}, \quad (3)$$

where Card denotes the cardinality of the abducibles set for i , if the abducibles set is disjoint. If the abducibles set is not disjoint, the QoI is given by

$$\text{QoI}_i = \frac{1}{(C_1^{\text{Card}} + \dots + C_{\text{Card}}^{\text{Card}})}, \quad (4)$$

where $C_{\text{Card}}^{\text{Card}}$ is a card-combination subset, with Card elements. The next element of the model to be considered is the relative importance that a predicate assigns to each of its attributes under observation, that is, w_i^k , which stands for the relevance of attribute k in the extension of predicate $_i$. It is also assumed that the weights of all the attribute predicates are normalized, that is

$$\sum_{1 \leq k \leq n} w_i^k = 1, \quad \forall_i, \quad (5)$$

where \forall denotes the universal quantifier. It is now possible to define a predicate $_i$ scoring function $V_i(x)$ so that, for a value $x = (x_1, \dots, x_n)$, defined in terms of the attributes of predicate $_i$, one may have

$$V_i(x) = \sum_{1 \leq k \leq n} w_i^k \times \frac{\text{QoI}_i(x)}{n}. \quad (6)$$

It is now feasible to rewrite the extensions of the predicate referred to above, according to productions of the type

$$\text{predicate}_i(x_1, \dots, x_n) :: \text{QoI} :: \text{DoC} \quad (7)$$

and evaluate the Degree of Confidence (DoC) given by $\text{DoC} = V_i(x_1, \dots, x_n)/n$ which denotes one’s confidence that the argument values of predicate $_i$ fit into its domain values. To be more general, let us suppose that the Universe of Discourse is described by the extension of the predicates:

$$a_1(\dots), a_2(\dots), \dots, a_n(\dots) \quad (n \geq 0). \quad (8)$$

Therefore one may have (where \perp denotes an argument value of the type unknown; the values of the others arguments stand for themselves):

$$\left\{ \begin{array}{l} \neg a_1(x_1, y_1, z_1) \leftarrow \text{not } a_1(x_1, y_1, z_1) \\ a_1(\perp, [10, 20], 15) :: 1 :: \text{DoC} \\ \underline{[5, 10]} \quad \underline{[5, 30]} \quad \underline{[10, 20]} \\ \text{attributes's domains for } x_1, y_1, z_1 \\ \vdots \\ \neg a_2(x_2, y_2, z_2) \leftarrow \text{not } a_2(x_2, y_2, z_2) \\ a_2([45, 54], [10, 12], \perp) :: 1 :: \text{DoC} \\ \underline{[30, 60]} \quad \underline{[6, 14]} \quad \underline{[2000, 6000]} \\ \text{attributes's domains for } x_2, y_2, z_2 \\ \vdots \end{array} \right.$$

⇓ 1st interaction: transition to continuous intervals

$$\left\{ \begin{array}{l} \neg a_1(x_1, y_1, z_1) \leftarrow \text{not } a_1(x_1, y_1, z_1) \\ a_1([5, 10], [10, 20], 15) :: 1 :: \text{DoC} \\ \underline{[5, 10]} \quad \underline{[5, 30]} \quad \underline{[10, 20]} \\ \text{attributes's domains for } x_1, y_1, z_1 \\ \vdots \\ \neg a_2(x_2, y_2, z_2) \leftarrow \text{not } a_2(x_2, y_2, z_2) \\ a_2([45, 54], [10, 12], [2000, 6000]) :: 1 :: \text{DoC} \\ \underline{[30, 60]} \quad \underline{[6, 14]} \quad \underline{[2000, 6000]} \\ \text{attributes's domains for } x_2, y_2, z_2 \\ \vdots \end{array} \right.$$

⇓ 2nd interaction: normalization: $\frac{Y - Y_{\min}}{Y_{\max} - Y_{\min}}$

$$\left\{ \begin{array}{l} \neg a_1(x_1, y_1, z_1) \leftarrow \text{not } a_1(x_1, y_1, z_1) \\ a_1\left(\left[\frac{5-5}{10-5}, \frac{10-5}{10-5}\right], \left[\frac{10-5}{30-5}, \frac{20-5}{30-5}\right], \left[\frac{15-10}{20-10}, \frac{15-10}{20-10}\right]\right) \\ \equiv a_1([0, 1], [0.2, 0.6], [0.5, 0.5]) :: 1 :: \text{DoC} \\ \underline{[0, 1]} \quad \underline{[0, 1]} \quad \underline{[0, 1]} \\ \text{attributes's domains for } x_1, y_1, z_1 \\ \vdots \\ \neg a_2(x_2, y_2, z_2) \leftarrow \text{not } a_2(x_2, y_2, z_2) \\ a_2\left(\left[\frac{45-30}{60-30}, \frac{54-30}{60-30}\right], \left[\frac{10-6}{14-6}, \frac{12-6}{14-6}\right], \left[\frac{2000-2000}{6000-2000}, \frac{6000-2000}{6000-2000}\right]\right) \\ \equiv a_2([0.5, 0.8], [0.5, 0.75], [0, 1]) :: 1 :: \text{DoC} \\ \underline{[0, 1]} \quad \underline{[0, 1]} \quad \underline{[0, 1]} \\ \text{attributes's domains for } x_2, y_2, z_2 \\ \vdots \end{array} \right.$$

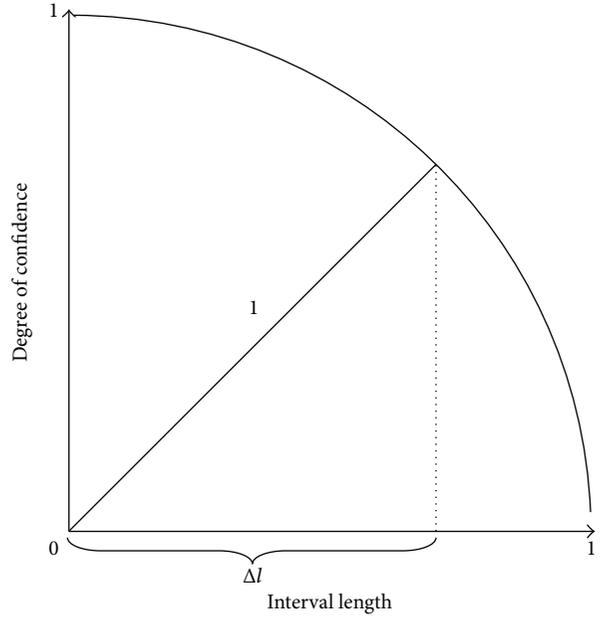


FIGURE 1: Evaluation of the Degree of Confidence.

The Degree of Confidence (DoC) is now evaluated using $\text{DoC} = \sqrt{1 - \Delta l^2}$, as it is illustrated in Figure 1. Here Δl stands for the length of the arguments intervals, once normalized.

As shown in Figure 1, one has the expected representation of the universe of discourse, where all the predicates' arguments are nominal. They speak for one's confidence that the unknown values of the arguments fit into the correspondent intervals referred to above

(9)

$$\left\{ \begin{array}{l} \neg a_{1_{\text{DoC}}}(x_1, y_1, z_1) \leftarrow \text{not } a_{1_{\text{DoC}}}(x_1, y_1, z_1) \\ a_{1_{\text{DoC}}}(0, 0.916, 1) :: 0.5 :: \text{DoC} \\ \underline{[0, 1]} \quad \underline{[0.2, 0.6]} \quad \underline{[0.5, 0.5]} \\ \text{attributes's values range for } x_1, y_1, z_1 \\ \underline{[0, 1]} \quad \underline{[0, 1]} \quad \underline{[0, 1]} \\ \text{attributes's domain for } x_1, y_1, z_1 \\ \vdots \\ \neg a_{2_{\text{DoC}}}(x_2, y_2, z_2) \leftarrow \text{not } a_{2_{\text{DoC}}}(x_2, y_2, z_2) \\ a_{2_{\text{DoC}}}(0.954, 0.968, 0) :: 0.6 :: 0.641 \\ \underline{[0.5, 0.8]} \quad \underline{[0.5, 0.75]} \quad \underline{[0, 1]} \\ \text{attributes's values ranges for } x_2, y_2, z_2 \\ \underline{[0, 1]} \quad \underline{[0, 1]} \quad \underline{[0, 1]} \\ \text{attributes's domain for } x_2, y_2, z_2 \\ \vdots \end{array} \right. \quad (10)$$

4. One's Model

Therefore, and in order to exemplify the applicability of our approach to default knowledge, we will look at the relational database model, since it provides a basic framework that fits

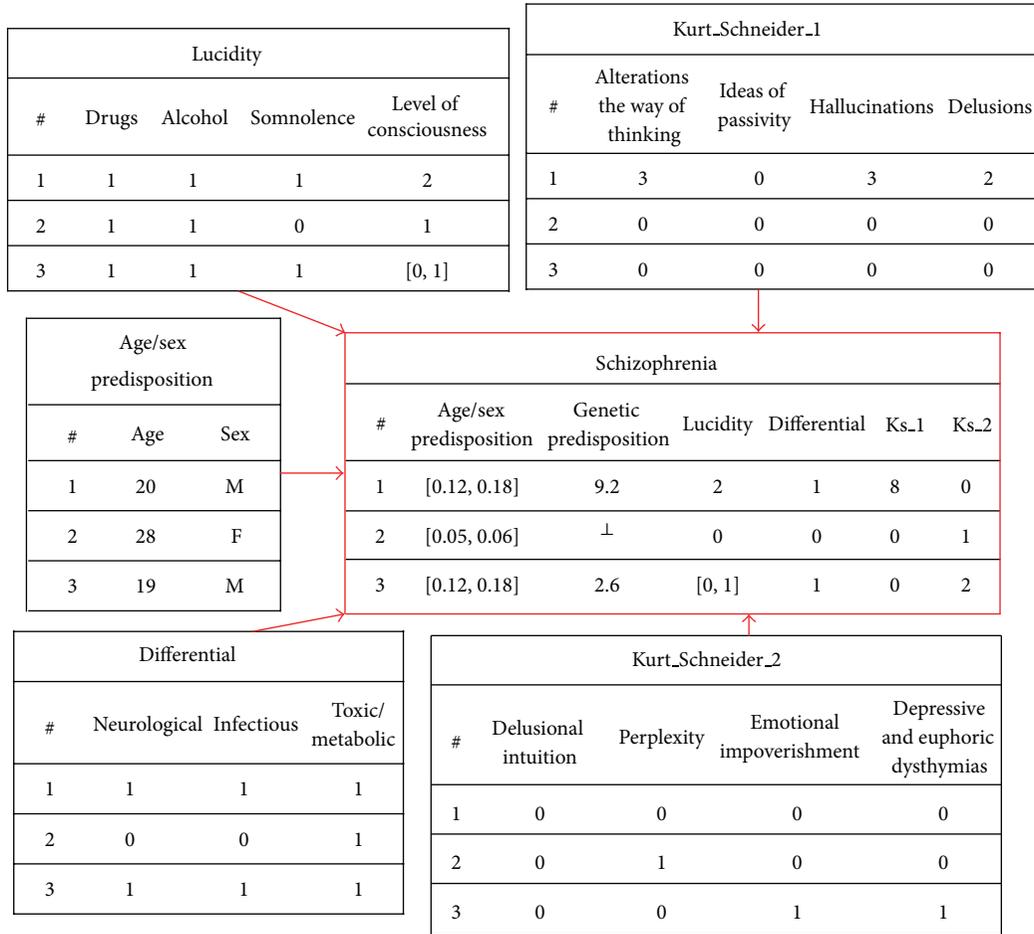


FIGURE 2: An extension of the relational database model.

into our expectations [50] and is understood as the genesis of the LP approach to knowledge representation and reasoning.

Consider the scenario where a relational database is given in terms of the extensions of the relations or predicates depicted in Figure 2, which stands for a situation where one has to manage information about schizophrenia disease detection. Under this scenario some incomplete data is also available. For instance, to patient “2” the genetic predisposition value is unknown, while to patient “3” the lucidity (in both the extensions of the schizophrenia relation or predicate) is in the interval [0, 1], although the domain values to genetic predisposition range in the interval [0.9, 91.5], and the ones to lucidity are bound by the interval [0, 2].

The psychiatrist fills the tables that link to the schizophrenia table while he/she executes the psychopathologic exam. Some symptoms may be detected by the psychiatrist; others may be perceived by additional exams (e.g., this happens with some of the attributes of the Differential/Lucidity table). At this point the psychiatrist may fill some attributes with the unknown symbol (\perp) and update the tables later.

The age/sex predisposition parameter, which is evidenced in the schizophrenia table in Figure 2, is based on Table 1 adapted from [1]. These predisposition values are clustered by age group and sex. Thus, the domain for this parameter

TABLE 1: Age/sex predisposition, adapted from [1].

Age/sex predisposition (Annual rate per 1000)		
Age group	Male	Female
0–14	[0, 0.18]	[0, 0.07]
15–24	[0.12, 0.18]	[0.06, 0.07]
25–34	[0.06, 0.12]	[0.05, 0.06]
35–44	[0.03, 0.06]	[0.03, 0.05]
45–54	[0.02, 0.03]	[0.03, 0.04]
55–64	[0.01, 0.02]	[0.02, 0.04]
>65	[0, 0.01]	[0, 0.02]

is in the range [0, 0.18] for males, and for females is in the range [0, 0.07]. The genetic predisposition parameter, which is also evidenced in the schizophrenia table in Figure 2, is based on Table 2 adapted from [51]; its domain is in the range [0.9, 91.5].

Lucidity is given in terms of the parameters or arguments that make a piece of Lucidity Table (Figure 2). The first three parameters are populated with a value between [0, 1], wherein 0 (zero) denotes the presence of drugs and 1 (one)

TABLE 2: Genetic predisposition, adapted from [51].

Genetic predisposition (%)	
General population	0.9
Foster brothers	1.8
Spouses	2.1
First cousins	2.6
Nieces and nephews	3.9
Grandchildren	4.3
Stepbrothers	7.1
Parents	9.2
Brothers	14.2
Dizygotic twin	14.5
Dizygotic twin of the same sex	17.6
Sons and daughters	16.4
Children of father and mother schizophrenics	39.2
Monozygotic twin living separately	77.6
Monozygotic twin living together	91.5

the absence of drugs, alcohol or somnolence. The level of consciousness ranges in the interval $[0, 2]$, wherein zero means that the patient is in a torpor state; one means that the patient presents obnubilation of the consciousness; and two means he is in a waking state. The differential diagnoses are represented in the differential table of the (Figure 2). Like the table lucidity, it is populated with values between $[0, 1]$, meaning the presence or absence of neurological, infectious or toxic/metabolic diseases. The lucidity and the differential diagnoses are important factors in the schizophrenia diagnosis, because if the patient is not lucid or he/she has one of the diseases previously mentioned, the diagnosis is compromised. In other words, if one of the parameters of the lucidity and differential tables is zero, the diagnosis is not settled properly. Thus, the lucidity and differential values in the schizophrenia table are calculated through the multiplication of each parameter of table lucidity by the ones of the differential one. In this way the domain of lucidity parameter in the schizophrenia table is in the range $[0, 2]$, while the domain to the differential value is in the range $[0, 1]$.

The tables Kurt_Schneider_1 and Kurt_Schneider_2 represent the schizophrenia symptoms of first and second order delineated by Kurt Schneider [1, 2]. Relative to the table Kurt_Schneider_1 (Ks.1) there are four clusters of symptoms. The former one, alterations to the way of thinking, ranges in the interval $[0, 3]$, with the meaning: 0 (zero), absence of alterations; 1 (one), obsessive ideas; 2 (two), escape of ideas; 3 (three), loudness, theft, broadcasting, insertion, interruption, and diffusion of the thinking. The ideas of passivity are categorized in the interval $[0, 2]$, with the connotation: (0) zero, absence of feelings, impulses or somatic sensations arising from outside that imposed on the patient; 1 (one), feelings, impulses or psychomotor activities with constant doubt or uncertainty; 2 (two), feelings, impulses, psychomotor activities, or somatic sensations strange to himself and controlled by external agents to himself. Hallucinations are rated in the interval $[0, 4]$, with the connotation: 0 (zero),

absence of hallucinatory phenomena; 1 (one), elementary hallucinations; 2 (two), visual hallucinations; 3 (three), gustatory, olfactory, and kinesthetic hallucinations; 4 (four), auditory hallucinations. The latest, delusions, is grouped in the interval $[0, 2]$: 0 (zero), absence of delusional ideas; 1 (one), delusional ideas assigning meanings of affective basis; 2 (two), delusional perception assigning a new meaning of autoreferential way. In Kurt_Schneider_2 (Ks.2) table all symptoms, showed in the (Figure 2), range in the interval $[0, 1]$, wherein 0 (zero) denotes the absence of the symptom and 1 (one) its presence. Once these two tables are populated, the values of Ks.1 and Ks.2 in the schizophrenia table are the sum of the values of respective symptoms. In this way these values range in the intervals $[0, 11]$ for Ks.1, and $[0, 4]$ for Ks.2. The higher these values are, more evident is the disease. It is of great significance to note that first order symptoms have a much higher influence than the second order ones in schizophrenia detection.

Let us now consider the extensions of the relations given in (Figure 2) to populate the extension of the schizophrenia predicate:

$$\begin{aligned} \text{schizophrenia: Age/Sex Predisposition, Genetic Predisposition,} \\ \text{Lucidity, Differential, Ks.1, Ks.2} \longrightarrow \{0, 1\}, \end{aligned} \quad (11)$$

where 0 (zero) and 1 (one) denote, respectively, the truth-values false and true. It is now possible to give the extensions of the predicate referred to above, which may be set in the form:

$$\begin{aligned} \{ \\ \neg \text{sch}_{(\text{izophrenia})} (AS_{(\text{predisposition})}, G_{(\text{predisposition})}, Luc_{(\text{idity})}, \\ \text{Dif}_{(\text{ferential})}, Ks.1, Ks.2) \\ \longleftarrow \text{not sch} (AS, G, Luc, Dif, Ks.1, Ks.2) \\ \text{sch} ([0.12, 0.18], 9.2, 2, 1, 8, 0) :: 1 :: \text{DoC} \\ \underline{[0, 0.18] \quad [0.9, 91.5] \quad [0, 2] \quad [0, 1] \quad [0, 11] \quad [0, 4]} \\ \text{attribute's values ranges} \\ \text{sch} ([0.05, 0.06], \perp, 0, 0, 0, 1) :: 1 :: \text{DoC} \\ \underline{[0, 0.07] \quad [0.9, 91.5] \quad [0, 2] \quad [0, 1] \quad [0, 11] \quad [0, 4]} \\ \text{attribute's values ranges} \\ \text{sch} ([0.12, 0.18], 2.6, [0, 1], 1, 0, 2) :: 1 :: \text{DoC} \\ \underline{[0, 0.18] \quad [0.9, 91.5] \quad [0, 2] \quad [0, 1] \quad [0, 11] \quad [0, 4]} \\ \text{attribute's values ranges} \\ \} . \end{aligned} \quad (12)$$

In this program, the first clause denotes the closure of predicate schizophrenia. The next three clauses correspond to the three cases presented in the extension of the schizophrenia relation in (Figure 2). For example, the second clause corresponds to the patient 2 (two), a 28 (twenty eight)

years old woman. Analyzing (Table 1), the psychiatrist fills its age/sex predisposition within the interval [0.05, 0.06], face a domain that ranges between [0, 0.07]. Looking at its genetic predisposition, the psychiatrist does not have any information. In this case the symbol “⊥” is used, which stands for an unknown value. It subsumes that this variable may take any value according to its domain, that is, in the range ([0.9, 91.5]).

It is now possible to have the arguments of the predicates extensions normalized to the interval [0, 1], in order to compute one’s confidence that the nominal value of the arguments under considerations fit into the intervals depicted previously. The normalization process presented below is with respect to patient 2 (two).

Consider

$$\begin{aligned}
 & \{ \\
 & \neg \text{sch}_{(\text{izophrenia})} \left(\text{AS}_{(\text{predisposition})}, \text{G}_{(\text{predisposition})}, \text{Luc}_{(\text{idity})}, \right. \\
 & \quad \left. \text{Dif}_{(\text{ferential})}, \text{Ks.1}, \text{Ks.2} \right) \\
 & \Downarrow \\
 & \text{sch} ([0.05, 0.06], \perp, 0, 0, 0, 1) :: 1 :: \text{DoC} \\
 & \underline{[0, 0.07] \quad [0.9, 91.5] \quad [0, 2] \quad [0, 1] \quad [0, 11] \quad [0, 4]} \\
 & \quad \text{attribute's domain} \\
 & \Downarrow \text{1st interaction: transition to continuous intervals} \\
 & \text{sch} ([0.05, 0.06], [0.9, 91.5], [0, 0], [0, 0], \\
 & \quad [0, 0], [1, 1]) :: 1 :: \text{DoC} \\
 & \underline{[0, 0.07] \quad [0.9, 91.5] \quad [0, 2] \quad [0, 1] \quad [0, 11] \quad [0, 4]} \\
 & \quad \text{attribute's domain} \\
 & \Downarrow \text{2nd interaction: normalization: } \frac{Y - Y_{\min}}{Y_{\max} - Y_{\min}} \\
 & \text{sch} ([0.71, 0.86], [0, 1], [0, 0], [0, 0], [0, 0], \\
 & \quad [0.25, 0.25]) :: 1 :: \text{DoC} \\
 & \underline{[0, 1] \quad [0, 1]} \\
 & \quad \text{attribute's domains} \\
 & \Downarrow \text{DoC calculation: } :: \text{DoC} = \sqrt{1 - \Delta I^2} \\
 & \text{sch}_{\text{DoC}} (0.989, 0, 1, 1, 1, 1) :: 1 :: 0.83 \\
 & \underline{[0.71, 0.86] \quad [0, 1] \quad [0, 0] \quad [0, 0] \quad [0, 0] \quad [0.25, 0.25]} \\
 & \quad \text{attribute's values range} \\
 & \underline{[0, 1] \quad [0, 1]} \\
 & \quad \text{attribute's domains} \\
 & \} . \\
 \end{aligned} \tag{13}$$

Its terms make the training and test sets of the Artificial Neural Network given below (Figure 3).

5. The Artificial Neural Network Topology

The previously presented diagnosis model works well to demonstrate how all the information comes together to form a diagnosis, but it was built with the pure objective of demonstration. In this section, more reliable ways to assemble this information are considered. Different Artificial Intelligence (AI) and data mining tools can be used. Neves et al. [36, 37] demonstrated how a Artificial Neural Network (ANN) can be successfully used to model data and capture complex relationships between inputs and outputs. ANNs simulate the structure of the human brain being populated by multiple layers of neurons. The knowledge representation model defined in this work is appropriate for use with ANNs.

Using a real case of 22-year-old male patient with no schizophrenic relatives, an experienced psychiatrist, during the diagnosis elaboration, verified a full lucidity state of the patient (drugs: 1; alcohol: 1, somnolence: 1; level of consciousness: 2). In the differential diagnoses the symptoms verified were: neurological: 1; infectious: 1; toxic/metabolic: ⊥ (it was not possible to detect). The patient also presented: theft, broadcasting, and diffusion of the thinking; absence of feelings, impulses or somatic sensations arising from outside that imposed on the patient; absence of hallucinatory phenomena; delusional perception assigning a new meaning of autoreferential way. Therefore, the values inserted in the Ks.1 table were: alterations the way of thinking: 3; ideas of passivity: 0; Hallucinations: 0; Delusions: 2. Regarding the Kurt Scheneider symptoms of second order, the psychiatrist detected a weak presence of emotional impoverishment symptoms and the absence of the other three symptoms (delusional intuition: 0; perplexity: 0; emotional impoverishment: [0.25, 0.75]; depressive and euphoric dysthymias: 0). In this way, the schizophrenia clauses obtained are sch([0.12, 0.18], 0.9, 2, ⊥, 5, [0.25, 0.75]) :: 1 :: DoC and sch_{DoC}(0.944, 1, 1, 0, 1, 0.992) :: 1 :: 0.82. In (Figure 3) it is shown how the normalized values of the interval extremes and theirs DoC values work as inputs to the ANN. The output translates the chance of having schizophrenia and the DoC that one has on such a result. In addition, it was built a database of study cases that were used to train and test the ANNs. As (Figure 3) demonstrates, it was detected schizophrenia in the patient with a DoC = 0.94.

6. Conclusions

The diagnosis of schizophrenia is a hard and complex task which needs to consider many different conditions with intricate relations among them. These characteristics put this problem into the area of problems that may be tackled by AI techniques. Despite that, little to no work has been done in that direction.

In this work the founding of a computational framework that uses powerful knowledge representation techniques to represent and structure the information. is presented. This representation is above everything else, very versatile and capable of covering every possible instance representation by considering incomplete and unknown data. This finding has

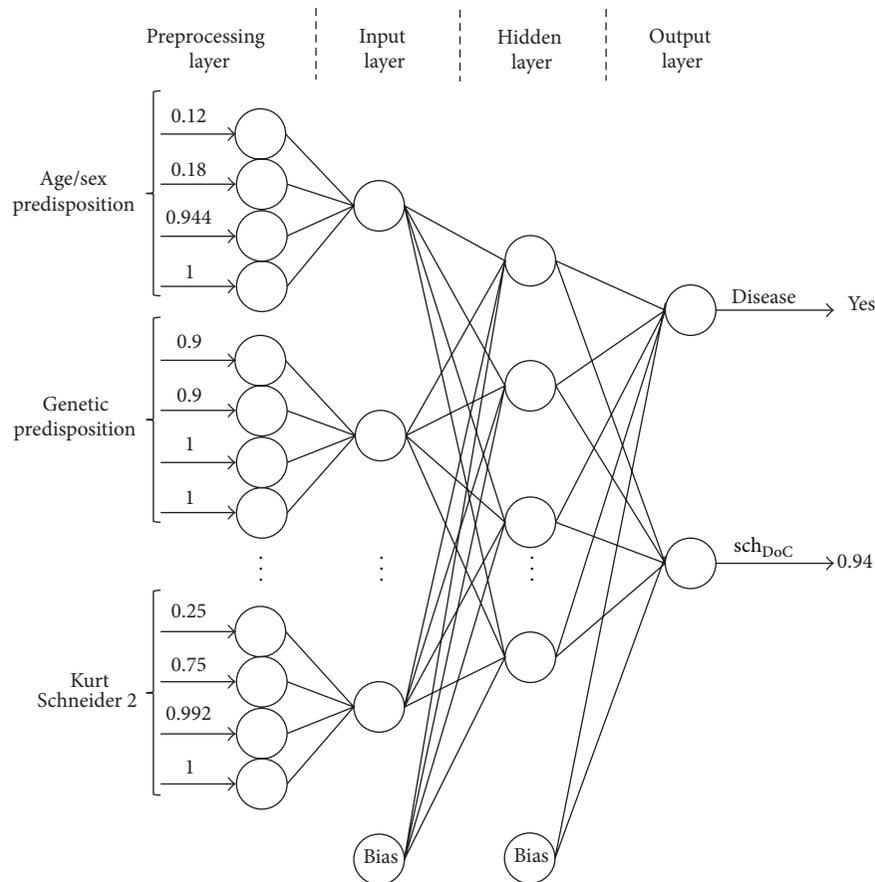


FIGURE 3: The ANN for the 22-year-old male.

several reasons, namely, that data is not equal to information; the translation of the raw measurements into interpretable and actionable read-outs is challenging; and read-outs can deliver markers and targets candidates without preconception; that is, knowing how human conditions and risk factors may affect schizophrenia detection.

7. Future Work

The knowledge representation and reasoning techniques presented above are very versatile and capable of covering every possible instance by considering incomplete, contradictory, and even unknown data. Indeed, the new paradigm for knowledge representation and reasoning enables one to use the normalized values of the interval boundaries and their DoC values, as inputs to different computational paradigms. The output translates the problem solution and the systems confidence on such a happening. Indeed the same problem must be approached using other computational frameworks like Case Based Reasoning [35] or Particle Swarm [38], among others.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is funded by National Funds through the FCT, Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within projects PEST-OE/EEI/UI0752/2014 and PEst-OE/QUI/UI0619/2012.

References

- [1] J. A. Liberman, T. S. Stroup, and D. O. Perkins, *Textbook of Schizophrenia*, The American Psychiatric Publishing, 1st edition, 2006.
- [2] S. Frangou and R. M. Murray, *Schizophrenia*, Martin Dunitz, 2nd edition, 2000.
- [3] R. D. Strous, M. Koppel, J. Fine, S. Nachliel, G. Shaked, and A. Z. Zivotofsky, "Automated characterization and identification of schizophrenia in writing," *The Journal of Nervous and Mental Disease*, vol. 197, no. 8, pp. 585–588, 2009.
- [4] J. Neves, J. Ribeiro, P. Pereira et al., "Evolutionary intelligence in asphalt pavement modeling and quality-of-information," *Progress in Artificial Intelligence*, vol. 1, no. 1, pp. 119–135, 2012.
- [5] G. J. Hunter and K. Beard, "Understanding error in spatial databases," *Australian Surveyor*, vol. 37, no. 2, pp. 108–119, 1992.
- [6] G. J. Hunter, "Managing uncertainty in GIS," in *Geographic Information Science*, vol. 2, pp. 633–641, 1999.
- [7] J. Zhang and M. F. Goodchild, *Uncertainty in Geographical Information*, CRC Press, New York, NY, USA, 2002.

- [8] J. Gottsegen, D. R. Montello, and M. F. Goodchild, "A comprehensive model of uncertainty in spatial data," in *Spatial Accuracy Assessment: Land Information Uncertainty in Natural Resources*, pp. 175–181, 1999.
- [9] L. Wang, J. Tao, R. Ranjan et al., "G-Hadoop: MapReduce across distributed data centers for data-intensive computing," *Future Generation Computer Systems*, vol. 29, no. 3, pp. 739–750, 2013.
- [10] L. Wang, Y. Ma, A. Zomaya, R. Ranjan, and D. Chen, "A parallel file system with application aware data layout policies in digital earth," *IEEE Transactions on Parallel and Distributed Systems*, 2014.
- [11] M. Menzel, R. Ranjan, L. Wang, S. Khan, and J. Chen, "Cloud-genius: a hybrid decision support method for automating the migration of webapplication clusters to public clouds," *IEEE Transactions on Computers*, no. 99, p. 1, 2014.
- [12] Z. Xu, X. Wei, X. Luo et al., "Knowle: a semantic link network based system for organizing large scale online news events," *Future Generation Computer Systems*, vol. 43–44, pp. 40–50, 2015.
- [13] Y. Liu, L. Chen, X. Luo, L. Mei, C. Hu, and Z. Xu, "Semantic link network based model for organizing multimedia big data," *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 3, pp. 376–387, 2014.
- [14] A. Pang, "Visualizing uncertainty in geo-spatial data," in *Proceedings of the Workshop on the Intersections between Geospatial Information and Information Technology*, 2001.
- [15] W. Shi, P. Fisher, and M. F. Goodchild, *Spatial Data Quality*, CRC Press, New York, NY, USA, 2003.
- [16] P. Fisher, A. Comber, and R. Wadsworth, "Approaches to uncertainty in spatial data," in *Fundamentals of Spatial Data Quality*, pp. 43–59, 2006.
- [17] R. Li, B. Bhanu, C. Ravishankar, M. Kurth, and J. Ni, "Uncertain spatial data handling: modeling, indexing and query," *Computers and Geosciences*, vol. 33, no. 1, pp. 42–61, 2007.
- [18] T. Hong, K. Hart, L.-K. Soh, and A. Samal, "Using spatial data support for reducing uncertainty in geospatial applications," *GeoInformatica*, vol. 18, no. 1, pp. 63–92, 2014.
- [19] X. Luo, Z. Xu, J. Yu, and X. Chen, "Building association link network for semantic link on web resources," *IEEE Transactions on Automation Science and Engineering*, vol. 8, no. 3, pp. 482–494, 2011.
- [20] Z. Xu, X. Luo, S. Zhang, X. Wei, L. Mei, and C. Hu, "Mining temporal explicit and implicit semantic relations between entities using web search engines," *Future Generation Computer Systems*, vol. 37, pp. 468–477, 2014.
- [21] Y. Liu, Q. Zhang, and L. Ni, "Opportunity-based topology control in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 3, pp. 405–416, 2010.
- [22] F. Hao, G. Min, J. Chen et al., "An optimized computational model for task-oriented multi-community-cloud social collaboration," *IEEE Transactions on Services Computing*, vol. 7, no. 3, pp. 346–358, 2014.
- [23] Y. Liu, L. Ni, and C. Hu, "A generalized probabilistic topology control for wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 9, pp. 1780–1788, 2012.
- [24] P. F. Fisher, "Models of uncertainty in spatial data," in *Geographical Information Systems*, vol. 1, pp. 191–205, 1999.
- [25] M. Schneider, "Uncertainty management for spatial data in databases: Fuzzy spatial data types," in *Advances in Spatial Databases*, vol. 1651 of *Lecture Notes in Computer Science*, pp. 330–351, Springer, Berlin, Germany, 1999.
- [26] L. Freire, A. Roche, and J.-F. Mangin, "What is the best similarity measure for motion correction in fMRI time series?" *IEEE Transactions on Medical Imaging*, vol. 21, no. 5, pp. 470–484, 2002.
- [27] T. Warren Liao, "Clustering of time series data?—a survey," *Pattern Recognition*, vol. 38, no. 11, pp. 1857–1874, 2005.
- [28] A. Kakas, R. Kowalski, and F. Toni, "The role of abduction in logic programming," in *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 5, pp. 235–324, Oxford University Press, Oxford, UK, 1998.
- [29] M. Gelfond and V. Lifschitz, "The stable model semantics for logic programming," in *Logic Programming: Proceedings of the Int'l Conference and Symposium*, R. Kowalski and K. Bowen, Eds., pp. 1070–1080, 1998.
- [30] L. M. Pereira and H. T. Anh, "Evolution prospection," in *New Advances in Intelligent Decision Technologies: Results of the First KES International Symposium IDT, Himeji, Japan, 2009*, K. Nakamatsu, Ed., vol. 199 of *Studies in Computational Intelligence*, pp. 51–63, Springer, 2009.
- [31] J. Neves, J. Machado, C. Analide, A. Abelha, and L. Brito, "The halt condition in genetic programming," in *Progress in Artificial Intelligence*, vol. 4874 of *Lecture Notes in Computer Science*, pp. 160–169, Springer, Berlin, Germany, 2007.
- [32] J. Neves, "A logic interpreter to handle time and negation in logic databases," in *Proceedings of the Annual Conference of the ACM on the Fifth Generation Challenge*, pp. 50–54, ACM, New York, NY, USA, 1984.
- [33] J. Kacprzyk, F. E. Petry, and A. Yazici, *Uncertainty Approaches for Spatial Data Modeling and Processing*, 2010.
- [34] D. Klimešová and E. Ocelíková, "Spatial data uncertainty management," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 209–212, 2008.
- [35] R. Mendes, J. Kennedy, and J. Neves, "The fully informed particle swarm: simpler, maybe better," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 204–210, 2004.
- [36] H. Vicente, J. C. Roseiro, J. M. Arteiro, J. Neves, and A. T. Caldeira, "Prediction of bioactive compound activity against wood contaminant fungi using artificial neural networks," *Canadian Journal of Forest Research*, vol. 43, no. 11, pp. 985–992, 2013.
- [37] A. T. Caldeira, M. R. Martins, M. J. Cabrita et al., "Aroma compounds prevision using artificial neural networks: influence of newly indigenous *Saccharomyces* spp. in white wine produced with *Vitis vinifera* cv Siria," in *Proceedings of the 6th International Conference on Simulation and Modelling in the Food and Bio-Industry (FOODSIM '10)*, pp. 33–40, June 2010.
- [38] D. Carneiro, P. Novais, F. Andrade, J. Zeleznikow, and J. Neves, "Using case-based reasoning and principled negotiation to provide decision support for dispute resolution," *Knowledge and Information Systems*, vol. 36, no. 3, pp. 789–826, 2013.
- [39] R. Kowalski, "The logical way to be artificially intelligent," in *Computational Logic in Multi-Agent Systems*, vol. 3900 of *Lecture Notes in Computer Science*, pp. 1–22, Springer, Berlin, Germany, 2006.
- [40] F. K. J. Sheridan, "A survey of techniques for inference under uncertainty," *Artificial Intelligence Review*, vol. 5, no. 1-2, pp. 89–119, 1991.
- [41] M. Ginsberg, *Readings in Nonmonotonic Reasoning*, 1980.
- [42] B. Rodrigues, S. Gomes, H. Vicente et al., "Systematic coronary risk evaluation through artificial neural networks based

- systems,” in *Proceedings of the 27th International Conference on Computer Applications in Industry and Engineering (CAINE '14)*, New Orleans, La, USA, October 2014.
- [43] S. Pereira, S. Gomes, H. Vicente et al., “Automatic diagnosis of attention deficit hyperactivity disorder through an artificial neuronal networks based system,” in *Proceedings of the IEEE International Conference on Imaging Systems and Techniques (IST '14)*, Santorini, Greece, October 2014.
- [44] V. Abelha, H. Vicente, J. Machado, and J. Neves, “An assessment on the length of hospital stay through artificial neural networks,” in *Proceedings of the 9th International Conference on Knowledge, Information and Creativity Support Systems (KICSS '14)*, Limassol, Cyprus, November 2014.
- [45] J. Y. Halpern, “Reasoning about knowledge: an overview,” in *Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowledge*, pp. 1-17, Morgan Kaufmann, 1986.
- [46] B. Kovalerchuk and G. Resconi, “Agent-based uncertainty logic network,” in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ '10)*, pp. 1-8, Barcelona, Spain, July 2010.
- [47] P. Lucas, “Quality checking of medical guidelines through logical abduction,” in *Research and Development in Intelligent Systems XX: Proceedings of AI2003, the Twenty-third SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pp. 309-321, Springer, London, UK, 2004.
- [48] A. Hommersom, P. J. Lucas, and P. van Bommel, “Checking the quality of clinical guidelines using automated reasoning tools,” *Theory and Practice of Logic Programming*, vol. 8, no. 5-6, pp. 611-641, 2008.
- [49] J. Machado, A. Abelha, P. Novais, and J. Neves, “Quality of service in healthcare units,” *International Journal of Computer Aided Engineering and Technology*, vol. 2, no. 4, pp. 436-449, 2010.
- [50] Y. Liu and M. Sun, “Fuzzy optimization BP neural network model for pavement performance assessment,” in *Proceedings of the IEEE International Conference on Grey Systems and Intelligent Services (GSIS '07)*, pp. 1031-1034, IEEE, November 2007.
- [51] W. Mayer-Gross, E. Slater, and M. Roth, *Clinical Psychiatry*, vol. 1, Baillière, Tindal & Cassel, London, UK, 1954.