Journal of Mathematical Chemistry 22 (1997) 143–147

Fuzzy sets and Boolean tagged sets

Ramon Carbó-Dorca

Institute of Computational Chemistry, University of Girona, Girona, Spain

Received 10 June 1997

The *fuzzy set* structure is analysed from the point of view of a new definition: *Boolean tagged sets*, which are constructed as a straightforward generalisation of the fuzzy set concept, but fully adapted to computational purposes. Boolean tagged sets are simply structured as sets whose members are *tagged* by a binary string, attached in turn to the vertices of a unit hypercube. Boolean tagged sets are so designed as to become useful in any field of applied mathematics, but they can also be seen as obviously prepared for chemical applications, associated to molecular information gathering and manipulation.

1. Introduction

Since Zadeh's definition of the fuzzy set concept [5], there has been a growing amount of literature dealing with the theory and applications of this interesting generalisation. Fuzzy sets are perfect tools to deal with the current everyday life situations, where no short cut between logical, {true, false}, sentence contents is present.

The aim of the present paper is to describe an alternative general way to deal with the same problems as fuzzy set theory does, while presenting an immediate connection with actual information structure and manipulation. The fact is that, owing to the current technological situation, information is gathered and manipulated, manufactured in practice, within a pattern, which can be reduced to the binary equivalent of the logical contents, $\{1, 0\}$, mentioned before.

Due to this situation, electromechanical devices, known as computers, deal with sequences of information units, bit strings, which in turn can be easily be seen as the set of vertices of unit n-dimensional cubes or hypercubes, with the origin placed at the vertex constructed with a $\{0\}$ bit string.

More than this, when one is thinking about any kind of *information* structure, it can be easily realised, with a simple imagination effort, how this information can be transported into a vertex of the n-dimensional cube.

Then, as a set of obvious examples, one can describe, using the vertex bit string of some appropriate dimension unit hypercube: the CPU contents of a given computer in a given time slice, or the information structure of a hard disk, or the contents of a stable CD-ROM, etc.

A not so typical example may correspond to the books of a library translated into binary code. Using a completely similar situation, the set of all molecules reported within chemical abstracts, with the known information on them attached forming a bit string, constitutes another obvious set, with the appropriate characteristics such as those that will be described here.

In order to describe this kind of set structure, this paper is organised in the following manner: first some definitions are given, then operations and vector spaces of this sort are described, and finally some applications are given.

2. Definitions

A *unit length n-dimensional cube or hypercube* is a set of 2^n vertices, whose elements can be constructed with the binary string vectors, associated to the integer sequence $\{0, 1, 2, \ldots, 2^n - 1\}$. Two vertices should be noted in the sequence above, namely: $\mathbf{0} = (0, 0, 0, \ldots, 0)$, the *zero vertex,* and $\mathbf{1} = (1, 1, 1, \ldots, 1)$, the *unit vertex*. Let us use K_n to denote such an n-dimensional polyhedron.

Hypercubes of this sort can be constructed iteratively as $K_{n+1} \leftarrow K_n$, using the algorithm

$$\forall v_k \in K_n\colon\ v'_{2k} = v_k \oplus (0) \wedge v'_{2k+1} = v_k \oplus (1) \Rightarrow \left\{v'_{2k}, v'_{2k+1}\right\} \in K_{n+1}.$$

Suppose any set, $S = \{s\}$, the *background set*, and a hypercube K_n. A *Boolean tagged set*, B_n, is constructed by means of the ordered pairs:

$$B_n(S) = S \times K_n = \left\{z \in B_n \mid z = [s, v_k];\ s \in S \wedge v_k \in K_n\right\}.$$

By the *order* of a Boolean tagged set will be understood the dimension, n, of the underlying hypercube, K_n. In order to simplify the notation, the equivalent symbols: $B_n \equiv B_n(S)$, will be used whenever no confusion could be generated. A *classical set* can be associated to a Boolean tagged set of unit order B_1.

3. Tagged classes

It is apparent that the *Boolean tag*, associated to every element of the set S, placed at the second moiety of the ordered pairs in Boolean tagged sets, can be understood as a *membership tag class*. A trivial example is given by the unit order Boolean tagged set, where the following interpretation holds:

$$\forall z \in B_1 \rightarrow \left\{z = [s, 1] \wedge s \in S\right\} \vee \left\{z = [s, 0] \wedge s \notin S\right\}.$$

The integer value, k, of a given Boolean tag, v_k, is defined by the symbol $\langle v_k \rangle = k$, so one will have, for example, $\langle \mathbf{0} \rangle = 0$ and $\langle \mathbf{1} \rangle = 2^n - 1$.

In general, now, and following the same path as in fuzzy set theory, observing any Boolean tagged set, B_n, only background elements bearing the unit vertex Boolean tag can be considered *true members* of the *background set* S. Likewise, only background elements bearing the zero vertex Boolean tag can be considered *true non-members* of S. The rest of the background elements may correspond to the *marginal elements* of B_n.

A *tagged class* C_k of a Boolean tagged set B_n is defined by means of the following condition:

$$\langle v_k \rangle = k \rightarrow \forall x = [t, v_k] \in C_k \subset B_n.$$

A tagged class will be attached to a subset T_k of the background set S, such that $\forall t \in T_k \subset S \rightarrow C_k = T_k \times \{v_k\}$.

The following relationships hold:

$$\begin{aligned} &\bigcup_k C_k = B_n \wedge C_p \cap C_q = \emptyset, \\ &\bigcup_k T_k = S \wedge T_p \cap T_q = \emptyset, \end{aligned} \tag{1}$$

that is, $\forall s \in S$: $\exists v_k \in K_n \rightarrow z = [s, v_k] \in C_k \subset B_n$. Thus Boolean tagged sets are defined in such a way that every element is a member of some disjoint subset, associated to a given hypercube vertex also representing an integer value in the range $k \in \{0, 1, 2, \ldots, 2^n - 1\}$.

A *non-degenerate Boolean tagged set* has the available 2^n classes with one or nil attached elements of the background set. On the other hand, degenerate Boolean tagged sets would give to the observer the meaning of lack of information on the tag moiety. While the presence of void tagged classes could signal that possible candidates of the background set remain still unknown.

4. Operations over Boolean tagged sets

Any operation defined over the elements of B_n has to be decomposed into two distinct operations, involving the *background* and the *tag* parts separately. That is, a general situation may be depicted in the following way:

$$\forall a, b \in B_n\colon\ a * b \rightarrow [\alpha, \mu] * [\beta, \nu] \rightarrow [\alpha \circ \beta, \mu \bullet \nu] \in B_m. \tag{2}$$

At the same time, transformations of the elements of B_n can be defined in three possible different manners, according to how the transformation affects the Boolean tagged set elements. There may be, thus, transformations of the background part only, or the tag part could be the uniquely affected or both parts may be changed. *Background*, *tag* or *total transformations* may be defined accordingly.

5. Boolean tagged vector spaces

A *Boolean tagged vector space* is simply a Boolean tagged set whose background set is a vector space. The structure of the vector space is preserved provided that the appropriate operations could be defined in the tag moiety. For the vector sum of the background part, there must be defined a corresponding composition on the tag. That is, if W_n corresponds to such a Boolean tagged vector space, then

$$\forall a, b \in W_n\colon\ a + b = [\alpha, \mu] + [\beta, \nu] = [\alpha + \beta, \mu \bullet \nu] \in W_n. \qquad (3)$$

In this way, choosing the tag part operation as $\bullet \equiv \wedge$, true members are preserved as such in the final sum result, and when summed up to true non-members the result is a true non-member. Sums performed over a tagged class of marginal set members preserve the tagged class association of the result. Within this trend, marginal members not of the same tagged class, when summed, yield true non-members.

When taking into account the product of a vector by a scalar, it is sufficient to consider this operation as a background set transformation, leaving the tag part invariant. As a consequence, linear combinations of Boolean tagged vector spaces are perfectly defined in this way.

Metric background vector spaces may merit some attention, because scalar products or norms will produce a scalar result in the background part. Thus, in the tag part some sound transformation must be defined too, providing the projection of the whole Boolean tagged space into a Boolean tagged unit order set with a scalar background set. Then, metric Boolean tagged vector spaces can easily be defined.

Distances can be also defined quite immediately over Boolean tagged vector spaces. The background part is immediate and can be obtained using some common form, as in the corresponding space and in the tag part a Minkowski formula may be used, which amounts to the same as counting the number of non-coincident bits in the pair of implied tags. The biggest distance number obtained in any usual way will be the one involving the extreme vectors **0** and **1**.

6. Applications

Obvious applications may be envisaged in the domain of chemistry, as well as in other areas, like Boolean tagged logic, which do not directly concern this journal.

The set S may be taken as a set whose elements are made of molecular structures. The hypercube tag elements can be identified to a set of chosen molecular descriptors, transformed into a corresponding set of bit strings. In doing so, the molecular classes may be easily compared and ordered.

Proceeding in this manner, it can be straightforwardly deduced how the Boolean tagged sets are to be considered in the backyard of all the attempts and procedures to classify molecules. Molecular classification and order can be achieved by means of studies essentially based on molecular properties. The molecular magnitudes, obtained through theoretical considerations or experimental measures, being *rational numbers* in the worst case, can be easily transformed into bit strings. A molecular set can be associated to Boolean tag descriptors, and transformed in this way into a Boolean tagged set. The Atomic Periodic Table is nothing but an early nice example of this situation.

A field where the author has been active in the last years, the realm of *quantum similarity measures* [1,3], provides us with a perfect example of the construction of a Boolean tagged molecular set. It is well known that molecular quantum similarity measures can produce a discrete representation of molecular structures within a given molecular set: the so-called *molecular point cloud* [2] of the set. Each molecule, by using this procedure and described in the molecular point cloud, is represented by a vector, a *point-molecule*, obtained by purely computational means. Due to this reason, the quantum similarity measure's molecular discrete description can be easily transformed into a bit string. A molecular point cloud is, from this point of view, a representative example of a potential Boolean tagged set. It is not by accident that in early examples of molecular point cloud representations in form of graphs, n-dimensional hypercubes [4] of the appropriate number of vertices were successfully used.

Acknowledgements

This work has been partially financed by a CICYT Research Project: #SAF 96-0158. Thanks are due to Profs. J. Karwowski and P.G. Mezey, for lively discussions on the subject, as well as thanks are due to Dr. E. Besalú for constructive criticism and advice.

References

[1] R. Carbó, ed., *Molecular Similarity and Reactivity: From Quantum Chemical to Phenomenological Approaches* (Kluwer, Dordrecht, 1995).
[2] R. Carbó, B. Calabuig, L. Vera and E. Besalú, Adv. Quantum Chem. 25 (1994) 253.
[3] R. Carbó-Dorca and P.G. Mezey, eds., *Advances in Molecular Similarity*, Vol. 1 (JAI Press, London, 1996).
[4] R. Carbó and B. Calabuig, Int. J. Quantum Chem. 42 (1992) 1695.
[5] L.A. Zadeh, Inform. Control 8 (1965) 338.