THE ANNALS OF "DUNAREA DE JOS" UNIVERSITY OF GALATI
FASCICLE III, Vol.31, No.2, ISSN 1221-454X
ELECTROTECHNICS, ELECTRONICS, AUTOMATIC CONTROL, INFORMATICS

MATHEMATICAL ANALYSIS OF UNCERTAINTY

Angel GARRIDO

Facultad de Ciencias, UNED
Departamento de Matemáticas Fundamentales
Senda del Rey, 9, 28040-Madrid, Spain
E-mail: algbmv@telefonica.net

Abstract: Classical Logic showed early its insufficiencies for solving AI problems. The introduction of Fuzzy Logic aims at this problem. There have been research in the conventional Rough direction alone or in the Fuzzy direction alone, and more recently, attempts to combine both into Fuzzy Rough Sets or Rough Fuzzy Sets. We analyse some new and powerful tools in the study of Uncertainty, as the Probabilistic Graphical Models, Chain Graphs, Bayesian Networks, and Markov Networks, integrating our knowledge of graphs and probability.

Keywords: Fuzzy Sets, Rough Sets, Fuzzy Analysis, Knowledge Representation, Artificial Intelligence.

1. FUZZINESS

Classical Logic showed early its insufficiencies for solving AI problems. We need a more flexible tool, making gradation of certainty possible, indicating different degrees of membership to a set or in the fulfilment of a relation. In other words, that allows for subtleties. For instance, in classical Set Theory, each element either belongs totally to the set, or it does not belong at all, without any possible intermediate situation. Also, relations can be either verified or not, but not partially verified. And this ought to be possible, because it is so in the real world, very different indeed to formal worlds. The ideas of set, relation, etc. must be modified in the sense of covering adequately the indetermination or imprecision of the real world, introducing so Fuzzy Logic. Other nuances are introduced by Modal Logic, with two new concepts: necessity and possibility. Or with the 3-valued Logic of Lukasiewicz-Moisil (with truth values: 0, 1/2 and 1), giving (through the corresponding generalization) the multi-valued Logic.

We define the "world" as a complete and coherent description of how things are or how they could have been. In the problems related with the "real world", which is only one of the "possible worlds", Monotonic Logic often fails. But such type of Logic is the classical in formal worlds, such as in Mathematics. It is a real problem, because the "common sense" logic is non-monotonic, and this is our usual logic.
An element of the Universe, U, can belong more or less to an arbitrary set C. It can belong to C in different degrees.
From 0, when it belongs absolutely nothing to C, to 1, when it belongs totally to C. Or in any intermediate degree, like: 0.5, 0.3, 0.1..., but always between 0 and 1, both values included.
Such "membership degree" value can be assigned by an adequate "membership function", which range is the closed unit interval, [0,1].

This paper was recommended for publication by Cornelia Tudorie

So, the application can be expressed by $f: C \rightarrow [0,1]$, informing about the "membership degree", of such element, x, of the universe U, to the set C.
In a Classical Set, therefore, the range of the function f should be reduced to the set $\{0, 1\}$.
Given n universes of the discourse, we define a *fuzzy relation, R,* as a membership function that associates each n-tuple, a value of the unit closed interval, [0, 1].
The fuzzy relation, R, can be defined through such "membership function".
The Cartesian product of two fuzzy sets, F and G, will be a fuzzy binary relation, through the minimum between the membership degrees. Sometimes, it is very useful to symbolize each fuzzy relation as one matrix, where the entries can be any real number between 0 (not related) and 1 (totally related, or simply, related).

There exists a clear analogy between the composition of fuzzy relations and the product of matrices.
To show this connection, it is sufficient to establish the correspondence

One + → one max

One • → one min

For this reason, the composition of fuzzy relations can also be called "max-min matrix product".

As a particular case of the previous definition for the composition between fuzzy relations, we can introduce the composition between a fuzzy set and a fuzzy relation. This can be very useful in the "Fuzzy Inference" where we attempt to obtain new knowledge from the available. Obviously, in such case, the fuzzy set can be represented by one row matrix, or a column matrix, depending on the order in the product.

The usual properties of the classical relations can be translated to fuzzy relations, but the transitive will be modified.

- R is *Reflexive* if

$$R(x, x) = 1$$

for each x in the set C, into the universe, U.

According to this, each element would be totally related with itself, when R is reflexive.

- R is *Symmetric* if

$$R(x, y) = R(y, x), \text{ for each pair } (x, y)$$

Therefore, the principal diagonal acts as a mirror, in the associate matrix.

- R is *Transitive*

Not in the usual way for relations or associate matrices, but when:

$$R(x, z) \geq \max(\min\{R(x, y), R(y, z)\})$$

All these mathematical methods can be very useful in Fuzzy Logic and in many branches of Artificial Intelligence; therefore, in Computational Intelligence.

We can introduce new generalized versions of Classical Logic. So, the Modus Ponens Generalized, or the Modus Tollens Generalized or the Hypothetic Syllogism.

To each Fuzzy Predicate, we will associate a Fuzzy Set: the defined by such property, that is, composed by the elements of the Universe such that totally or partially verify such condition.

So, we can prove that the class of fuzzy sets with the operations $\cup$, $\cap$ and c (path to the complementary) does not constitute a Boolean Algebra, because neither the Contradiction Law nor the Third Excluded Principle work in it. Geometrical and algebraic proof is easy, by a counterexample: it suffices taking an element with membership degree that belongs to (0,1).

2. ROUGHNESS

The concept of Rough Set was introduced by the polish mathematician Zdzislaw Pawlak in 1982. From then, some theoretical advances with the corresponding applications are emerging.
It is possible to apply Rough concepts to a wide range of problems, as the prediction of financial risk, voice recognition, image processing, medical data analysis and so on.
Taking object, attribute or decision values, we will create rules for them: upper and lower approximations and boundary approximation. Each object is classified in one of these regions, in this way:

For each rough set, $A \subseteq U$, we dispose of

- *Lower Approximation of A:*
Collection of objects which can be classified with full certainty as members of A.

- *Upper Approximation of A:*
Collection of objects which may possibly be classified as members of A.

Obviously, this class is wider than the aforementioned, containing between both the Rough set.

Rough Set Theory is a model of Approximate Reasoning. According to it we will interpret knowledge as a way to classify objects.
We dispose of *U, the universe of discourse,* composed by objects, and an *equivalence relation* on *U*, denoted *R.*
The *procedure* is to search a collection of subsets in *U* (categories), such that all the elements of the same class possess the same attributes with the same values.
So, we obtain a covering of *U* by a set of categories.
The elementary knowledge is encoded by a pair *(U,R),* composed of "elementary granules of knowledge". They constitute a partition in equivalence classes, into the quotient set, U / R.
Given two elements, it is possible to define when they are mutually indiscernible. For this, we say that is the *Indiscernibility Relation.*
Therefore, it is possible to introduce the application which assigns to each object its corresponding class.
Then, such indistinguishability allows us to introduce the *Fibre of* a_R, defined by the aforementioned relation *R*.
So, the collection of such fibres, in the finite case, produces a union: this union of fibres is called a *granule of knowledge*.
The pair *(U, R)* will be a *Knowledge Base*.

We say that an object, or category, is *R-rough,* if it is not *R*-exact.
For each *R*-rough set, $Y \subseteq U$, we define two associate *R*-exact sets

- the *R-lower approximation of Y*

$$\underline{R}\ Y = \{x \in U : [x]_R \subset Y\} \quad (1)$$

- the *R-upper approximation of Y*

$$\overline{R}\, Y = \{x \in U : [x]_R \cap Y \neq \emptyset\} \quad (2)$$

So, we can represent the Rough set, *Y*, through the pair ($\underline{R}\, Y$, $\overline{R}\, Y$).
Observe that:

$$\underline{R}\ Y \subseteq Y \subseteq \overline{R}\ Y,\ \forall Y \subseteq U \quad (3)$$

And furthermore:

$$Y \text{ is } R\text{-exact} \Leftrightarrow \underline{R}\ Y = \overline{R}\, Y \quad (4)$$

Given a Knowledge Base (KB):

$$K \equiv (U, R)$$

we will take the collection of classes

$$E_K = \{R \text{ - exact sets on } U\}$$

closed with respect to the usual set operations $\cup$, $\cap$ and *c* (complement).
It verifies the known properties of a *Boolean Algebra*. More concretely, we can call it a *Field of Sets*.
But it is not the case when we deal with *R*-rough sets. Because, for instance, the union of two *R*-rough sets can be a *R*-exact set.
The coincidence of this Rough Set Theory with the Classical Theory of Sets occurs when we only work with *R*-exact sets.
An interesting generalization of Rough Set will be the

Generalized Approximation Space,

denoted

GAS

It consists of a triple

$$(U, \mathrm{I}, \nu)$$

where *U* will be the *Universe; I,* the *uncertainty function,*

$$I: U \rightarrow \wp(U)$$

and ν the *Rough Inclusion Function.*

An example of this type of Rough Inclusion Function will be the *Generalized Rough Membership Function.* So, given any subset, we have both GAS - approximations, lower- and -upper.

3. APPROXIMATIONS

Its names induce to thinking of almost the same concept. But they are very different approaches to uncertainty in the set of data. It depends of the nature of vagueness in the problem, or the convenience in applications.
Both resources cover distinct aspects of the Approximate Reasoning. For this reason, both

paradigms address the Boundary Problem in Non-Crisp cases.
Dubois and Prade (1992) establish the relation between

1) Rough Fuzzy Set

and

2) Fuzzy Rough Set.

In the case 1), we will pass from fuzzy sets, through filtering, by the classical equivalence relations to quotient spaces, which are fuzzy sets.
In the case 2), we imitate the rough set approximation, but now with fuzzy similarity (instead of equivalence) relations.
We work within the collection of fuzzy sets on *U*, endowed with the operations: *max* and *min*.
So:

{Fuz (U, [0, 1]), max, min}

This produces a *Zadeh Lattice.*
And provided of the path to complementary operator

{Fuz (U, [0, 1]), max, min, c}

it is a *Brouwer-Zadeh Lattice.*

This lets us introduce now the adequate Rough Approximation to Fuzzy Sets.

Our purpose is double:

- Given $A \in Z(U)$, to induce a fuzzy set in U / R, by *A*.
- To reach the approximation of *A*, relative to *R*, according to the Rough Set Theory.

The notion of *Fuzzy Rough Set* is dual to the concept above.
We consider again the family of fuzzy sets in the universe *U*, with values in the closed unit interval

Fuz (U, [0, 1])

And we need to analyze the fuzzy notion of equivalence relation and then, the fuzzy partition induced. Respect to the equivalence relation, the closest concept is the T-Fuzzy Similarity Relation.

In the past, the relationship between Fuzzy and Rough concepts were studied by some mathematicians and computer scientists, as Pawlak, Nakamura, Pedricz, Dubois and Prade, Pal, Skowron and so on.

4. BAYESIAN NETWORKS

Another very promising and ever increasing way to handle uncertainty is based on Probabilistic Graphical Models; in particular, by Chain Graphs, Bayesian and Markov Networks.

Recall that every *graph,* in turn, is another pair

$$G = (V, E)$$

where *V* is the set of vertices, or nodes, and *E* the set of links or edges between them. And in the particular case of DAGs, we must add the condition of being a directed graph and without cycles.

Two nodes, *a* and *b*, are said to be *connected* in the graph, if there are paths from *a* to *b* and from *b* to *a*.
Both paths connecting *a* and *b* may be one and the same undirected path, or two distinct paths.
Connection defines a number of disjoint subsets of connected nodes, $\{C_i\}_{i=1,2,\ldots,n}$, referred to as *connectivity components*:

$$E = \cup C_i = C_1 \cup \ldots \cup C_n$$

Graph theory defines a number of different types of subsets of variables related to a specific node, as its parents, its neighbours and so on.

A *Chain Graph (*denoted *CG)* is a generalization of both classical types: Undirected and Directed Graphs, that is, it includes UGs and DAGs, being represented by undirected and directed edges. Therefore, they are mixed graphs, composed by directed and undirected edges.

Two CGs are *Markov Equivalent,* if they represent the same statistical model.

A *Bayesian Network,* or *Bayesian Net* (BN) is a pair (*G*, P), where *G* is a directed acyclic graph, or DAG, their structure and a probability distribution P, associated with each random variable, represented by a node into the graph.
So, a DAG is a BN, relative to a set of random variables if the joint distribution of the node values can be expressed as the product of local distributions of each node and its parents:

$$P(X_1, X_2, ..., X_n) = \Pi P(X_i / pa_{Xi}) \quad (5)$$

Therefore, Bayesian Networks, also called Belief Networks, are a Probabilistic Graphical Model (PGM) that represents a set of variables and their probabilistic independencies. In other words, they encode joint distribution probabilities.
Formally, a BN is a DAG whose nodes represent random variables, and whose missing edges encode CIs (conditional independencies) between the variables.

We say that two BNs are *equivalent* (denoted by the symbol X), if both represent the same joint probability distribution.
The following *properties* hold:

- *Reflexive: B* X *B,* ∀ *B*
- *Symmetrical: if B* X *B′⇒ B′*X *B*
- *Transitive: if B* X *B′ and B′*X *B′′⇒ ⇒ B* X *B′′*

Therefore, it is an Equality or Equivalence Relation**,** defined on the BNs set. On such mathematical object, it is well established a partition in equivalence classes, as we will later see.

Let *S* and *S'* be two of such structures of BNs on *V*. Then, we say that *S is equivalent to S'*, which is denoted by S X S', if for each parameterization, θ, of *S*, there exists another parameterization, θ', of *S'*, such that

$$P(V / S, \theta) = P(V / S', \theta') \quad (6)$$

Let *C* be a class of DAGs Markov equivalent among them. Then, their essential graph would be the smallest graph greater than every DAG that belongs to the class. If we denote the essential graph as *G**, this is equivalent to saying

$$G^* = \cup \{G: G \in C\} \quad (7)$$

where such graph union is reached by the union of the nodes and edges of *G*

$$V(G^*) = \cup V(G) \quad (8)$$

$$E(G^*) = \cup E(G) \quad (9)$$

The directed edges connecting the same pair of nodes, but showing opposed directions, in two graphs belonging the same class, *C*, are substituted by a line. Therefore,

$$G^* = \min [\max \{G: G \in S\}] =$$

$$= \cap [\cup \{G: G \in S\}] \quad (10)$$

Being S the equivalence class that represents *G*, with min ⇔ inf and max ⇔ sup, for infinite graphs.
So, *G** will be the lesser of the upper bounds for every graph of the class represented.

Many efforts are devoted in the last years to develop this rich field, in the intersection of Graph Theory and Probabilistic Methods.

And also many different lines of advance in the research, with new and interesting support by more geometrical tools, such as

- *Petri Nets*
- *Voronoi-Thiessen-Dirichlet Diagrams,*

  or its dual
- *Delaunay Tesellations,*

  and also by
- *Hypergraphs*.

Or algebraic tools, in the last times, as

- *Imsets*
- *Graphoids*
- *Semigraphoids*
- *Matroids*

And so on. All of them very useful in many other fields.

5. CONCLUSIONS

It is possible to establish stronger mathematical foundations for research with *Probabilistic Graphical Models,* connecting with the aforementioned tools and also by the essential theory of *Random Graphs,* created by Paul Erdös and Alfred Rényi, and from then cultivated by members of the Hungarian graph school, as Bela Bollobás and Laszlo Lovász. Or by numerous Romanian scientists with great contributions to Computer Science as the pioneer Grigor Moisil, Solomon Marcus, Florin *G.* Filip, and many others.

I hope this paper contributes to a clearer vision, being these concepts necessary to advance in AI. Because both theories are two complementary ways of reaching the summit: solving more efficiently a fundamental problem of data, Uncertainty and Vagueness.

6. REFERENCES

Anderson, Madigan, and Perlman (1997): A characterization of Markov equivalence classes for acyclic digraphs. Ann. Stat. 25, N. 2, p. 505-541.

Cattaneo (1998). Abstract approximation spaces for rough theories, in Polkowski and Skowron (eds.): *Rough Sets in Knowledge discovery. Methodology and Applications,* Vol. Nr. 18, pp. 59-98. *Studies in Fuzziness and Soft Computing,* Physica-Verlag, Heidelberg.

Ibid. (1997), Generalized rough sets. Preclusivity fuzzy-intuitionistic (BZ) lattices, *Studia Logica,* 58, pp. 47-77.

Chickering (1995): "A transformational characterization of equivalent Bayesian Networks structures". *UAI'95,* p. 87-98.

Dubois and Prade (1992). Putting rough sets and fuzzy sets together, in Slowinsky (ed.): *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Set Theory,* pp. 203-232, Kluwer Publ., Dordrecht, Holland,

Fariñas del Cerro and Prade (1986). Rough Sets, twofold fuzzy sets and modal logic-fuzziness in indiscernibility classes and partial information, in Di Nola and Ventre (eds.): *The Mathematics of Fuzzy Systems,* Verlag TÜV Rheinland, Köln.

Garrido (2009): Chain Graphs and Directed Acyclic Graphs improved by Classification and Essential Graphs. *SIC (Studies in Informatics and Control),* vol. 14, special nr. 4, 6 pp., invited paper to the special issue in honor of the 60th birthday of Prof. Neculai Andrei. ICI Bucharest, accepted, to appear in March.

Gillispie and Perlman (2001): "Enumerating Markov Equivalence Classes of Acyclic Digraphs", *UAI'2001*, p. 171-177.

Mantaras and Valverde (1988). New results in fuzzy clustering based on the concept of indisguishability relation, *IEEE Trans. on Pattern Analysis and Machine Intelligence,* pp. 754-757.

Oxley (1992): *Matroid Theory.* Oxford University Press, New York.

Roverato and Studený (2006): "A graphical representation of equivalence classes of AMP chain graphs", *JMLR* 7, p. 1045-1078.

Ziarko (1994). *Rough Sets, Fuzzy Sets and Knowledge Discovery*, Springer-Verlag, Berlin.