INTRODUCTION TO FUNCTIONAL SET THEORY AND ITS PRINCIPLES

DJAMEL BOUCHAFFRA

Department of Computer Science & Engineering
131 Dodge Hall, Oakland University
Rochester, MI, 48309, USA
Telephone: (248) 370-2242
dbouchaffra@ieee.org

ABSTRACT:
We propose a novel mathematical domain that we named "functional set theory" (f-set theory). It extends traditional set theory by considering all possible mappings between objects in a traditional set. Because of our need to search structural information, f-set theory focuses more on relationships between objects and less on the objects themselves. We developed the notions of "role" assigned to mappings between objects and membership values assigned to objects. We introduce some functional-set theoretic operations as well as the notion of functional reasoning. We show some significant impacts of f-set theory in the area of artificial intelligence, and soft computing.

1. INTRODUCTION

From Aristotle to Cantor, the focus of human research was mainly behind the idea of counting, membership and infinity. The notion of infinity was one of the most important foci and challenges for the Greek scientists at that time. It was with Cantor's work however, that set theory came to be put on a proper mathematical basis. Cantor stated: "a set is a collection into a whole of definite and separate objects of our intuition or thought" [Dauwben, 1979; Johnson, 1972]. However, to remove paradoxes in the set axiomatization conducted by Russell and Whitehead, Russell added: "... such that given an object *x* and a set *S*, it is possible to state unequivocally that *x* either *belongs* to the set *S*, or *x* does *not belong* to *S*" [Garciadiego, 1992]. *However, since many years the use of set theory had focused more on objects than on relations between objects.* In other words, the notion of structure[1] formed by the objects (or elements) of a set has always been lacking. In fact, no inter-relationships between objects are emphasized in many applications of set theory. Cantor's definition requires the objects to be separate *but did not exclude inter-relationships between the objects.* In addition, the Russel and Whitehead axiomatization of set theory did remove paradoxes but limited significantly the interaction of an object with the set that contains it. Fuzzy set theory developed more recently by Zadeh extended the notion of membership of an element within a set [Zadeh, 1962] but suffers similarly from the inability to capture structural information. Fuzzy set

---

[1] A structure refers to a certain arrangement of objects in a definite pattern of organization.

theory does not account for structural information between the elements that form the fuzzy set. The "subjective" membership functions introduced in fuzzy set theory do not unfold the hidden structure formed by the objects themselves. Ideally researchers would like to have a solution to a pattern recognition (PR) problem that consists of the following stages: (i) find a feature vector x, (ii) train a system using a set of training patterns whose classification is *a priori* known, and (iii) classify unknown incoming patterns. Unfortunately, for most practical problems, this approach is not feasible. The reason is that a pattern contains some *relational information* from which it is difficult and sometimes impossible to derive appropriate feature vector. *Therefore, the analytical approaches which process the patterns only on a quantitative basis but ignore the inter-relationships (or structure) between the components of the patterns quite often fail.* Several approaches have been attempted to circumvent this serious obstacle. The first alternative that accounts for structural information is syntactical pattern recognition [Fu, 1982]. This concept establishes a connection between formal languages and pattern recognition, but it is quite difficult to represent continuous and complex patterns using a finite set of rules. A second alternative is the Bayesian belief networks [Buntine, 1996; Pearl, 1988] that use conditional probabilities and priors to express causal and sometimes non causal relationships. However, the structural information we are considering is more general and cannot always be represented through a Bayesian network.

The thrust in this paper is to develop a new mathematical domain that the author named *"Functional Set Theory" (FST)* which answers this challenge. We show that object of the traditional set theory can be enriched by adding its connections (or mappings) with other objects within the same set. Each mapping is assigned to a *role* "played" in a functional set (f-set). This role introduces the notion of *fuzzy membership* assigned to each object of the f-set. The roles impact the membership values of the objects. Because of the existence of the mathematical properties of the mappings between objects and their roles, FST will unravel the interaction between the traditional use of set theory and fuzzy set theory. An example of impact is in artificial intelligence and pattern recognition [Duda, Hart and Stork, 2001] where the hidden Markov models will be transformed into structural hidden Markov models (SHMM's). These models introduced by the author and his Ph.D student has been recently established within the computational intelligence and pattern recognition community [Bouchaffra and Tan, 2004, 2003[1], 2003[2]]. The *structural* hidden Markov models help us better understand the shape of a sequence of patterns in a traditional HMM. We believe this new vision strengthened by "everyday" more powerful computers will leapfrog the current state-of-the-art models.

2. INTRODUCTION TO FUNCTIONAL SET THEORY AND ITS PRINCIPLES

Our goal is to develop a new mathematical theory that incorporates structural information within the f-set axiomatic. Our goal is to develop a definition of a structure and devise ways to learn it. Our development of functional set theory is driven by the followings: (i) the need to create structural relationships between the objects that constitute the f-set. Instead of just being able to relate each element of a set to the set itself, we will be able to characterize and understand the type of structural relationships that exist between the elements themselves. (ii) the need of capturing imprecision, or fuzziness within a set through a non crisp membership operator. (iii) the need of viewing structures at different levels of resolution. Therefore, objects in a set are not necessarily atomic but are either atomic or are recursively f-sets. We now define an f-set (a functional set):

Definition 1: *Let E be a ground set of objects a_i of cardinality n, an f-set Ω is a finite collection of functions $f_{i;j}$: $S_i \rightarrow S_j$, where S_i and $S_j \in P(E)$ (power set of E) and whose element membership values in the f-set Ω depend on the functions $f_{i;j}$.*

A possible representation of an f-set is: $\Omega = \{f_{1;2}\langle S_1;S_2\rangle, f_{3;4}\langle S_3;S_4\rangle, \ldots, f_{u;v}\langle S_u;S_v\rangle\}$ where: $f_{i;j} \in F(S_i; S_j)$ (set of all mappings from the fuzzy set S_i to the fuzzy set S_j).

2.1 Some Examples of F-Sets

The first example of an f-set is the following: Let $E = \{1, 2, 4\}$ then $P(E) = \{\varnothing, \{1\}, \{2\}, \{4\}, \{1, 2\}, \{1, 4\}, \{2, 4\}, E\}$, a possible f-set Ω is formed by mappings from one member of $P(E)$ to another member of $P(E)$. A possible mapping is $f_1 : \{1, 2\} \rightarrow \{1, 4\}$ such that $\forall n \in \{1, 2\}, f_1(n) = n^2 \in \{1, 4\}$. Another mapping is $f_2 : \{1, 2\} \rightarrow \{2, 4\}$ such that: $\forall n \in \{1, 2\}, f_2(n) = 2n \in \{2, 4\}$. In case there are no possible mappings between members of $P(E)$, identity functions that map each member to itself are used. The second example describes an f-set where the objects are non mathematical and the mappings are graphical representations rather than mathematical functions: Let E = {fax/scanner1, fax/scanner2,computer1,computer2,cell-phone1,cell-phone2,rugby-ball} be the ground set. Some members of $P(E)$ are: S_1 = {fax/scanner1,fax/scanner}, S_2 = {computer1,computer2}, S_3 ={cell-phone1,cell-phone2}, and S_4 = {rugby-ball}, then possible elements of this f-set are: $f_{1;2} = Plug\langle S_1; S_2\rangle$: "plug a fax/scanner to a computer", $f_{2;1} = Transmit\langle S_2; S_1\rangle$: "transmit information from a computer to a fax/scanner", $f_{3;2} = Connect\langle S_3; S_2\rangle$: "connect a cellular phone to a computer using a wireless communication protocol" and $f_{2;3} = Redirect\langle S_2; S_3\rangle$: "redirect information from a computer server to a cell-phone". The isolated object "rugby-ball" is mapped to itself. We can write: $f_4 = Id\langle S_4\rangle$. Figure 1 provides an illustration of this f-set.

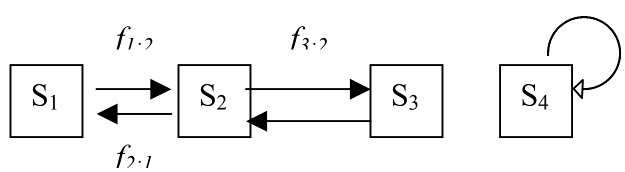

Figure 1: A graphical representation of an f-set.

The third example is in handwriting recognition. Let's consider E as a ground set that contains images of digits from 0-9 (handwritten by different persons) with their feature representations using various feature extraction techniques. This ground set contains different ways of writing a digit. For example, some zeros are round, while others may be elliptical, flat, or some other shape. A possible collection of mappings between members of $P(E)$ are the different feature extraction techniques that are applied on these images. These mappings are for example: a polynomial representation of the external contour, which attempts to approximate each stroke of the external contour of the digit by a polynomial expression, or a vector representation such as Fourier descriptors or a chain code representation. All these functions map a digit image to a digit representation in a feature space. Similarly, other mappings from the feature space to itself are obtained by mapping a digit feature representation to one of the ten template feature representations. These mappings are classification techniques such as "Neural Network", "Genetic Algorithms", and others. The f-set Ω is a *pattern recognition system* composed of feature extraction functions, and pattern classification functions.

2.2 Roles of Mappings and Object Memberships

The notion of mapping between two members of *P*(*E*) is crucial in FST. The "role" the mappings play provides the most relevant information about the inner structure contained in an f-set. However, a role is different from a membership. Memberships apply to objects with respect to an f-set whereas roles apply to mappings. *If the role value assigned to a mapping between two members of P*(*E*) *is high then the membership values of these objects in the f-set are also high and vice versa. In other words, unlike fuzzy set theory, functional set theory focuses on the roles played by the functions and not only on the objects themselves.* These membership values *are derived from these role values*. Because the role value assigned to the identity mapping on a member of *P*(*E*) is very low, therefore the objects contained in this fuzzy set have very low membership values to belong to the whole f-set. Conversely, the objects of fuzzy sets that are sources or targets of several mappings with *highest roles* have high membership values to belong to the f-set. These latter fuzzy sets contribute prominently to the global structure of the f-set. We can express physical structures such as the "statue of liberty", logical such as "a particular combination of words makes a sentence" or temporal structures such as "a sequence of objects ordered by time. For example, any group of organs of a human body considered as a fuzzy set has a membership value different from 0. The vital biological functions[2] between the heart and the brain considered both as fuzzy sets will provide these two organs with high membership values to belong to the whole human body f-set. Conversely, the finger and the arm will be assigned low membership values because the biological functions between these two organs are not vital. If the mappings are dynamic functions with respect to time, then using functional set theory, we will be able to predict the parts of cybernetic systems (including species) that will change or even disappear over time because of a change in their environment.

Definition 2: *The membership of an object a_i in the f-set Ω is determined by the roles of the mappings that relate two fuzzy sets that contain a_i (as a source or as a target). We write $\mu(a_i \in \Omega) = \alpha$ to state that the membership value of the object a_i in Ω is equal to α.*

2.3 Functional-Set Theoretic Operations

For now, we will introduce only the complement and the intersection for now. Given an f-set Ω and an f-subset $A \subseteq \Omega$, the complement of A in Ω is the f-subset B of all mappings that belong to Ω but not to A. The intersection of two f-sets A and B is the f-set C containing those mappings in both A and B. If: $A = \{f_1\langle S_1;\ S_2\rangle,\ f_2\langle S_3;\ S_4\rangle\}$ and $B = \{f_1\langle S_1;\ S_2\rangle\}$, then the intersection $C = A \wedge B = \{f_1\langle S_1;\ S_2\rangle\}$ with the following membership values: $\mu_{A \wedge B}(a_i) = \min\{\mu_A(a_i);\ \mu_B(a_i)\}\ \forall\ a_i \in S_i$.

2.4 The Notion of Functional Class

Our goal in this section is to extract well-organized patterns that are parts of a whole entity. This is done through the definition of "a functional class" (f-class).

Definition 3: *An f-class is an f-subset of an f-set Ω that involves only members of P*(*E*) *that are related through at least one mapping. That is, all non-related members (or fuzzy sets) are dropped. Figure 1 depicts one f-class that involves $\{S_1, S_2, S_3\}$ and another f-class that involves $\{S_4\}$.*

---

[2] We mean very complex biochemical interactions between the two organs.

Definition 4: *An f-structure of order β ($\beta \in \Re$) is an f-class that is represented by the mappings between members of P(E) with role values $\geq \beta$. The lists of mappings that have* the highest roles *are called* the invariants *(or descriptors) of the f-structure.*

"A human face" can be viewed as an f-structure since it is made of "a nose", "a mouth", "two eyes", etc., all viewed as f-substructures. The structural invariants are *particular qualities or properties* that mappings between fuzzy subsets of objects from the same f-class share. *They represent a particular class of functions with high roles*. The structural invariants capture the core of the structure. For example, these invariants might characterize features pertaining to the ancestors of species and are found in a higher level in a taxonomy tree of species. They are key answers to the questions: "what makes a hawk close to an eagle?" and "what makes a tiger close to a lion?"

3. LEARNING WITHIN FUNCTIONAL SET THEORY

The learning process within the FST paradigm consists roughly of (i) determining whether components or groups of objects are forming different whole entities (or f-classes) or whether they form a single entity, and (ii) discovering the parts of the entity that are most "important" in the sense they have higher fuzzy membership values in the f-class. These important parts constitute *the structure*, they are determined by the *invariants* of the f-class. The learning algorithm starts with random mappings that relate members of *P*(*E*) where *E* is a ground set. As the system evolves, it gets more relevant information, by adding, modifying or removing some mappings between members of *P*(*E*). The mappings are improved with respect to a criterion. At this stage some f-classes have emerged from the initial configuration (the initial cloud). Once the f-classes are identified, the algorithm will focus on finding the structure of each f-class. Through these mapping updates, the roles are being increased. In a supervised mode, the expert will provide better mappings within an iteration of the learning algorithm in order to enhance their roles. The process will continue iteratively and stop after a predefined learning threshold is reached. At this stage, the algorithm has identified the "optimal mappings" called "invariants" as well as the "optimal fuzzy membership values" of the objects in the f-class. *Therefore, the structure has emerged and the f-set algorithm has learned.* Reasoning under FST is called *functional reasoning* (FR) by the author. It is conducted in order to extract the different f-classes contained in a functional set. The power set *P*(*E*), enables view the cloud of data from different views in order to detect eventual "best connections" between these views. The idea is similar to the puzzle game principle. By arranging the puzzle pieces in a certain way, a larger well-organized pattern will be unraveled.

4. IMPACTS OF F-SET THEORY

There are many areas on which f-set theory has an impact. Some of them are:

Logic: Functional set theory brings a new vision of logic that enables human to better communicate with machines. The logical formalism of FST is more appropriate to human reasoning than the binary logical formalism. In a long-term, we will be able to design a new type of chips based on FST that will open a door to a new generation of more intelligent machines that better understand humans.

Fuzzy Set Theory: The incorporation of FST will extend fuzzy set theory that also does not account for structural information. *There are no relations or mappings between objects in the fuzzy set theory.* FST introduces the notion of "role" of a mapping that impacts the degree of membership of objects that are being mapped in the f-set.

Hidden Markov Models: Traditional HMM's *have a clear conceptual framework and the ability to learn statistically, but they are unable to account for structural information of the sequence.* Because the symbols of an input sequence are assumed to be conditionally independent, therefore HMM's make no use of structure, either topological or conceptual. Structural hidden Markov models (SHMM's) developed by the author and his student extend HMM's by searching this underlying structure. *A structural HMM is then composed of several f-structures that are connected through mappings to form a bigger f-structure.* For example, if $E = \{o_1, o_2, \ldots, o_k, c_1, c_2, \ldots, c_p\}$ is the ground set, an f-structure might be composed of mappings between some members S_i of $P(E)$ (group of evidences o_i) and some conclusions c_k. SHMM's merge statistics with syntax in a natural and seamless way. We have applied the concept of SHMM to help design engineers map external car designs to customer's preferences. SHMM has outperformed the classical HMM with 10**%** of difference in prediction accuracy [Bouchaffra, and Jun, 2004].

5. CONCLUSION

We have introduced in this paper a new mathematical domain that we called f-set theory. Our goal is to emphasize the relationships between objects rather than the objects themselves. We believe that this "functional approach" is a must for a better understanding of the notion of structure. We have introduced the notion of role assigned to mappings between fuzzy sets as well as the notion of fuzzy membership value assigned to an object in an f-set. We introduced the notion of f-class from which we derived the notion of a structure. Our intuition is driven by the fact that mappings provide more information about structures than the objects alone. This research is promising since it impacts so many areas such as mathematics, computer science, and biology.

REFERENCES

D. Bouchaffra and J. Tan "Introduction to the Concept of Structural Hidden Markov Model: Application to Mining Customers' Preferences for Automotive Designs", in Proceedings of the 17th International Conference on Pattern Recognition (ICPR), Cambridge, UK, 23-26-07-2004.

D. Bouchaffra and J. Tan, "Structural HMM modeling and its Applications in Automotive Industry", in: Proceedings of the 5th International Conference on Enterprise Information Systems, Angers, France 23-26, 2003[1].

D. Bouchaffra and J. Tan, "Mapping Designs to User Perceptions using a Structural HMM: Application to Kansei-Engineering", in: Proceedings of the International Conference on Computational Intelligence for Modeling, Control and Automation - CIMCA'2003, Vienna, Austria, 12-14 February 2003[2].

W. Buntine, "A Guide to the Literature on Learning Probabilistic Networks from Data", in: IEEE Transactions on Knowledge and Data Engineering, 8(2): 195-210, 1996.

J. W. Dauwben, "Georg Cantor's Creation of Transfinite Set Theory: Personality and Psychology in the History of Mathematics", Papers in Mathematics, New York, 1979.

R. Duda, P. Hart, and D. Stork, "Pattern Classification", Wiley, New York, 2001.

K.S. Fu, "Syntactic Pattern Recognition and Applications", Prentice-Hall, Englewood Cliffs, N.J., 1982.

A. Garciadiego, "Bertrand Russell and the Origins of the Set Theory Paradoxes", Basel 1992.

P.E. Johnson, "A History of Set Theory" (Boston, Mass., 1972).

J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Morgan Kaufmann, San Mateo, CA, 1988.

L. Zadeh, "Fuzzy Sets", in: Information and Control, 8(3), 338-53, 1962.