Medical Diagnosis by Possibilistic Classification Reasoning

Mohammad Homam Alsun; Laurent Lecornu; and Basel Solaiman
Dept. ITI, Telecom Bretagne/ INSTITUT TELECOM, Brest, France
{mohammad.alsun,laurent.lecornu, basel.solaiman}@telecom-bretagne.eu

Clara Le Guillou; and Jean Michel Cauvin
Medical Information Departement, CHU Brest, Brest, France
{Clara.LeGuillou, jean-michel.cauvin}@chu-brest.fr

Laurent Lecornu; Clara Le Guillou;
Jean Michel Cauvin; and Basel Solaiman
Inserm, U650, IFR 148 ScInBioS, Brest, France

Abstract - In medicine, diagnostic reasoning refers to the approaches used by physicians with the aim of achieving a medical diagnosis concerning a given patient. This paper presents a new approach of medical decision support systems. The proposed approach is based on the use of possibility theory as a global framework, including knowledge representation (as a possibilistic pair of measures: Necessity, Possibility); and, building a possibilistic medical knowledge base (to be exploited in order to make a diagnostic decision (classification of new medical cases). The efficiency validation of the proposed approach is conducted using an Endoscopic Knowledge and Case Base systems. Obtained results confirm that the proposed approach constitutes an efficient tool in terms of medical knowledge representation and possibilistic diagnostic reasoning.

Keywords: Possibilistic reasoning, Diagnostic reasoning by classification, Decision support system, Possibilistic representation.

1 Introduction

Diagnostic reasoning is the general term referring to different approaches of diagnostic hypothesis generation (based on the availab[1]le heterogeneous data) with the aim to achieve a diagnosis tasks or to explain a set of observed abnormal manifestations (findings or symptoms). It is obvious that the physician is the direct responsible of health and life of his patients. For this reason, diagnosis delivering is an extremely important, although difficult, task. Medical Diagnosis Decision Support Systems (such as Classification Based Reasoning Systems, Case Based Reasoning Systems, Machine Learning Systems and Medical Data Mining Systems) aim to reduce diagnosis error risks and to help physicians making high quality and reliable medical decisions [1-2].

In this paper, the Medical Diagnosis using Classification Reasoning approach is considered. This reasoning approach is based on the comparison of the available information acquired from a given patient and the medical a prior knowledge formulated by the physician (i.e. Expert Medical Vision) with the aim to assign possible diagnoses facing this particular case. The medical a prior knowledge concerns the features to be observed, as well as the description of all potential diagnoses (or diseases) to be considered. This knowledge is assumed to be available in a medical knowledge base.

In fact, whatever the nature of the medical knowledge, diagnosis is described (by physicians) using all features potential modalities, by observing the relationship (Feature Modality) – (Diagnosis), that refers to the compatibility degree between a given feature and a particular diagnosis. This relation expression way constitutes an information that can suffer from one of the different imperfection types (ambiguity, imprecision, uncertainty, etc.) [3-4].

Therefore, the mathematical model chosen, in order to represent this information (relationship), has to be able to manage different types of information imperfections, as well as that the reasoning mechanism selected must be adapted with the considered representation model.

In this paper, the medical knowledge is modeled in a possibilistic framework, in which the description of the relationship (established by the physician) will be represented as a possibilistic pair of measures (Necessity, Possibility). The possibility measure refers to which degree the medical knowledge is possible, and the necessity measure refers to which degree this knowledge is certain [5-8]. The proposed model focuses on building the possibilistic knowledge base that will be exploited to perform the classification reasoning of a new medical case. In fact, the use of the possibility theory as a global framework is motivated by its capacity to handle different

[1] This work was partially supported by a TECSAN/ANR project under the name Med1dex associating Brest CHU, Telecom Bretagne, INSERM Unit U650 LaTIM and PRISMEDICA Company.

types of information imperfections, as well as the diagnostic reasoning made thanks to the two possibilistic measures (Necessity and Possibility).

The remainder of this paper is organized as follows: main aspects of the possibility theory are briefly introduced in section 2. Section 3 presents a knowledge model representation allowing physicians to express their medical knowledge. Section 4 is devoted to the detailed description of the proposed possibilistic system. The particular Endoscopic application allowing the efficiency validation of the proposed approach is presented in section 5. Section 6 concerns the presentation and analyze of the obtained results. Finally, section 7 presents some conclusions related to the proposed approach, as well as some propositions for future work.

2 Possibility theory

Possibility theory (introduced by L. Zadeh in 1978 [9] and then developed by Dubois and Prade in 1988 [10]) offers an interesting tool allowing to deal with different forms of information imperfections (ambiguity, imprecision, incompleteness, etc.).

Let $\Omega = \{x_1, ..., x_i, ..., x_N\}$ denotes an exhaustive and exclusive Universe of discourse. At the semantic level, the basic function in possibility theory is a *possibility distribution* denoted as $\pi : \Omega \to [0,1]$, which assigns to each possible singleton x_i from Ω a value in $[0,1]$. This possibility distribution represents the occurrence possibility degree of x_i. If, for some x_i, $\pi(x_i) = 1$, then x_i is said to be a "totally possible singleton"; and if $\pi(x_i) = 0$, then x_i is said to be an "impossible singleton".

Based on the possibility distribution, the information concerning the occurrence of an event $A \in \mathcal{P}(\Omega)$ (where $\mathcal{P}(\Omega)$ is the power set of Ω) is represented through two set functions: a Possibility Measure (denoted as $\Pi(.)$ and a Necessity Measure (denoted as $\mathrm{N}(.)$).

The *possibility measure* $\Pi(.)$ is defined as follows:

$$\Pi : \mathcal{P}(\Omega) \to [0,1]$$
$$A \to \Pi(A) = \max_{x \in A} \pi(x) \tag{1}$$

If the possibility measure of an event $A \in \mathcal{P}(\Omega)$ is equal to unity (i.e. $\Pi(A) = 1$), then A is said to be totally possible. If $\Pi(A) = 0$, then A is said to be totally impossible. The second measure, called the *necessity measure* $\mathrm{N}(.)$, is defined as follows:

$$\mathrm{N} : \mathcal{P}(\Omega) \to [0,1]$$
$$A \to \mathrm{N}(A) = 1 - \max_{x \notin A} (1 - \pi(x)) \tag{2}$$

If the necessity measure of an event $A \in \mathcal{P}(\Omega)$ is unity (i.e. $\mathrm{N}(A) = 1$), then A is totally certain. If $\mathrm{N}(A) = 0$, then A is totally uncertain.

In the case where A is a fuzzy set (with the membership function μ_A) defined over Ω, both possibility and necessity measures can be given by the next general formulas:

$$\Pi(A) = \max_{x \in \Omega} [\min(\pi(x), \mu_A(x))] \tag{3}$$

$$\mathrm{N}(A) = 1 - \Pi\left(\bar{A}\right) = \min_{x \in \Omega} [\max(\pi(x), 1 - \mu_A(x))] \tag{4}$$

Notice that the definitions given in (3) and (4) can be considered as the most general due to the inclusion of the crisp definition of A (where the membership function will be considered as identical to the caracteristic function defining the crisp set A).

3 Medical knowledge representation

This section is devoted to the presentation of how physicians describe different diagnoses (findings) using "predefined features" of qualitative nature (derived from quantitative, binary … observations or facts.).

In fact, a pathology is described by physicians using all features potential modalities, through "frequency-based linguistic variables" representing the physician vision, *vs*. statistical experience, while observing the relationship: (Feature Modality) – (Diagnosis). Some of these linguistic variables are: a feature modality is ("*frequently*", "*rarely*", "*never*", etc.) observed when dealing with a given diagnosis.

In the following subsections, the mathematical models representing the physician description will be detailed concerning the medical knowledge and the case bases.

3.1 Medical knowledge base

The Medical knowledge base is assumed to encapsulate the Expert Physician knowledge related to different considered diagnoses.

Let $C = \{C_1, ..., C_m, ..., C_M\}$ denotes the set of M diagnoses. A diagnosis C_m is characterized using a set of G features $P = \{P_1, ..., P_g, ..., P_G\}$, and each feature P_g can take K_g potential modalities defined by the set $V_g = \left\{v_1^g, ..., v_k^g, ..., v_{K_g}^g\right\}$.

The diagnosis $C_m, m = 1, ..., M$ is presented in the medical knowledge base by the following model:

$$C_m = \left\{\left(P_i, v_j^i, D\left(v_j^i, C_m\right)\right); i = 1, ..., G; j = 1, ..., K_i\right\} \tag{5}$$

where P_i denotes the feature "i", v_j^i: the j^{th} modality ($j = 1, ..., K_i$) of the feature "i", and $D\left(v_j^i, C_m\right)$ represents the frequency description related to the diagnosis C_m (defined by the expert) of the j^{th} modality of the feature "i".

Usually, the frequency based description $D\left(v_j^i, C_m\right)$ is expressed by a probabilistic nature based linguistic

variable, referring to the experts "assessment" of the frequency of occurrence of v_j^i related to the diagnosis C_m.
Indeed, the ideal description approach (from a mathematical point of view), is to attribute to each couple: (Feature Modality) – (Diagnosis), its exact occurrence probability value. Nevertheless, these values are rarely known by physicians in terms of exact and precise values. For this reason, and in order to express this "imperfect" knowledge of the probabilistic values, physicians use a qualitative description by means of natural linguistic variables [11]. The most popular approach in this context is the Qualitative Theory of Uncertainty formalizing a multi-valued logic with linguistic truth degrees. A qualitative scale running from *"Impossible"* to *"Certain"*, allows representing the uncertainty (for details see [12]). In short, this theory offers the expert the opportunity to express his uncertainty by using linguistic values, more indicative than numerical ones used in possibility or probability theories.
In the same spirit, [13] defined two levels of relations; the relation Feature/Diagnosis, and the relation Modality/Diagnosis.
The relation Feature/Diagnosis can be described using three potential statutes: *"Impossible"*, *"Without Interest"* or *"With Interest"*. For a relation Feature/Diagnosis given as *"Impossible"* (resp. *"Without Interest"*), all relations Modality/Diagnosis concerning this feature, will be automatically described as *"Impossible"* (resp. *"Without Interest"*).
For a relation Feature/Diagnosis given as *"With Interest"*, the corresponding relations Modality/Diagnosis will be described using a qualitative description by choosing one of L values ranging from *"Never"* to *"Essential"* as follows (Figure.1): $D = \{D_1, ..., D_l, ..., D_L\}$; where D_1 = "*Never*",…, D_L = "*Essential*".

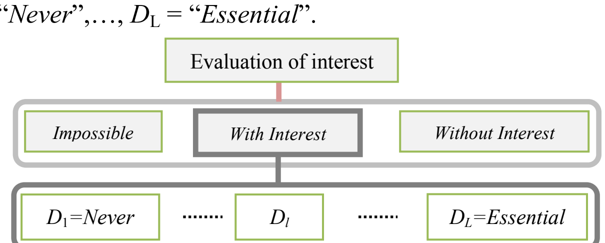

Figure 1. Synthesis of different relations

Table.1 shows an example of an expert description in a medical knowledge base.

Table 1. ***Example of a physician description in Medical Knowledge Base***

	P_1			P_2	
	v_1^1	v_2^1	v_3^1	v_1^2	v_2^2
C_1	*Essential*	*Never*	*Never*	*Without*	*Without*
C_2	*Rare*	*Usual*	*Usual*	*Never*	*Essential*
C_3	*Except*	*Usual*	*Usual*	*Impossible*	*Impossible*

In this example, the physician describes a set of three diagnoses (diseases) $C = \{C_1, C_2, C_3\}$ using two features: P_1 (with three modalities: $V_1 = \{v_1^1, v_2^1, v_3^1\}$) and P_2 (with two modalities: $V_2 = \{v_1^2, v_2^2\}$).
As noted, the feature P_2 is evaluated by the physician as *"Without Interest"* (resp. *"Impossible"*) concerning the first diagnosis C_1 (resp. C_4). Therefore, all the relations (Modality(v_j^2)/Diagnosis(C_1)) concerning this feature and the diagnosis C_1 are automatically described as *"Without Interest"* (resp. *"Impossible"* (concerning C_4)). In this example, 5 linguistic variables are used:
D = {D_1 = *Never*, D_2 = *Exceptional*, D_3 = *Rare*, D_4 = *Usual*, D_5 = *Essential*}.

3.2 Medical Case Base

A Medical Case Base contains a set of N medical cases $B = \{B_1, ..., B_n, ..., B_N\}$, where each case B_n is characterized using the same set of features $P = \{P_1, ..., P_g, ..., P_G\}$ used in the construction of the Medical Knowledge Base. Each feature P_g can take one and only one of its potential modalities included in the corresponding set of potential modalities V_g. A case $B_n, n = 1, ..., N$ is thus represented in the medical case base by the following medical model:

$$B_n = \left\{\left(P_i, x^{i,n}\right), d_n; i = 1, ..., G^{'}; x^{i,n} \in V_i; G^{'} \leq G\right\} \quad (6)$$

where $G^{'}$ is the number of observed features in the case B_n, d_n is the diagnosis associated with the case B_n.

An illustrative example of an Endoscopic Medical Case Base is shown in Table 2.

Table 2. *Example of medical description in an Endoscopic Case Base*

	P_1 = Object Type	P_2 = Origin	*d*
B_1	*Not Homogenous Simple*	*Parietal*	*Tumor*
B_2	*Homogenous*	*Parietal*	*Spot*
B_3	*Not Homogenous Multiple*	*Luminal*	*Food*

In this example, three cases are shown using two features: P_1 = "Object Type" with three modalities {*Homogenous, Not Homogenous Simple, Not Homogenous Multiple*} and P_2 = "Origin" with two modalities {*Parietal, Luminal*}.

4 Classification reasoning based on possibilistic reasoning

In the light of the realization of a classification reasoning based on possibilistic reasoning, two main issues must be performed: the construction of the possibilistic knowledge base and, the development of the corresponding possibilistic reasoning mechanism. In the next subsections, the following issues will be addressed:

- What is the possibilistic model to be used in the knowledge base building?
- How to obtain the diagnostic decision using the two possibilistic quantities (possibility and necessity measures)?
- How to represent graphically the obtained decisions?

4.1 Possibilistic knowledge base building

In order to build a possibilistic knowledge base, the linguistic description formulated by the physicians (representing the two levels of relations; the Feature/Diagnosis relation, and the Modality/Diagnosis relation) should be transformed into a possibilistic description.

Each of the linguistic values is represented by a pair of two values; the Necessity and Possibility. $\forall l \in [1,...,L]; D_l \equiv (\mathrm{N}_l, \Pi_l); (\mathrm{N}_i, \Pi_i) \in [0,1]\times[0,1]$. So, a possibilistic numeric scale corresponding to the qualitative scale is established.

In other words, the possibilistic correspondence is established between different qualitative physician descriptions and the pairs of Necessity, Possibility measures (called possibilistic pair). In fact, the values of possibility and necessity measures are not chosen arbitrarily, but based on expert guidance. For instance, the relation Feature/Diagnosis given as *"impossible"* is represented by the possibilistic pair $(\mathrm{N},\Pi)=(0,0)$; and the relation Modality/Diagnosis given as *"Usual"* by the possibilistic pair $(\mathrm{N},\Pi)=(0.75,1)$, etc. The values have been chosen respecting the ranking of the qualitative scale. Therefore, each diagnosis C_m is be represented in the Possibilistic Medical Knowledge Base by the following model:

$$C_m = \left\{\left(P_i, v_j^i, \left(\mathrm{N}\left(v_j^i, C_m\right), \Pi\left(v_j^i, C_m\right)\right)\right)\right\} \quad (7)$$

where $i=1,...,G$ and $j=1,...,K_i$.

4.2 Possibilistic reasoning

4.2.1 Case-diagnosis similarity estimation

Let us suppose that we are faced with a new case B_{new} described by $G^{'}$ observed features. Each feature P_i is associated with the modality $x^{i,new}$ ($x^{i,new}$ is one of the potential modalities within the set V_i). Therefore, the new case will be represented as follows:

$$B_{new} = \left\{\left(P_i, x^{i,new}\right); i=1,...,G^{'}, G^{'} \le G, x^{i,new} \in V_i\right\} \quad (8)$$

The objective of the diagnostic decision support system is to compare the new case with all diagnoses in the medical knowledge base, and thus, to "estimate" the potential diagnosis (diagnoses). This comparison may be conducted by distance estimation, similarity estimation, etc. In the proposed approach, the case-diagnosis similarity will be considered and estimated by means of two measures:

- The possibility measure of case- diagnosis similarity that tells us the possibility level that the considered case is similar to the stored diagnosis knowledge; and
- The necessity measure of case-diagnosis similarity that defines the certainty of this similarity.

Let us consider the diagnosis C_m represented as already shown in the equation (7). To compute the possibility and necessity measures of similarity likeness between the case B_{new} and C_m, the following steps are conducted:

1. For each element $\left(P_i, x^{i,new}\right)$ belonging to B_{new}, the corresponding element $\left(P_i, v_j^i = x^{i,new}, \left(\mathrm{N}\left(v_j^i, C_m\right), \Pi\left(v_j^i, C_m\right)\right)\right)$ from the description of diagnosis C_m is extracted from the Medical Knowledge Base. Finally, the result of this extracted is given as:

$$B_{new} = \left\{\left(P_i, x^{i,new}, \left(\mathrm{N}\left(x^{i,new}, C_m\right), \Pi\left(x^{i,new}, C_m\right)\right)\right)\right\} \quad (9)$$

   where $i=1,...,G^{'}; G^{'} \le G; x^{i,new} \in V_i$.

   This operation is repeated for each diagnosis.

2. The total similarity between the considered case and the stored diagnosis knowledge is then computed by means of combination of these (necessity, possibility) pairs, using a combination operator. In the literature, several combination operators are presented. We can define the following combination operators: *average value*, *minimum*, *maximum, geometric mean,* respectively given in the following equations,

$$Sim\left(B_{new}, C_m\right) = \frac{1}{G^{'}} \sum_{i=1}^{G^{'}} Sim\left(x^{i,new}, C_m\right) \quad (10)$$

$$Sim\left(B_{new}, C_m\right) = \min_{i=1}^{G^{'}} \left(Sim\left(x^{i,new}, C_m\right)\right) \quad (11)$$

$$Sim\left(B_{new}, C_m\right) = \max_{i=1}^{G^{'}} \left(Sim\left(x^{i,new}, C_m\right)\right) \quad (12)$$

$$Sim\left(B_{new}, C_m\right) = \sqrt[G^{'}]{\prod_{i=1}^{G^{'}} Sim\left(x^{i,new}, C_m\right)} \quad (13)$$

where Sim can refer to the necessity or possibility measure. The combination operator selection is generally related to the desired application purpose.

4.2.2 Uncertainty measurement and representation

In order to measure and represent the uncertainty associated with the potential decision of considering the case B_{new} as being labelled with the diagnosis C_m, a confidence index Ind [14] is defined as a combination of the possibility and necessity of similarity measures as follows:

$$Ind_{B_{new},C_m} = \Pi\left(B_{new},C_m\right) + \Pi\left(B_{new},C_m\right) - 1 \qquad (14)$$

Since the possibility and the necessity measures range in the interval $[0,+1]$, the value of the confidence index, ranges in the interval $[-1,+1]$. $Ind_{B_{new},C_m} = -1$, is obtained when $\Pi\left(B_{new},C_m\right) = \mathrm{N}\left(B_{new},C_m\right) = 0$ meaning that the diagnosis C_m is impossible to be a potential solution for the case B_{new}. On the other side, $Ind_{B_{new},C_m} = +1$, means that the similarity is totally possible and totally certain ($\Pi\left(B_{new},C_m\right) = \mathrm{N}\left(B_{new},C_m\right) = 1$). In this case, the diagnosis C_m represents the optimal solution for B_{new} (Figure 2). In addition to interest of the confidence index in measuring the uncertainty level in terms of similarity between the case B_{new} and a considered diagnosis, it can be represented graphically as shown in Figure 2.

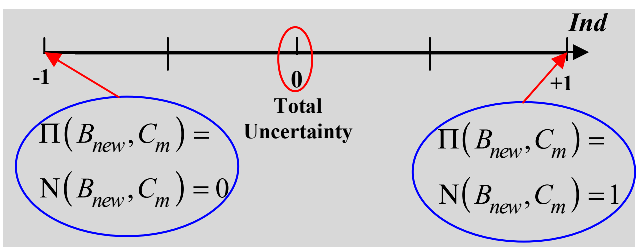

Figure 2. The confidence index *Ind.*

5 Medical application

The Medical Knowledge Base used for the evaluation of the proposed approach is an Endoscopic Knowledge Base [15-16]. This Base consists of a set of 89 endoscopic findings (diagnoses). The diagnosis is described using a set of 33 features corresponding to 206 modalities. The linguistic variables which are used to describe the relation Feature/Diagnosis are {*Impossible, Without Interest, With Interest*} and the qualitative scale used to express the relation Modality/Diagnosis is : {*Never, Exceptional 2, Exceptional 1, Rare 2, Rare 1, Usual 2, Usual 1, Essential*}. Furthermore, the linguistic variable *Doubtful* that is an intermediate between *Never* and *Exceptional*, is added when the expert has an ignorance about the modality reality. We note that there are two importance levels for the three variables *Exceptional, Rare, Usual*.

The case base used in this study comes from a system of endoscopic image analyzes [16-17]. It is an assistance system for the decision-making of the diagnosis of endoscopic findings. These findings are described by the physicians from the endoscopic images through symbolic terms, which are defined by the Minimal Standard Terminology of the SEGE (European Company of Gastro-enterology). The case (or object) in the base represents a description of the image (using a set of 33 features, 24 features to describe an object and 9 features to describe a potential sub-object) of an endoscopic lesion.

In addition to the fact that the sub-object features depend on the “non-homogenous state” of the feature’s type, there are some other relationships between modalities and feature (for example a diagnosis whose Density is *essentially* “unique” has no Spatial Organization feature, a diagnosis whose Shape is *essentially* “ring-tube” has no Minor Axis feature, …) or between modalities of different features (for example, modalities of Relief and Thickness features or modalities of object sizes and sub-object sizes, etc.).

6 Experiments and results

First, each linguistic variable (mentioned in Subsection 3.1) used by physicians to describe both relations Feature/Diagnosis and Modality/Diagnosis, is transformed to an adequate possibilistic pair. Table 3 and Table 4 show all these linguistic variables with the corresponding pair of necessity and possibility.

Table 3. *Linguistic variable and its corresponding possibilistic pair for the Feature/Diagnosis relation*

Feature/Diagnosis Relation	
Linguistic variable	(N , Π)
Impossible	(0, 0)
With interest	See Table 4
Without interest	(0, 1)

Table 4. *Linguistic variable and its corresponding possibilistic pair for the Modality/Diagnosis relation*

Modality/Diagnosis Relation	
Linguistic variable	(N, Π)
Essential	(1, 1)
Usual.1	(0.80, 1)
Usual.2	(0.75, 1)
Rare.1	(0.30, 1)
Rare.2	(0.25, 1)
Exceptional.1	(0.10, 1)
Exceptional.2	(0.05, 1)
Doubtful	(0.05, 0.5)
Never	(0, 0)

Second, the possibilistic knowledge base is constructed by replacing the linguistic description by the corresponding possibilistic pair.

Once the possibilistic knowledge base is computed, and in order to have a simple and clear representation of our results, a small subset CB of cases belonging to different diagnose will illustrate this representation.

Suppose that $CB = \{B_1, B_2, B_3\}$ is a case base consisting of three cases (endoscopic lesions) belonging to the following diagnoses respectively {Esophagus Normal, Dilated Lumen, Ring}.

In order to compute the similarity between each case and the 89 diagnoses included in the possibilistic knowledge database, possibility and necessity of case-diagnosis similarity are computed by the proposed approach (presented in Subsection 4.2.1). In fact, at this level the possibilistic reasoning upon the possibility measure intends to exclude some diagnosis. Therefore, if there is at least one of the local possibility measures, equals to zero, then the corresponding diagnosis is rejected as an impossible solution. Conversely, if there is no local possibility measure that equals to zero, then the corresponding diagnosis is considered as a potential solution. As the influence of all local possibility and necessity values will be taken into account, and for this reason, the *geometric mean* presented in the equation (13) verifying these conditions, is considered as the combination operator of the local possibility and necessity measures. These potential solutions (having a total possibility measure unequal to zero) will be arranged using the value of confidence index combining the two measures: possibility and necessity.

Table 5 shows the obtained results including the potential diagnoses for each case in this small base as well as their possibilistic pairs and confidence index. A graphic representation which helps us to easily interpret the obtained results is presented (as an example for the cases B_1 and B_3) in Figure 3 respectively (a) and (b).

Table 5. ***Potential solutions for the cases included in the CB with its confidence index***

Case	Case diagnosis	Potential diagnoses (Solutions)			
		Diagnosis	[N, Π]	*Ind*	*Each one of the rest of the 89 diagnoses has the pair* [N, Π]= [0, 0], and thus the confidence index : *Ind* = -1
B_1	Esophagus Normal	Esophagus Normal	[0.86, 1]	0.86	
		Spot Esophagus	[0.44, 1]	0.44	
		Previous Surgery	[0.36,0.81]	0.17	
B_2	Dilated Lumen	Dilated Lumen	[0.67, 0.97]	0.64	
B_3	Ring	Ring	[0.69,1]	0.69	
		Web	[0.34,0.88]	0.22	

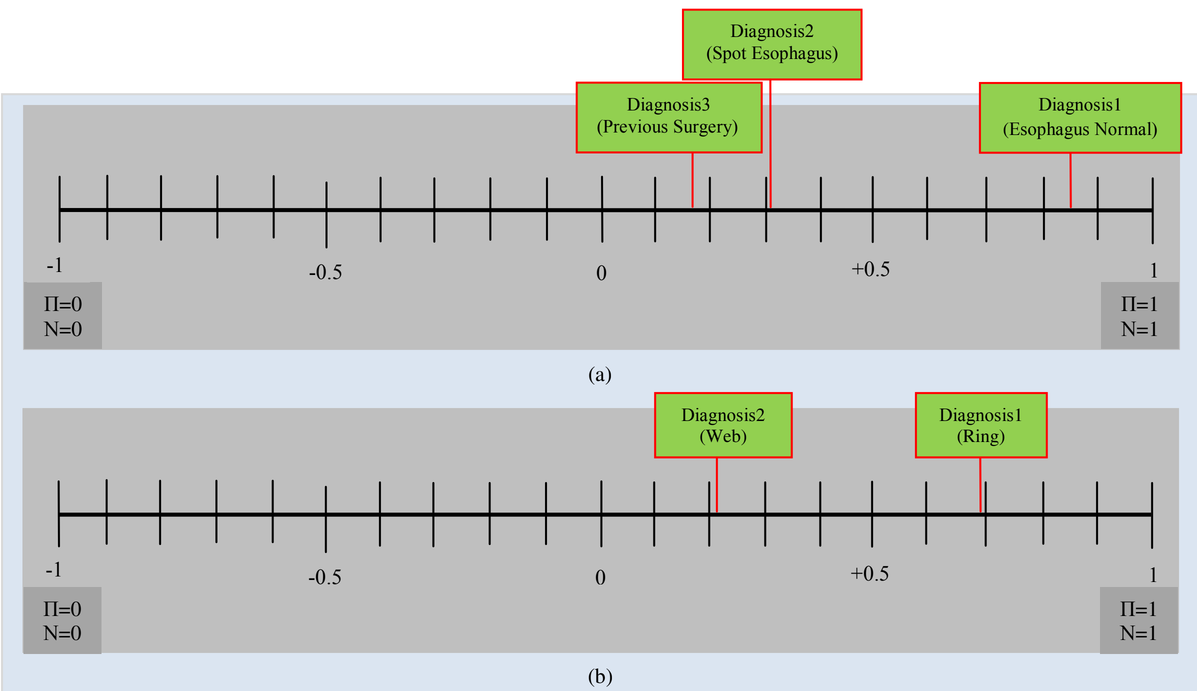

Figure 3. Graphic representation of potential diagnoses of B_1 and B_3 using the confidence index.

Obtained results are extremely interesting because the potential diagnoses are restricted within a small set containing two or three diagnoses. In this example, the correct diagnosis of any case is always one of potential solutions (diagnoses), and often this is the solution having the greatest similarity and the greatest confidence index (for example B_1). Frequently, this is the only solution (for example B_2). The reality confirmation of this analyze is obviously shown in Figure 3 (for example, the diagnosis "Esophagus Normal" that represents the right diagnosis of the case B_1 is the potential solution the closest to the optimal solution (confidence index is equal to the +1).

After validating the proposed approach on several examples, a classification process of a database containing

4450 cases (lesions) has been applied; promising results presented in Table 6 are obtained. The classification rate performed is equal to 97.55%, involving 61.42% of cases in which the correct diagnosis is classified as a unique potential solution, 36.31% in which the right diagnosis is as obtained as the first proposed solution, and the rest 2.45% of cases, the correct diagnosis of course, is a potential solution, but in another order. This result is coherent with other methods as [17].

Table 6. *Obtained results for all the base case*

	Total	Not found	Sole	First	Other
Cases number	4450	0	2725	1616	109
Percentage		100%	61.24%	36.31%	2.45%

7 Conclusion

This paper reported a possibilistic diagnostic reasoning classification system. The proposed system is based on the use of possibility theory (possibility and necessity measures) in order to, in the first step, build a medical knowledge base, and in the second step, to perform a medical diagnostic reasoning by exploiting this possibilistic knowledge base. The choice of this form of representation and reasoning is motivated by three mains issues: the capacity of the possibility theory to deal with all information types; as well as, with all types of information imperfections; the decision making based on two quantities (possibility and necessity measures). Promising results are obtained using the proposed system.
As a future work, the proposed model will be exploited to study the relations between features by comparing their local possibility and necessity values in order to distinguish the key features of each diagnosis (disease). Another perspective involves the exploitation of the proposed model in order to construct a Case Based Reasoning-based medical diagnostic reasoning system.

References

[1] K. Aftarczuk, A. Kozierkiewicz et al., *Using Representation Choice Methods for a Medical Diagnosis Problem*, Springer, computer science, vol. 4253, pp. 805-812. 2006.

[2] T. Wetter, *Medical Decision Support Systems*, Springer, computer science, vol. 1933, pp. 458-466. 2000.

[3] P. Szolovits. *Uncertainty and decision in medical informatics*, methods of information in medicine, vol. 34, pp. 111-121, 1995.

[4] DK. Owens, HC. Sox, *Medical Decision Making: Probabilistic Medical Reasoning*, Addison Wesley USA, chapter 3, 1990.

[5] B. Bouchon-Meunier, *La logique floue et ses applications*, Addison Wesley France, chapter 2, 1995.

[6] D. Dubois, H. Fargier et al., *Possibility theory in constraint satisfaction problems: handling priority, preference and uncertainty*, Applied Intelligence, vol. 6, pp. 287-309, 1996.

[7] D. Dubois, *et al.*, *Fuzzy set-based models in case-based reasoning*, 2nd International Conference on Case-Based Reasoning, ICCBR'97, USA, 1997.

[8] D. Dubois, *et al.*, *Probability-Possibility transformations, triangular fuzzy sets and probabilistic inequalities*. Reliable Computing, Vol. 10: pp. 273–297, 2004.

[9] A. Zadeh, *Fuzzy Sets as the Basis for a Theory of Possibility*, Elsevier, Fuzzy Sets and Systems, Vol. 100, Supplement. 1, pp. 9-34, 1999.

[10] D. Dubois and H. Prade, *Possibility Theory: an Approach to Computerized Processing of Uncertainty*, Plenum press, International Journal of general Systems, Vol. 15, Issue. 2, pp. 168-170, 1989.

[11] C. Leguillou, J.M. Cauvin *et al.*, *Knowledge representation and Cases Indexing in Upper Digestive Endoscopy*, IEEE Conferences, Vol. 1, pp. 9-12, 2000.

[12] H. Akdag, M. Deglas et al., *Qualitative theory of uncertainty*, IOS Press, Vol. 17, Issue. 4, pp. 333-362, 1992.

[13] JM. Cauvin, C. Le Guillou *et al.*, *Computer-Assisted Diagnosis System in Digestive Endoscopy*, IEEE Transactions on Information Technology in Biomedicine, Vol. 7, Issue. 4, pp. 256-262, 2003.

[14] Sh. Kikuchi, V. Perincherry, *Handling Uncertainty in Large Scale Systems with Certainty and Integrity,* MIT Engineering Systems Symposium, Cambridge, 2004.

[15] C. Le Guillou, JM. Cauvin, *From Endoscopic Imaging and Knowledge to Semantic Formal Images*, Springer, computer science, vol. 4370, pp. 189-201. 2007.

[16] C. Le Guillou, JM Cauvin et al., *Upper Digestive Endoscopic Scene Analyze.* 23rd Conference of the IEEE Engineering in Medicine and Biology Society, vol. 4, pp. 3855-3858, 2001.

[17] A. Zemerline, L. Lecornu, B. Solaiman., *Rule-Based Diagnostic System Fusion*, 10th International Conference on Information Fusion, 2007.