Structural verification through similarity measures for fuzzy rule bases representing clinical guidelines

Massimo Esposito¹, Domenico Maisto¹

Abstract —Clinical practice guidelines are expected to promote more consistent, effective, and efficient medical practices and improve health outcomes, especially if provided in the form of clinical decision support. However, most clinical guidelines, especially when expressed in the form of condition-action recommendations, embody different kinds of structural errors that compromise their practical value. With this respect, this paper presents a novel method for verifying the reliability of condition-action clinical recommendations encoded in the form of fuzzy rules, with the final aim of determining inconsistency, redundancy and incompleteness anomalies in a very simple and understandable fashion. The method is based on general definitions of inconsistency, redundancy and incompleteness for fuzzy clinical rules in terms of similarity between antecedents and consequents, bringing them near the imprecise character of fuzzy decision support systems. A key issue relies on the formalization of fuzzy degrees for these anomalies that can be simply interpreted by the final users as measurements suggesting the modifications to be performed to the clinical rules in order to eliminate or mitigate the existing undesired effects. The method has been profitably assessed on two sample sets of clinical rules: the first one identified from the relevant clinical literature and the second one extracted automatically by machine learning techniques from a widely known clinical database. The achieved results prove simplicity and usability of our method in detecting structural anomalies and in adjusting a rule base by exploiting information carried out during the verification phase.

I. INTRODUCTION

The increase of expert knowledge is characterizing medical domain determining, as a consequence, an increase of the specialization phenomenon. Specialization, in its turn, produces a constantly growing number of clinical guidelines.

Clinical practice guidelines are standardized specifications for care, developed by a formal process that incorporates the best scientific evidence of effectiveness with experts’ opinions [15]. They are expected to promote more consistent, effective, and efficient medical practices and improve health outcomes when followed [37], especially if provided in the form of clinical decision support [28] by means of a computer-based system designed to help people to make clinical decisions [32][9].

In order to build Decision Support Systems (DSSs) from clinical practice guideline, this latter is encoded into a logical formalism starting from natural language and embedded into a decision support implementation used in practice. Each clinical guideline may contain both a narrative section with complex control flow and a set of condition-action clinical rules.

Condition-action clinical rules represent elementary, isolated care recommendations, which specify one or at most a few conditions which are linked to specific actions [30], whereas narrative sections describe a coherent and unified care process in terms of a series of branching or iterative decisions unfolding over the time [26]. Clinical condition-action rules, which cover most diagnostic and therapeutic guideline recommendations, can be distilled from narrative guidelines [30], although this discards the control flow structure.

Focusing on condition-action clinical recommendations rather than time-oriented guidelines, up to now, decision support implementations have been widely studied [23], [1], [2], especially to find much concrete solutions for the process of rule formalization.

However, actually, one prerequisite for their broad acceptance and efficient application to medical settings is their power to guarantee a high level of quality and reliability [7]. Moreover, as shown in numerous studies, most condition-action clinical recommendations embody different kinds of semantic incongruences — commonly denoted as structural anomalies — that compromise their practical value [30], [31].

As a consequence, a detection of structural anomalies is crucial for condition-action clinical recommendations. This process, known as verification, is an extremely complicated task since, on the one hand, it can imply important consequences that could even have an effect on patient’s life and, on the other hand, it involves the treatment of medical knowledge that, by its nature, usually implies a great amount of subjectivity and imprecision.

The situation is further complicated by the fact that the process of correcting such anomalies might not be definitive but periodically repeated, since clinical rules might be subject to change due to medical progress in the treatment of individual diseases.

Consequently, the development process of a clinical rule often relies on expert panels for combining scientific evidence and expert opinion. Hence, conflicting clinical rules appear to be unavoidable, even for groups looking at the same scientific evidence and using defensible expert judgment procedures [11], and competing clinical rules addressing the same medical problem can confuse clinicians and undermine confidence in the guideline concept [3].

Besides, both the development of a DSS from expert knowledge and its updating are processes involving some
potential knowledge gaps caused by medical knowledge transfer from the expert into the rule base through a computable format established by the knowledge engineer [30].

A potential straight consequence of this situation is that the condition-action rule set encoding clinical recommendations could be not complete, i.e., no rule could be verified for some specific clinical observations and the related DSS could have no corresponding response.

This situation, also called as “missing rule” phenomenon, denotes inability of the considered rule base to formalize expert knowledge in a satisfying way.

In contrast to the intensive efforts made to develop DSSs based on condition-action clinical recommendations, the issue of providing mechanisms to perform their verification in order to detect structural anomalies and to ensure the overall consistency has been widely neglected thus far.

Taking into account such considerations, this paper presents a novel method for verifying the reliability of condition-action clinical recommendations encoded in the form of fuzzy rules, with the final aim of determining inconsistency, redundancy and incompleteness anomalies in a very simple and understandable fashion. The method is based on general definitions of inconsistency, redundancy and incompleteness for fuzzy clinical rules in terms of similarity between antecedents and consequents.

A key issue relies on the formalization of fuzzy degrees for these anomalies that can be simply interpreted by the final users as measurements suggesting the modifications to be performed to the clinical rules in order to eliminate or mitigate the existing undesired effects.

The method has been profitably assessed on both a sample set of clinical rules that has been identified from the relevant clinical literature and a set of rules automatically extracted from clinical database through machine learning techniques.

The remainder of the paper is structured as follows. In Section II, both the motivations for adopting the fuzzy logic formalism to encode clinical rules and an overview of existing methods for verifying fuzzy rules are reported. In Section III, the proposed method for verifying clinical rules encoded by exploiting fuzzy logic formalism is presented. Two clinical scenarios for validating the method are diffusely described in Section IV. A discussion concludes the paper in Section V.

II. MOTIVATION AND RELATED WORK

In the last few years, the computer science community has been challenged to support the transition to evidence-based medicine by helping to create the technologies necessary to make clinical knowledge more accessible, manageable and updatable [6].

In more detail, from an information management perspective, a condition-action clinical rule must be encoded into a decision support implementation. With this respect, the most critical activity is the selection of a mediating representation in a computable format [30].

Two specific issues affecting this activity are: 1) the inherent vagueness of the natural language used in clinical rules; 2) the uncertainty that arises when patients with highly similar clinical characteristics receive very different recommendations from the guideline [17], [13]. In other words, medical data measured or collected about a patient can be characterized according to fuzziness [24].

Up to now, it has been argued that the best manifestation of evidence-based medicine is to remove fuzziness from clinical decisions. As an instance, the clinical rules adopted for the classification of hypertension by the 2003 ESH/ESC guidelines, drawn up by the European Society of Hypertension (ESH) and the European Society of Cardiology (ESC), are still based on cutoff values, i.e., crisp values used as thresholds to discriminate between “normal” and “abnormal” measures of a considered feature. The reason consists in reckoning that the use of cutoff values simplifies diagnostic and treatment approaches in daily practice [18].

Nevertheless, modeling the intrinsic fuzziness in clinical rules and trying to manage it could be more appropriate than forcing healthcare providers to work with sharp, but strict, crisp logic formulations [36].

With this respect, in this work, Fuzzy Logic [40], [41] has been identified as the most suitable approach to describe vagueness and imprecision, since it enables to explicitly represent natural vagueness rather than abolishing it, by means of a precise mathematical language. The most notable advantage in the adoption of fuzzy logic for encoding condition-action clinical rules is that they can be written in a form similar to natural language and provide recommendations consistent with the human thinking. A consequence of this transparency is that fuzzy rules are easy to develop and understand.

Representing condition-action clinical recommendations in the form of fuzzy rules implies the adoption of verification methods specifically devised and developed to work on fuzzy decision support systems. Until now, different approaches have been adopted to study the problem of verifying a fuzzy rule base. They can be divided into global and local approaches.

Global (or dynamic) methods rely on the idea that any anomaly can be detected by involving the rule evaluation in the analysis [38], [7], [39]. If an anomaly occurs within a fuzzy rule base, it imposes some constraint on the results of the inference. Thus, the inference operator is tightly involved in the verification process.

Local (or static) methods try to detect anomalies in a fuzzy rule base by using similarity, affinity or matching measures [4], [5], [16]. These methods examine subsets (e.g., a couple) of rules within the rule base and assess the degree of overlapping between the fuzzy sets associated to the propositions forming either conditions or conclusions.

The conceptual differences between the two verification approaches have repercussions for the achievement of different results as consequence [35]. This is the major outcome we need to deal with if we want to compare local and global approaches to the verification. In general, the
choice to use a specific approach depends on the particular application case analyzed [35].

For what concerns verification of condition-action clinical rules, global approaches do not result the most appropriate since the encoding of condition-action recommendations into fuzzy logic does not generate rule chaining. In fact, rule chaining expresses the dependencies among rules. More specifically, it expresses the dependencies among the actions of a rule and the conditions of other rules.

Condition-action clinical rules essentially involve 1-level fuzzy rules without chaining, since the action inferred by a rule is just a suggestion to be reported to the clinicians and, thus, it cannot generate any feedback for activating other clinical rules.

Therefore, local approaches result the most suitable to verify clinical rules encoded in fuzzy logic. However, the existing approaches proposed in literature can hardly lead to a practical verification process of fuzzy rule bases due to the lack of generality and the poor understandability of the suggested algorithms.

In such a direction, the proposed method essentially performs a local verification based on a general definition of the incompleteness, redundancy and incompleteness concepts for fuzzy clinical rules in terms of similarity between antecedents and consequents.

A gradual hint is integrated in these definitions, in accordance with the imprecise character of fuzzy decision support systems. The strength of this method relies on the formalization of the aforementioned anomalies in fuzzy degrees, which can be simply interpreted by the final users as suggestions for modifying the clinical rules, in order to eliminate or mitigate the undesired effects potentially caused by including conflicting, redundant or incomplete knowledge.

III. A FUZZY KNOWLEDGE BASE VERIFICATION METHOD

The proposed local method has been developed with the aim of detecting some kinds of structural anomalies within a Fuzzy Rule Base (FRB) encoding condition-action clinical recommendations.

Structural anomalies are not actual errors. Rather, they are symptoms pointing out that an inference process carried out by a decision support system on an FRB encoding these clinical rules could produce errors.

Structural anomalies can be thought as characteristics of the clinical guideline knowledge, independently of the knowledge coding formalism. However, each decision support implementation encodes clinical rules in a specific knowledge representation formalism that also defines both syntax and semantics of the anomalies.

In general, three kinds of structural anomalies may generate errors during a 1-level inference process on a set of formal rules. Usually, these kinds of anomalies, due to either rule conflicting or repeated or missing conditions and conclusions, are denoted as inconsistency, redundancy and incompleteness, respectively [22].

Incompleteness results in a conflict in the rule derivation. Redundancy increases the size of the rule set with unnecessary rules causing useless derivations and, consequently, additional time consumption during the inference process. Incompleteness prevents to activate certain rule derivations for some value of the variables.

In our method, inconsistency, redundancy and incompleteness are detected by means of a fuzzy measure used to estimate the similarity between the fuzzy sets representing the linguistic terms adopted in both the conditions and conclusions of rule base.

In fact, in fuzzy systems, differently from the classical systems, two or more sets can be ‘similar’, in the sense that there might be an overlapping between their membership functions. Thus, it is possible to define a fuzzy similarity measure that associates to every couple of fuzzy sets their degree of overlapping [29].

Among several definitions of fuzzy similarity measure [29], we have chosen to adopt a very common definition based on set-theoretic considerations. Given any two fuzzy sets A and B defined in a universe U, the fuzzy similarity measure used is defined by

\[ \sigma(A, B) = \frac{M(A \cap B)}{M(A) + M(B) - M(A \cap B)} \]  

where \( M(\cdot) \) associates to any fuzzy set the integral on U of its own membership function. \( \sigma \), as it is possible to see by its definition in (1), varies in [0, 1] and is 0 or 1 when the fuzzy sets are not at all or completely overlapped, respectively.

A. Representation and Interpretation of the clinical guidelines as a fuzzy rule-base

As assumption of the local verification approach designed, we have stated that the fuzzy rules constituting an FRB need to have both antecedents and consequents in Conjunctive Normal Form (CNF).

In the CNF, the rule antecedents are formed by conjunctions of a set of clause, each of them composed of a disjunction of a set of propositions with linguistic terms defined for an input variable.

Another important assumption is on the kind of inference process chosen to assess the collection of fuzzy rules constituting an FRB. We assume that the FRB has an implicative interpretation, i.e., rules are conjoined and their aggregation is realized through t-norms [34].

Let us start to present condition-action rule formalization by introducing some elements of the adopted symbology.

Given the collections of linguistic variables \{x\} and \{y\}, with \( u = 1, \ldots, n \) and \( v = 1, \ldots, m \) — representing the inputs and the actions of the FRB, respectively —, and their relative fuzzy sets \( A_v \) and \( B_u \), let us consider a generic fuzzy rule \( R_i \):

\[ R_i: \quad \text{IF } A_v(x_1, \ldots, x_n) \text{ THEN } B_u(y_1, \ldots, y_m) \]
The condition \( A_i(x_1, \ldots, x_n) \) is written as:
\[
A_i(x_1, \ldots, x_n) = A_i^1(x_1) \land \cdots \land A_i^n(x_n)
\]
with
\[
A_i^u(x_i) = (x_i \in A_i^u) \lor (x_i \in A_i^u) \lor \cdots \lor (x_i \in A_i^u)
\]
where the number \( J_i \), depending on \( i \) and \( u \), is the cardinality of the collection \( \{A_i^u\} \), with \( j = 1, \ldots, J_i \), composed of the linguistic terms \( A_i^u \) present in the clause \( A_i(x_n) \).

On the other hand, the consequent is defined as a series of propositions connected by conjunctions:
\[
B_i(y_1, \ldots, y_m) = (y_1 \in B_i^1) \land \cdots \land (y_m \in B_m^m)
\]

The statement in (2) can also be expressed through the implication connective as follows:
\[
R_i: \quad A_i(x_1, \ldots, x_n) \rightarrow B_i(y_1, \ldots, y_m)
\]
which offers a more compact way to represent the (2).

The choice of these specific representation and interpretation has several motivations. First of all, this form has a high degree of compactness and knowledge synthesis. Furthermore, it is worth noting that, in an implication-based FRB, the presence in \( B_i(y_1, \ldots, y_m) \) of predicates in disjunction can be always bypassed by splitting the relative rule in two different rules, each one with the consequents formed by a clause of the original-rule consequent.

Finally, we can indicate as \( \{R_i\} \), with \( i = 1, \ldots, L \), the set of \( L \) fuzzy rules with antecedents and consequents in CNF representing the FRB associated to the clinical condition-action rules extracted from the examined guidelines, while we can denote as \( \{A_i^u\}_{i=1}^{H_u} \), with \( h = 1, \ldots, H_u \), the set of \( H_u \) linguistic terms defined over the input variables \( \{x_u\} \) and actually used in \( \{R_i\} \).

B. Inconsistency and Redundancy measures

By assuming that fuzzy rules are represented in the form expressed in definition (3), it is possible to introduce a novel method to statically check whether a couple of rules is conflicting or unnecessary (in the sense that one rule is subsumed or generalized by another). Such a method relies on the extension of similarity concept from linguistic terms (fuzzy sets) to both rule conditions and conclusions.

As matter of fact, by using definition (1) regarding similarity between fuzzy sets, we construct a definition of similarity for rule antecedents (SRA) and for rule consequents (SRC).

Let us consider two fuzzy rules:
\[
R_i: \quad A_i(x_1, \ldots, x_n, \ldots, x_p) \rightarrow B_i(y_1, \ldots, y_m)
\]
\[
R_i: \quad A_i(x_1, \ldots, x_n, \ldots, x_p) \rightarrow B_i(y_1, \ldots, y_m)
\]
then SRA and SRC of these rules are defined as follows:

\[
SRA(i,k) = T(\left\{ \sigma(A_i^u, A_k^u) \right\}_{i,j \in U})
\]

(4)

\[
SRC(i,k) = T(\left\{ \sigma(B_i^u, B_k^u) \right\}_{i,j \in U})
\]

(5)

where \( T(\cdot) \) and \( S(\cdot) \) are the T-norm and the T-conorm (or S-norm) implementing the logical connectives \( \land \) and \( \lor \). Practically, we propose to compute the SRA by firstly measuring, for each input variable, the similarity between the terms in the disjunctive clause and making, subsequently, their S-norm; secondly, by calculating the T-norm of the S-norm values achieved over the input variables.

To better illustrate the similarity measures proposed, we provide a simple instance. Let us consider the rules
\[
R_1: \quad (x_i \in A_i^1) \lor (x_i \in A_i^2) \rightarrow (y_i \in B_i^1)
\]
\[
R_2: \quad (x_i \in A_i^3) \land (x_i \in A_i^4) \rightarrow (y_i \in B_i^2)
\]
where the \( \land \) and \( \lor \) are in this case the connectives introduced by Zadeh [40]. By applying the (3) and the (4), we have:
\[
SRA(1,2) = \min \left\{ \max \left[ \sigma(A_i^1, A_i^2), \sigma(A_i^1, A_i^1), \sigma(A_i^2, A_i^1) \right] \right\}
\]
\[
SRC(1,2) = \sigma(B_i^1, B_i^2)
\]

At this point, we can describe how to estimate a measure of the consistency and the redundancy of two fuzzy condition-action clinical rules and, subsequently, how to extend it to an FRB as a whole.

In general, fuzzy condition-action clinical rules are considered inconsistent if they have similar antecedents, but dissimilar consequents. That is to say, differently from the classic logic, two rules can be also inconsistent if their premise parts and consequence parts are necessarily not, respectively, the same and different ones. On the other hand, it is reasonable to affirm that two rules could contradict each other if their premises have little similarity and their consequences are pretty similar.

In the same way, two fuzzy condition-action clinical rules are considered redundant if both their antecedents and consequents are similar. However, this structural anomaly can occur in different ways based on the relationships among the antecedents.

Therefore, from the aforementioned considerations, the concepts of consistency and redundancy are not concrete in fuzzy logic and they can only be described by a degree.

A definition of consistency that seems to satisfy the discussed requirements has been provided by [14]. Given two fuzzy rules \( R_i \) and \( R_k \), their consistency is:
\[ Cons(R_i, R_j) = \exp \left[ -\frac{(SRA(i,k) - 1)^2}{SRC(i,k)} \right] \] (6)

Analogously, we propose a new definition for the redundancy. Let \( R_i \) and \( R_j \) be two fuzzy rules, their degree of redundancy is given by:

\[ Red(R_i, R_j) = \exp \left[ -\frac{(SRA(i,k) + SRC(i,k) - 2)^2}{(SRA(i,k) \cdot SRC(i,k))} \right] \] (7)

Definitions (6) and (7) have some fundamental properties. Both of them have value 1 if the rules have exactly both the same antecedents and consequents. When the rules have the antecedents and consequences totally different, consistency has value equal to 1 and redundancy reaches its lowest value of 0.

If the antecedents are exactly the same but the consequents are totally different, then the consistency and the redundancy go to 0. In the opposite case, i.e., when antecedents are totally different and consequents are exactly the same, redundancy is 0 but consistency has the value 1. In all the other cases, when both antecedents and consequents are similar in some degree, consistency and redundancy ranges in [0, 1].

Finally, starting from (6) and (7), we are able to calculate the degrees of inconsistency and redundancy for a fuzzy rule base as a whole by summing up the degrees of inconsistency and redundancy of a set of rules \( \{R_i\} \), with \( i = 1, ..., L \):

\[ Inc(\{R_i\}) = \frac{2 \sum_{i} \sum_{k} [1 - Cons(R_i, R_j)]}{L(L-1)} \] (8)

\[ Red(\{R_i\}) = \frac{2 \sum_{i} \sum_{k} Red(R_i, R_j)}{L(L-1)} \] (9)

As for their correspondent versions for couples of rules, \( Inc(\{R_i\}) \) and \( Red(\{R_i\}) \) are still two indexes varying in the range \([0,1.0]\) because they are normalized over the number of distinct comparisons in the rule base.

In our opinion, such a method permits to naturally conciliate the satisfaction of the fuzzy formalism with the intuition and the common sense of human beings.

C. Incompleteness measure

Completeness is a necessary property for an FRB encoding guidelines. More in general, in the context of artificial intelligence, completeness assesses how much a deductive system can represent every situation belonging to the encoded domain.

Usually, scientific literature distinguishes among several forms of completeness in Fuzzy Logic: the completeness of a fuzzy set collection defined over the same linguistic variable and the completeness of an FRB (a fuzzy model, in general) [19].

An FRB is said complete when a proper action is returned as output to every potential input [19]. When this does not happen, inference process for FRB generates an undesirable behavior depending on its interpretation form. In the case of implicational interpretation, adopted as assumption in our method, the result of the inference process is random when no rule fires because an input; consequently, the output of the FRB is undefined [19].

Usually, incompleteness in FRB is due to the fact that a fuzzy set collection \( F \) defined over the input variables \( x \) does not totally cover the corresponding universe of discourse \( U \) [20]. More formally we could say that a FRB defined over \( U \) is incomplete if:

\[ \exists x \in U \quad \forall A \in F, \quad \mu_{A}(x) = 0 \]

However, completeness property is not enough to avoid some situations that may influence the efficacy of the inference process.

As example, Figure 1 shows two cases where completeness does not entail a proper behavior of the FRB in output. In pane (a), there is a lack of physical meaning associated to the considered fuzzy sets, although their partitioning of the universe of discourse is complete.

On the contrary, in pane (b) it is depicted a case where two fuzzy sets are distinct, their partitioning is complete but there could be an improper behavior of the FRB, for instance a weak firing of some rule, with that element of the universe where their membership function values are close to 0.

Additionally, it is worth noticing that this completeness measure does not give any information about the intelligibility of the FRB [19].

In fact, the intelligibility, for a generic FRB, strongly depends on the interpretability of the fuzzy sets delineated for its variables. In its turn, a fuzzy set collection can be designated as “interpretable” if the fuzzy sets making it up are sharply distinguishable [19], i.e., they do not overlap too much. For the above reasons, it is convenient to adopt a
measure able to assess how much some linguistic terms are distinguishable.

In the approach we propose, the two concepts of coverage and distinguishability are merged in order to avoid the aforementioned problematic situations. As a result we obtain a definition of incompleteness for FRBs based on the similarity measures of the fuzzy sets involved in their rules.

As shown in Section III.A, by presuming that $A^h_{x_u}$ denotes a fuzzy set belonging to the collection $\{A^h_{x_u}\}_{h}$ defined over the variable $x_u$ and used in the rule base $\{R_i\}$, we can indicate as $[A^h_{x_u}]_s$ the $\alpha$-cut of $A^h_{x_u}$.

This being said, the measures adopted to estimate coverage and distinguishability of a rule base $\{R_i\}$ with respect to a single variable $x_u$ are:

$$ C_\alpha(x_u, \{R_i\}) = \sigma \left( \operatorname{supp} \mu \left( \bigcup_{j=1}^{n_x} [A^h_{x_u}]_s \right) \bigcap U_{x_u} \right) \quad (10) $$

$$ D_\alpha(x_u, \{R_i\}) = 1 - S \left( \sigma \left( [A^h_{x_u}]_s \bigcup [A^h_{x_u}]_s^{\alpha} \right) \right) \quad (11) $$

respectively, where

$$ \operatorname{supp} \mu \left( \bigcup_{j=1}^{n_x} [A^h_{x_u}]_s \right) $$

stands for a set of elements of the universe of discourse $U_{x_u}$ corresponding to $x_u$ such that the membership function $\mu$ of the union of the fuzzy sets $A^h_{x_u}$ defined over $U_{x_u}$ is greater than a parameter $\alpha \in [0,1]$, and

$$ \sigma \left( [A^h_{x_u}]_s \bigcup [A^h_{x_u}]_s^{\alpha} \right) $$

is the similarity measure between the $\alpha$-cuts of two neighboring fuzzy sets.

We call the (10), already introduced in [12], as $\alpha$-coverage measure of $\{R_i\}$ with respect to $x_u$ because it represents the degree of the rule-base coverage, constrained by a parameter $\alpha \in [0,1]$, over the universe of discourse $U_{x_u}$, through the linguistic terms actually expressed in $\{R_i\}$. It is worth underlining that although this definition is stronger than the simple coverage, it can be viewed as a valid requirement to limit “weak firing” phenomenon shown in the pane (b) of Figure (1) [25].

In this regard, the (10) can be considered as an extension to a fuzzy rule base of the measure proposed by [20] in order to assess the coverage of a finite collection of fuzzy sets defined over the same linguistic variable.

$C_\alpha(x_u, \{R_i\})$ has values in the real interval $[0,1]$: when $C_\alpha$ is equal to 1 (or, contrarily, equal to 0), it means that, the $\alpha$-cuts of the fuzzy sets $\{A^h_{x_u}\}_{h}$ defined over $x_u$ and used in $\{R_i\}$, have a support with values greater (less) than over $U_{x_u}$.

On the other hand, we propose the (11) as $\alpha$-distinguishability measure, with $\alpha \in [0,1]$. Practically, it consists in an assessment of the overlapping degree of the $\alpha$-cuts of the neighboring fuzzy sets defined over the same variable.

As in the case of the $\alpha$-coverage, the value of $D_\alpha(x_u, \{R_i\})$ ranges in $[0,1]$ and it is equal to 1 when all the $\alpha$-cuts of the neighboring fuzzy sets defined over $x_u$ and expressed in $\{R_i\}$ are not overlapped at all, while it is equal to 0 if the $\alpha$-cuts completely overlap.

Hence, up to now, we have introduced two new measures revealing information about two different and important features of the rule base completeness. These features are connected to the capacity of linguistic terms to represent input space of a rule base: the coverage and the distinguishability of the fuzzy sets computed through the (10) and the (11), respectively.

When both $\alpha$-coverage and $\alpha$-distinguishability have high values, it means that the fuzzy set collection expressed in the rule base constitutes an input space partitioning with good completeness characteristics.

Moreover, the parameter $\alpha$ has a fundamental importance in our definitions. In fact, $\alpha$-coverage and $\alpha$-distinguishability assume complementary behaviors as function of the parameter $\alpha$. For $\alpha \rightarrow 1$, $\alpha$-coverage decreases and $\alpha$-distinguishability increases while they act oppositely for $\alpha \rightarrow 0$.

From these considerations, we can define as $\alpha$-incompleteness measure of a rule base $\{R_i\}$ with respect to the input variable $x_u$ the following non-linear expression:

$$ \text{Incomp}_\alpha(x_u, \{R_i\}) = \exp \left( - \frac{C_\alpha + D_\alpha}{4 - (C_\alpha + D_\alpha)^2} \right) \quad (12) $$

As it is possible to see, the measure (12) has values in $[0,1]$, with value 1 when both $C_\alpha$ and $D_\alpha$ are 0 and value 0 when $C_\alpha$ and $D_\alpha$ are 1.

Extending the $\alpha$-incompleteness definition over the entire set $\{x_u\}_i$ of input variables defined in $\{R_i\}$ is immediate. It is sufficient to sum up and to average the expression (12) over every input variable:

$$ \text{Incomp}_\alpha(\{R_i\}) = \frac{1}{n} \sum_{u} \text{Incomp}_\alpha(x_u, \{R_i\}) \quad (13) $$

with $n$ to denote the cardinality of $\{x_u\}_i$, i.e., the number of input variables defined in an FRB.

As a finally remark, incompleteness is a global property of the considered rule base. Therefore, its evaluation makes sense only if it is carried out over the rule base as whole, differently by the inconsistency and redundancy measures whose estimations stem from a local computation over couples of rules.
D. A procedural definition for the proposed verification method

The presented verification method can be easily implemented in an algorithmic procedure by following these instructions:

1. For every couple of rules belonging in the FRB
   1.1. Use the similarity measures defined in equation (1) to calculate the SRA and SRC of the rules by adopting equations (4) and (5);
   1.2. Estimate Consistency and Redundancy degrees of the examined rules through the equations (6) and (7), respectively;
2. Compute Consistency and Redundancy degrees of the FRB as a whole by means of equations (8) and (9)
3. Fixed the parameter \( \alpha \), for every variable used in the FRB
   3.1. Calculate the \( \alpha \)-coverage and the \( \alpha \)-distinguishability measures of the FRB respect to considered variable by using equations (10) and (11);
   3.2. Estimate the \( \alpha \)-incompleteness degree for the FRB respect to the considered variable by exploiting the equation (12);
4. Compute the \( \alpha \)-incompleteness degree for the FRB respect to the universe of discourse \( U \) through definition (13).

IV. APPLICATIVE SCENARIOS

The verification method described above has been validated with practical cases. For this purpose we have considered two different scenarios.

In the first scenario, some clinical recommendations extracted from the GOLD guideline [27] for the Chronic Obstructive Pulmonary Disease (COPD) are verified in order to detect inconsistency, redundancy and incompleteness.

In the second scenario, clinical crisp rules extracted from the widely used Wisconsin Breast Cancer Dataset (in the following, WBCD) [21] for diagnosing breast masses have been fuzzified by an automatic procedure [7]. We assess the completeness of this fuzzification and propose suitable solutions.

A. The GOLD guidelines

GOLD guideline [27] includes the best evidence related to diagnosis, management and prevention of Chronic Obstructive Pulmonary Disease (COPD), formalized by the US National Heart, Lung, and Blood Institute and the World Health Organization in 2006.

COPD is a preventable and treatable disease with some significant extra-pulmonary effects, which can lead, in its severe forms, to respiratory failure, hospitalization and eventually death from suffocation.

Spirometry is the standardized and reproducible test that objectively confirms the presence of airflow obstruction, and, thus, is indispensable in establishing the diagnosis of COPD. Characteristically, spirometry shows a decreased forced expiratory volume in one second (FEV\(_1\)) and a decreased \( FEV_1/FVC \) (forced vital capacity) ratio [33].

Starting from the degrees of \( FEV_1 \) and \( FEV_1/FVC \) and in addition to presence of Chronic Respiratory Failure (CRF), shown in Table I, the GOLD guideline has developed a set of clinical rules for classifying the COPD stage, as reported into Table II.

These clinical rules, which involve rigid boundaries, have been modified according to a smooth fashion and encoded in terms of fuzzy rules. More in detail, three linguistic variables have been assigned to \( FEV_1 \) predicted, \( FEV_1/FVC \) ratio and CRF, respectively, for modeling the conditions in the rules. Additionally, the linguistic variable Stage has been defined to model the action part in the clinical rules. In accordance with the knowledge formalized into Table I, physicians have identified the linguistic terms shown in Table III.

For what concerns the definition of the membership functions for these terms, the degrees are given depending on the intervals provided for each parameter into Fig. 2. In particular, for each term, a transition region representing its fuzzy boundary has been identified as shown in Fig. 2.

A trapezoidal membership function has been used for the variables \( FEV_1 \) predicted, \( FEV_1/FVC \) ratio and Stage, whereas a singleton membership function has been applied to CRF, since it models a categorical concept.

These linguistic variables and terms have been used to write a set of seven “if-then” rules aimed at identifying the appropriate COPD stage, as shown in the Table IV. This set of rules has been intentionally encoded wrongly, by including both an inconsistent and a redundant rule.
Furthermore, they keep some aspects of incompleteness already present in the set of clinical recommendations used for this scenario and reported in Table II.

The proposed verification method has been thus applied to the set of rules shown in Table IV, assuming, in order to calculate the similarity measures, the use of the Zadeh implementation for the connectives \{\&, \vee\}.

Equation (6) and (7) have permitted to identify the third rule as redundant with respect to the second one with a redundancy degree equals to 1, whereas the fifth rule has been detected as inconsistent with both the first, the second and the third ones with a consistency degree equals to 0, approximately. By using the (8) and (9), the inconsistency degree of the whole rule base is equal to 0.14 (3/21), whereas the redundancy degree is equal to 0.047 (1/21).

For what concerns the \(\alpha\)-incompleteness degree of the rule set, we have applied to each input variable the definition (12) by setting the value of the parameter \(\alpha\) formerly to 0 and then to 0.5. That has been done to illustrate how our measure varies as function of \(\alpha\).

Table V reports the values of the \(\alpha\)-coverage and the \(\alpha\)-distinguishability for each linguistic variable used in the conditional part of the rules (definitions (10) and (11)) and the relative \(\alpha\)-incompleteness measure.

To compute the incompleteness degree one needs to pay attention to the fact that only the fuzzy set “Low” is used in the rules as a term of the variable \(FEV/FVC\) ratio – this determines a not complete coverage of the universe of discourse – and that a singleton fuzzy set is always assumed to have coverage and distinguishability equal to 1, for all the values of \(\alpha\).

Besides, as it is possible to notice, a greater value has been attained for \(\alpha=0.5\); this is due to the properties of convexity and orthogonality of the fuzzy sets involved in the rules.

As a result, by exploiting such indications provided in terms of consistency and redundancy degrees, the fuzzy rule base has been updated by deleting from Table IV both the third and the fifth one, as reported in Table VI.

In terms of completeness, it is possible to conclude that the extracted rules have an incompleteness degree, intrinsic to the chosen rule base, depending on the parameter \(\alpha\). As consequence to increase the completeness of the rule base we have to conveniently change it by either adding or modifying a rule.

In this specific case, we do not propose any modification to the rule base since the set of clinical recommendations, shown in Table I, which we extracted the fuzzy rules from are intrinsically incomplete. That could be deduced by observing that the proposition “\(FEV/FVC \geq 70\%\)” misses in Table I.
Thus, the verification of the former rule base (Table IV) has allowed to detect structural anomalies and, successively, to adjust the rule base as shown in Table VI, so that it results coherent and consistent with respect to the GOLD guideline.

### B. Clinical rules from Wisconsin Breast Cancer Database

The verification method has been further validated with another practical case, consisting in some clinical rules aimed at accurately diagnosing breast masses and extracted from the widely used Wisconsin Breast Cancer Dataset, (in the following, WBCD).

Particularly, in this specific test case we want to deeply test the \( \alpha \)-incompleteness measure introduced in Section III.C.

WBCD was built by means of image processing on fine needle aspiration (FNA) of breast masses collected at the University of Wisconsin [21]. The samples contain visual assessment of the nuclear features of fine needle aspirates collected from patients’ breasts and can be obtained from UCI (University of California at Irvine) machine learning repository.

The version of WBCD used consists of 10 features obtained from FNA and shown in Table VII, each one with the related range. The two outputs are benign and malignant. All the instances were properly recorded without any missing value. The diagnosis class is distributed with 320 benign samples and 194 malignant samples.

Clinical rules used for the verification were extracted by using a decision tree algorithm, i.e. J48, packaged within the WEKA toolkit [10]. During the execution, J48 algorithm has also selected which features needs to be used in the rules in order to maximize their classification rate. This algorithm was applied only to the mean values calculated for each the feature of the WDBC database.

The resulting rules are reported in Table VIII.

Similarly to the previous test, also in this case, these clinical rules, which involve rigid boundaries, have been modified according to a smooth fashion and encoded in terms of fuzzy rules.

More in detail, among all the features included into the WBCD dataset, only the ones appearing into the conditions of the rules reported into Table VIII have been formalized as fuzzy linguistic variables, namely Concave Points, Texture and Area. Additionally, the linguistic variable Diagnosis has been defined to model the action part in the clinical rules.

In accordance with the knowledge formalized into Table VIII, the linguistic terms shown in Table IX have been firstly identified and, then, their membership functions have been automatically calculated through a procedure presented in [7] depending on the intervals provided for each feature as reported in Fig. 3.

A trapezoidal membership function has been used for the variables Concavity Points, Texture and Area, whereas a singleton membership function has been applied to the categorical variable Diagnosis.

These linguistic variables and terms have been used to write a set of four “if-then” rules aimed at identifying the appropriate diagnosis of breast masses, as shown in the Table X.

In terms of completeness, Equation (12), with \( \alpha=0 \), has permitted to identify the variable Concave Points and Area as incomplete with degrees equal to 0.21 and 0.14, respectively, since they do not completely cover their universe of discourse. The values of the \( \alpha \)-coverage and \( \alpha \)-distinguishability measures are shown in Table XI, together with the value of the \( \alpha \)-incompleteness of both the single variable and the rule base as whole.

As a result, by exploiting such indications for this couple of variables, provided in terms of incompleteness degree, the membership functions for the terms of both the variables Concave Points and Area shown in Fig. 3 have been correctly modified, as reported in Fig. 4.
V. CONCLUSIONS AND FUTURE WORKS

In this paper, a solution to determine the degree of inconsistency, redundancy and incompleteness for clinical rule bases encoded into a fuzzy decision support system is presented.

In order to do that, the classic definitions of inconsistent, redundant and incomplete rule base have been reformulated in terms of degrees by taking in account the measures of similarity between antecedents and consequents.

The degrees of inconsistency, redundancy and incompleteness obtained should be interpreted as measurements suggesting the modifications to be applied to the clinical rules in order to eliminate or mitigate the undesired effects eventually caused by including contradictory or redundant knowledge.

In this sense, experts obtain, by means of the degree of either inconsistency or redundancy or incompleteness, information about the direction to which they should look for possible improvements of the set of clinical recommendations encoded into a decision support system.

The method has been profitably assessed on a two scenarios. In the first sample, a set of condition-action clinical recommendations aimed at identifying Chronic Obstructive Pulmonary Disease (COPD) has been chosen from the relevant clinical literature.

The second scenario is a sample of a fuzzy rule base automatically extracted from a widely known clinical
In the future, we will investigate an extension of our local verification method with the aim of working with bases of rules without a fixed form and composed of a number of variables different for each rule.

REFERENCES


