Different Generalizations of Bags

Sadaaki Miyamoto

Abstract We consider four different generalizations of bags (alias multisets). We first discuss Yager's fuzzy bags having different sets of operations. It is shown that one is not a generalization of fuzzy sets but a mapping of them into fuzzy bags, since operations are inconsistent between the two, while the other includes fuzzy sets as particular cases. Third type is called real-valued bags which is simpler than the former two and is a kind of the reduction of fuzzy bags. Finally, the fourth generalization called G-bags includes all three except the first type. It is a minimal extension of the second and the third generalizations. Bag relations are defined for the third type of real-valued bags, which can further be generalized for G-bags.

Keywords bags · fuzzy bags · t-norms · bag relations

1 Introduction

Bags alias multisets have long been discussed by computer scientists as a basic data structure not only theoretically [4,7,1] but also as a framework in many applications such as information retrieval.

Fuzzy bags have also been studied by a number of researchers [22,23,5,6,18–20,8–12,14,21]. Moreover, real-valued bags [15] and generalized bags [13] have been proposed. We thus find four different proposals to extend bags. First, Yager proposed fuzzy bags should be crisp bags of $X \times [0,1]$ where X is a basis set. Second, Miyamoto (e.g., [8]) proposed another set of basic relations and operations for fuzzy bags. Third, real-valued bags have also been proposed [15]. Lastly, a minimal extension that includes the second type of fuzzy bags and real-valued bags is given as generalized bags [13].

This paper gives an overview of these studies with the purpose to make clear where the essential differences are in these different proposals. Among these proposals, three are *generalizing* bags and fuzzy sets in the sense that basic set operations are compatible

Sadaaki Miyamoto
Department of Risk Engineering, University of Tsukuba, Ibaraki 305-8573, Japan
Tel.: +81-29-853-5346
Fax: +81-29-853-5809
E-mail: miyamoto@risk.tsukuba.ac.jp

between the original and generalized frameworks. In contrast, the other fuzzy bags have incompatibility of operations between fuzzy bags and fuzzy sets, and hence it is not a generalization, but a mapping of fuzzy sets into fuzzy bags.

No proposals are wrong among these four, but we note real-valued bag is simpler and convenient when we define bag relations, and the last generalization called G-bags is a minimal extension of the others.

The rest of this paper is organized as follows. Section 2 presents two kinds of fuzzy bags. Section 3 discusses real-valued bags, and G-bags are considered in Section 4. Bag relations are discussed in Section 5. Section 6 briefly mentions application possibilities. Finally, Section 7 gives conclusions.

2 Bags and Fuzzy Bags

Throughout this paper a basis set of objects is denoted by X. Unless stated otherwise, X is assumed to be a finite set for simplicity.

2.1 Crisp bags

A (crisp) bag M of X is characterized by a function $C_M(\cdot)$ which is called the count function of M, whereby a natural number including zero corresponds to each $x \in X$: $C_M \colon X \to \{0, 1, 2, \dots\}$.

When $X = \{x_1, \dots, x_n\}$, we may express a crisp bag as

$$M = \{k_1/x_1, \dots, k_n/x_n\}$$

or

$$M = \{\overbrace{x_1, \dots, x_1}^{k_1}, \dots, \overbrace{x_n, \dots, x_n}^{k_n}\}.$$

In this way, an element of X may appear more than once in a bag.

Example 1 Let $X = \{a, b, c, d\}$ and let $C_M(a) = 2$, $C_M(b) = 3$, $C_M(c) = 1$, $C_M(d) = 0$. We can write $M = \{a, a, b, b, b, c\} = \{2/a, 3/b, 1/c\}$ by ignoring zero occurrence of d.

The two expressions: $M = \{k_1/x_1, \dots, k_n/x_n\}$ and $M = \{\overbrace{x_1, \dots, x_1}^{k_1}, \dots, \overbrace{x_n, \dots, x_n}^{k_n}\}$ will lead to different generalizations of crisp bags.

Why bags are important in a variety of information processing can be illustrated using a simple mapping and linear sequential processing.

Example 2 Suppose $X = \{x_1, x_2, \dots, y_1, y_2, \dots, z_1, z_2 \dots\}$ and $Y = \{x, y, z\}$. Consider a mapping $f \colon X \to Y$: $f(x_i) = x$, $f(y_j) = y$, $f(z_k) = z$.

Suppose a sequence $(x_1, x_2, y_1, z_2, y_2, z_1, x_3)$ comes as an input to f and then the output is (x, x, y, z, y, z, x):

$$x_1, x_2, y_1, z_2, y_2, z_1, x_3 \to \boxed{f} \to x, x, y, z, y, z, x$$

We now regard the input sequence as a set, i.e., we do not care its order. Thus, the input is $\{x_1, x_2, x_3, y_1, y_2, z_1, z_2\}$. Then the output is not a set, but a bag $\{x, x, x, y, y, z, z\} = \{3/x, 2/y, 2/z\}$.

Such 'don't care about the order' typically arises in relational databases. Indeed, we have bags of tuples as the result of a SELECT operation.

Basic operations for crisp bags: The followings are basic relations and operations for crisp bags.

1. (inclusion): $M \subseteq N \Leftrightarrow C_M(x) \leq C_N(x), \quad \forall x \in X$.
2. (equality): $M = N \Leftrightarrow C_M(x) = C_N(x), \quad \forall x \in X$.
3. (union): $C_{M \cup N}(x) = \max\{C_M(x), C_N(x)\}$.
4. (intersection): $C_{M \cap N}(x) = \min\{C_M(x), C_N(x)\}$.
5. (addition or sum): $C_{M+N}(x) = C_M(x) + C_N(x)$.
6. (scalar multiplication): $C_{\alpha M} = \alpha C_M(x)$, where α is a nonnegative integer.
7. (Cartesian product): Let P is a bag of Y. $C_{M \times P}(x, y) = C_M(x) C_P(y)$.

We use $\vee$ and $\wedge$ for max and min, respectively. Note that the above relations and operations are similar to those for fuzzy sets. However, bags have the addition operation that fuzzy sets do not have, and the Cartesian product for bags is different from that for fuzzy sets.

Example 3 Suppose X, Y, and $f\colon X \to Y$ are the same as those in Example 2. Let $A = \{x_1, x_2, x_3, y_1, y_2, z_1, z_2\}$ and $M = \{x, x, x, y, y, z, z\}$. Then we can express

$$M = \{f(x_1)\} + \{f(x_2)\} + \cdots + \{f(z_2)\} = \sum_{x \in A} \{f(x)\}$$

using the addition. Generally, functional image of a set using the addition results in a bag. In contrast, the ordinary image of a set is expressed by the union:

$$f(A) = \bigcup_{x \in A} \{f(x)\}.$$

2.2 Fuzzy bags

Let us consider a *fuzzy* variation of Example 2.

Example 4 Suppose X, Y, and f are given in Example 2. Consider a 'fuzzy' sequence $(0.1/x_1, 0.5/x_2, 0.7/y_1, 0.9/z_2, 0.7/y_2, 0.2/z_1, 0.5/x_3)$ as an input to f. The output is then $(0.1/x, 0.5/x, 0.7/y, 0.9/z, 0.7/y, 0.2/z, 0.5/x)$. If the input sequence is regarded as a fuzzy set $\{0.1/x_1, 0.5/x_2, 0.5/x_3, 0.7/y_1, 0.7/y_2, 0.2/z_1, 0.9/z_2\}$, i.e., we don't care about its order, the output is $\{0.1/x, 0.5/x, 0.5/x, 0.7/y, 0.7/y, 0.2/z, 0.9/z\}$.

Yager defined a fuzzy bag to be a crisp bag of $X \times [0, 1]$. By this definition, $0.1/x$ in the above output is written as $(x, 0.1)$ and we write

$$F = \{0.1/x, 0.5/x, 0.5/x, 0.7/y, 0.7/y, 0.2/z, 0.9/z\}$$
$$= \{1/(x, 0.1), 2/(x, 0.5), 2/(y, 0.7), 1/(z, 0.2), 1/(z, 0.9)\}.$$

We thus have $C_F((x, 0.1)) = 1$, $C_F((x, 0.5)) = 2$, and so on.

Although this definition is simple and useful, we have a problem when generalizing bags, which leads to different definitions.

Example 5 Suppose X, Y, and f as in the last example. Let two fuzzy set be $A = \{0.5/x_1\}$ and $B = \{0.7/x_1\}$ which satisfies $A \subset B$. The output of f is $\tilde{A} = \{1/(x, 0.5)\}$ and $\tilde{B} = \{1/(x, 0.7)\}$. Obviously $\tilde{A} \subset \tilde{B}$ does not hold.

Thus, fuzzy sets appear to be generalized into the collection of fuzzy bags, but there is fundamental inconsistency between fuzzy set operations and fuzzy bag operations.

This means that, actually, fuzzy bags above defined do not *generalize* fuzzy sets, or in other words, fuzzy sets are *mapped* into the collection of fuzzy bags, and there is no consistency between the original space of fuzzy sets and the subspace of 'fuzzy sets' after mapping.

Here is a question: is there any other way that has consistency between basic operations of fuzzy sets and those of fuzzy bags. The answer is affirmative by introducing another set of relations and operations as below.

2.3 Fuzzy bags as generalization of fuzzy sets

Miyamoto redefined fuzzy bags [8–12, 14]. Basically, fuzzy bags have the form of crisp bags of $X \times [0, 1]$ as Yager defined, but with a different set of basic relations and operations. Generally, we do not write $\{1/(x, 0.1), 2/(x, 0.5), 2/(y, 0.7), 1/(z, 0.2), 1/(z, 0.9)\}$ but we write

$$F = \{(x, 0.1), (x, 0.5), (x, 0.5), (y, 0.7), (y, 0.7), (z, 0.2), (z, 0.9)\}$$
$$= \{\{0.1, 0.5, 0.5\}/x, \{0.7, 0.7\}/y, \{0.2, 0.9\}/z\}$$

Generally, we define a fuzzy bag is characterized by a bag of $[0, 1]$, i.e., for each $x \in X$, we have a bag $\{\mu, \mu', \ldots, \mu''\}$ of $[0, 1]$. Moreover the bag is arranged into a sequence of numbers of the decreasing order, e.g.,

$$F = \{\{0.1, 0.5, 0.5\}/x, \{0.7, 0.7\}/y, \{0.2, 0.9\}/z\}$$
$$= \{(0.5, 0.5, 0.1)/x, (0.7, 0.7)/y, (0.9, 0.2)/z\}$$

We thus assume that a fuzzy bag is characterized by such a sequence of decreasing order, called *a membership sequence*, denoted by

$$(\mu_A^1(x), \mu_A^2(x), \ldots, \mu_A^p(x))$$

where $\mu_A^1(x) \geq \mu_A^2(x) \geq \cdots \geq \mu_A^p(x)$. In other words, we have

$$C_A(x) = (\mu_A^1(x), \mu_A^2(x), \ldots, \mu_A^p(x)),$$

where p generally depends on x, but p can be taken as a constant for all x by appending zeros at the tail of the sequence, when we handle a finite X and a finite number of fuzzy bags.

We now can define the basic relation of inclusion and operations for fuzzy bags.

1. **inclusion:**
   $A \subseteq B \Leftrightarrow \mu_A^j(x) \leq \mu_B^j(x), \ j = 1, \ldots, p, \ \ \forall x \in X.$
2. **equality:**
   $A = B \Leftrightarrow \mu_A^j(x) = \mu_B^j(x), \ j = 1, \ldots, p, \ \ \forall x \in X.$
3. **addition:**
   $A + B$ is defined by the addition operation in $X \times [0, 1]$ for crisp bags [22].
4. **union:**
   $\mu_{A \cup B}^j(x) = \max\{\mu_A^j(x), \mu_B^j(x)\} = \mu_A^j(x) \vee \mu_B^j(x), \ j = 1, \ldots, p.$
5. **intersection:**
   $\mu_{A \cap B}^j(x) = \min\{\mu_A^j(x), \mu_B^j(x)\} = \mu_A^j(x) \wedge \mu_B^j(x), \ j = 1, \ldots, p.$

6. α-**cut:** The α-cut for a fuzzy bag A, denoted by $[A]_\alpha$, is defined as follows.
$\mu_A^1(x) < \alpha \Rightarrow C_{[A]_\alpha}(x) = 0,$
$\mu_A^j(x) \geq \alpha,\ \mu_A^{j+1}(x) < \alpha \Rightarrow C_{[A]_\alpha}(x) = j, \quad j = 1, \ldots, p.$
(Note: The strong α-cut is also defined in a similar manner, but we omit the detail.)

7. **Cartesian product**
Let $A = \{(x, \mu)\}$ and $F = \{(y, \nu)\}$. Then the Cartesian product $A \times F$ is all the combination of x in A and y in F with the membership $\min\{\mu, \nu\}$:
$A \times F = \{((x, y), \min\{\mu, \nu\})\}.$

We give a few propositions below of which the proofs are omitted to save space (see, e.g., [8]).

Proposition 1 *Assume A and B are fuzzy bags of X. The necessary and sufficient condition for $A \subseteq B$ is that for all $\alpha \in (0, 1]$, $[A]_\alpha \subseteq [B]_\alpha$. Moreover, the condition for $A = B$ is that for all $\alpha \in (0, 1]$, $[A]_\alpha = [B]_\alpha$.*

Proposition 2 *Assume A and B are fuzzy bags of X and F is a fuzzy bag of Y. Take an arbitrary $\alpha \in (0, 1]$. We then have*

$$[A \cup B]_\alpha = [A]_\alpha \cup [B]_\alpha,$$
$$[A \cap B]_\alpha = [A]_\alpha \cap [B]_\alpha,$$
$$[A + B]_\alpha = [A]_\alpha + [B]_\alpha,$$
$$[A \times F]_\alpha = [A]_\alpha \times [F]_\alpha.$$

Proposition 3 *Assume A, B, and C are fuzzy bags of X. The followings are valid.*

$$A \cup B = B \cup A, \quad A \cap B = B \cap A,$$
$$A \cup (B \cup C) = (A \cup B) \cup C,$$
$$A \cap (B \cap C) = (A \cap B) \cap C,$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C),$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

The collection of all fuzzy bags by the latter set of operations thus form a distributive lattice.

Let us recall that a fuzzy set can be regarded as a fuzzy bag. To distinguish the both, let a fuzzy set be $A = \{(x, \mu), (x', \mu'), \ldots\}$ and when it is embedded into fuzzy bags, it is written as $\mathcal{E}(A)$, i.e., $\mathcal{E}(A) = \{(x, \mu), (x', \mu'), \ldots\}$ the members and membership values are unchanged but it is regarded as a fuzzy bag. Note also that fuzzy sets can be characterized by the membership sequence with $p = 1$: $\mu_{\mathcal{E}(A)}^1(x) = \mu_A(x)$.

We then have the next proposition.

Proposition 4 *Let A and B are ordinary fuzzy sets. Then the followings are valid.*

1. $A \subseteq B \iff \mathcal{E}(A) \subseteq \mathcal{E}(B)$
2. $\mathcal{E}(A \cup B) = \mathcal{E}(A) \cup \mathcal{E}(B), \quad \mathcal{E}(A \cap B) = \mathcal{E}(A) \cap \mathcal{E}(B).$

Thus the basic relation and operations for fuzzy sets are compatible with those for fuzzy bags by the latter definition.

3 Real-valued Bags

Let us consider another way to generalize crisp bags that is in a sense simpler than fuzzy bags. For this purpose we consider Example 4 again.

Example 6 Let X, Y, f, and input

$$G = \{0.1/x_1, 0.5/x_2, 0.5/x_3, 0.7/y_1, 0.7/y_2, 0.2/z_1, 0.9/z_2\}$$

are the same as those in Example 4. Recall that f with the input G and output G' is defined by the addition:

$$G' = \sum_{(x,\mu)\in G} \{(f(x), \mu)\}.$$

Let us suppose that another definition of the addition is adopted that adds all membership values for each $x \in X$:

$$\begin{aligned}
G' &= \{(0.1 + 0.5 + 0.5)/x, (0.7 + 0.7)/y, (0.2 + 0.9)/z\} \\
&= \{1.1/x, 1.4/y, 1.1/z\}.
\end{aligned}$$

If the input is crisp, e.g., $K = \{x_1, x_2, x_3, y_1, y_2, z_1\}$, the output K' from the above definition is $K' = \{3/x, 2/y, 1/z\}$ which coincides with the original definition of $\sum_{x\in G}\{f(x)\}$ for crisp bags.

We thus have real-valued membership values that is nonnegative but not necessarily in the unit interval.

Definition 1 A real-valued bag B of X is characterized by a count function $C_B \colon X \to [0, +\infty]$ with the following relations and operations.

1. (inclusion): $B \subseteq B' \Leftrightarrow C_B(x) \leq C_{B'}(x), \quad \forall x \in X$.
2. (equality): $B = B' \Leftrightarrow C_B(x) = C_{B'}(x), \quad \forall x \in X$.
3. (union): $C_{B\cup B'}(x) = \max\{C_B(x), C_{B'}(x)\}$.
4. (intersection): $C_{B\cap B'}(x) = \min\{C_B(x), C_{B'}(x)\}$.
5. (addition or sum): $C_{B+B'}(x) = C_B(x) + C_{B'}(x)$.
6. (scalar multiplication): $C_{\alpha B} = \alpha C_B(x)$, where α is a nonnegative real number.
7. (α-cut): Let $\alpha \in [0, +\infty]$ and B is a real-valued bag of Y. α-cut $[B]_\alpha$ of B is a crisp set: $[B]_\alpha = \{x \in X : C_B(x) \geq \alpha\}$

The above definition appears to be the same as that of crisp bags, except that the range of $C_B(\cdot)$ is now $[0, +\infty]$ instead of $\{0, 1, 2, \ldots\}$. In particular, we assume the membership can be infinity: $C_B(x) = +\infty$.

It is clear that embedding fuzzy sets into real-valued bags is also well-defined: for a fuzzy set A, it is embedded into a real-valued bag $\mathcal{E}_R(A)$, i.e., $C_{\mathcal{E}_R(A)}(x) = \mu_A(x)$. We then have

Proposition 5 *Let A and B are fuzzy sets. Then the followings are valid.*

1. $A \subseteq B \iff \mathcal{E}_R(A) \subseteq \mathcal{E}_R(B)$.
2. $\mathcal{E}_R(A \cup B) = \mathcal{E}_R(A) \cup \mathcal{E}_R(B), \quad \mathcal{E}_R(A \cap B) = \mathcal{E}_R(A) \cap \mathcal{E}_R(B)$.
3. $[\mathcal{E}_R(A)]_\alpha = \mathcal{E}_R([A]_\alpha), \quad \forall \alpha \in [0, 1]$.

All basic relations and operations of fuzzy sets and real-valued bags are consistent except the complementation discussed below.

The following propositions which are analogous to Propositions 1–3 hold for real-valued bags and also for fuzzy sets. The proofs are easy and omitted.

Proposition 6 *Assume A and B are real-valued bags of X. The necessary and sufficient condition for $A \subseteq B$ is that for all $\alpha \in (0, +\infty]$, $[A]_\alpha \subseteq [B]_\alpha$. Moreover, the condition for $A = B$ is that for all $\alpha \in (0, +\infty]$, $[A]_\alpha = [B]_\alpha$.*

Proposition 7 *Assume A and B are real-valued bags of X. Take an arbitrary $\alpha \in (0, +\infty]$. We then have*

$$[A \cup B]_\alpha = [A]_\alpha \cup [B]_\alpha,$$
$$[A \cap B]_\alpha = [A]_\alpha \cap [B]_\alpha,$$
$$[A + B]_\alpha = [A]_\alpha + [B]_\alpha.$$

Proposition 8 *Assume A, B, and C are real-valued bags of X. The followings are valid.*

$$A \cup B = B \cup A, \quad A \cap B = B \cap A,$$
$$A \cup (B \cup C) = (A \cup B) \cup C,$$
$$A \cap (B \cap C) = (A \cap B) \cap C,$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C),$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

The collection of all real-valued bags of X thus form a distributive lattice.

3.1 Complementation and s-norms

Complementation of crisp bags has a fundamental problem, i.e., there is no 'maximum bag,' while complementation of an ordinary set is $X - A$ (universe X minus A) where the universe is the maximum set.

To consider a complementation of a real-valued bag, however, has an advantage that will be shown below. The infinity point $+\infty$ is useful for this purpose. That is, we introduce the maximum bag **Infinity**:

$$C_{\mathbf{Infinity}}(x) = +\infty, \quad \forall x \in X. \tag{1}$$

A negation operator $\mathcal{N} \colon [0, +\infty] \to [0, +\infty]$ with the next properties is then used to define a complementation operation:

(i) $\mathcal{N}(0) = +\infty, \quad \mathcal{N}(+\infty) = 0$.
(ii) $\mathcal{N}(x)$ is strictly monotonically decreasing on $(0, +\infty)$.
(iii) $\mathcal{N}(\mathcal{N}(x)) = x$.

A typical example is

$$\mathcal{N}(x) = \frac{const}{x} \qquad (const > 0) \tag{2}$$

A complement $\bar{B}$ of a real-valued bag B is then defined:

9.(complement):
$$C_{\bar{B}}(x) = \mathcal{N}(C_B(x)).$$

This operation justifies the introduction of real-valued bags, since even if we start from crisp bags, the result of complementation is generally real-valued.

We have the following two propositions of which the proofs are easy and omitted.

Proposition 9 *For arbitrary real-valued bags B and C, the next properties are valid:*

$$\overline{(\overline{B})} = B \tag{3}$$

$$\overline{B \cup C} = \bar{B} \cap \bar{C}, \qquad \overline{B \cap C} = \bar{B} \cup \bar{C}. \tag{4}$$

Proposition 10 *Let an empty bag $\emptyset$ be*

$$C_{\emptyset}(x) = 0, \quad \forall x \in X. \tag{5}$$

Then we have

$$\bar{\emptyset} = \mathbf{Infinity}, \quad \overline{\mathbf{Infinity}} = \emptyset. \tag{6}$$

s-norms and t-norms

We introduce two functions $s(a,b)$ and $t(a,b)$ like those in fuzzy sets, but the boundary conditions are different.

Definition 2 Two functions $s\colon [0, +\infty]^2 \to [0, +\infty]$ and $t\colon [0, +\infty]^2 \to [0, +\infty]$ having the next properties (I)–(IV) are called an s-norm and a t-norm for real-valued bags, respectively.

(I)[monotonicity] For $a \leq c$, $b \leq d$,

$$s(a,b) \leq s(c,d), \quad t(a,b) \leq t(c,d).$$

(II)[symmetry]

$$t(a,b) = t(b,a), \quad s(a,b) = s(b,a).$$

(III)[associativity]

$$s(s(a,b),c) = s(a,s(b,c)), \quad t(t(a,b),c) = t(a,t(b,c)).$$

(IV)[boundary condition]

$$s(+\infty, +\infty) = +\infty, \quad s(a,0) = s(0,a) = a,$$
$$t(0,0) = 0, \quad t(a, +\infty) = t(+\infty, a) = a.$$

A purpose to introduce such norms for real-valued bags is to generalize the intersection and union operations. First we note that $s(a,b) = a + b$, $s(a,b) = \max\{a,b\}$, and $t(a,b) = \min\{a,b\}$ satisfy the above conditions (I)–(IV). Thus the s-norms and t-norm represent the addition, union, and intersection. We moreover introduce a generating function $g(x)$ for s-norm.

Definition 3 A function $g\colon [0, +\infty] \to [0, +\infty]$ is called a generating function for s-norm if it satisfies the next (i)–(iii):

(i) it is strictly monotonically increasing,

(ii) $g(0) = 0$, $\quad g(+\infty) = +\infty$,

(iii) $g(x + y) \geq g(x) + g(y)$, $\quad \forall x, y \in [0, +\infty]$.

We have the next two propositions.

Proposition 11 *Let*

$$s(a, b) = g^{-1}(g(a) + g(b)). \tag{7}$$

Then $s(a, b)$ is an s-norm.

An example of the generating function is

$$g(x) = x^p \quad (p \geq 1). \tag{8}$$

Proposition 12 *Let $s(a, b)$ is an s-norm and $\mathcal{N}$ is a complementation operator. Then*

$$t(a, b) = \mathcal{N}(s(\mathcal{N}(a), \mathcal{N}(b))) \tag{9}$$

is a t-norm. Suppose $t(a, b)$ is a t-norm, then

$$s(a, b) = \mathcal{N}(t(\mathcal{N}(a), \mathcal{N}(b))) \tag{10}$$

is an s-norm.

If a pair of t-norm and s-norm has the property stated in the last proposition, we say (s, t) has the *duality* of s-norm and t-norm.

We apply s-norm and t-norm to define bag operations $B\mathcal{S}B'$ and $B\mathcal{T}B'$:

$$C_{B\mathcal{S}B'}(x) = s(C_B(x), C_{B'}(x)). \tag{11}$$

$$C_{B\mathcal{T}B'}(x) = t(C_B(x), C_{B'}(x)). \tag{12}$$

Let us consider typical examples.

Example 7 The standard operators

$$s(a, b) = \max\{a, b\} \tag{13}$$

$$t(a, b) = \min\{a, b\} \tag{14}$$

are an s-norm and a t-norm, respectively. This pair has the duality where $\mathcal{N} = const/x$.

Example 8 Using generating function $g(x) = x^p$, we have

$$s(a, b) = (a^p + b^p)^{\frac{1}{p}}, \tag{15}$$

$$t(a, b) = (a^{-p} + b^{-p})^{-\frac{1}{p}}. \tag{16}$$

This example includes the addition, max, and min operations. First, $s(a, b) = a + b$ is a particular case of (15) for $p = 1$. Moreover $s(a, b) = \max\{a, b\}$ and $t(a, b) = \min\{a, b\}$ are obtained from (15) and (16) when $p \to +\infty$.

The s-norms and t-norms are applied to define operations for bag relations discussed later.

4 Fourth Generalization

We now have two different types of generalizations of bags and fuzzy sets: fuzzy bags with the second type of operations and real-valued bags. A natural question is: what is a minimal extension that includes the both. To answer this question, we introduce 'region-valued bags' in this section and call them G-bags (generalized bags).

A G-bag (generalized bag) A is characterized by a count function as a region in $\boldsymbol{R}^2$ [13]:

$$C_A(x) = \mathbf{N}(x, A), \quad \forall x \in X \tag{17}$$

where $\mathbf{N}(x, A)$ is a closed set of $[0, +\infty)^2$. Let $\nu(y, z; x, A)$ is the characteristic function of the region $\mathbf{N}(x, A)$:

$$\nu(y, z; x, A) = \begin{cases} 1, & (y, z) \in \mathbf{N}(x, A), \\ 0, & (y, z) \notin \mathbf{N}(x, A). \end{cases}$$

This region should have the following properties:

(N1) For any $y \in [0, +\infty)$,

$$\eta(y) = \{z \in [0, +\infty) : \nu(y, z; x, A) = 1\}$$

is either an empty set or a closed interval $[0, \mu(y)]$ where $\mu(y) = \max\{y' : y' \in \eta(y)\}$.
(N2) For all $y \in [0, +\infty)$, $\mu(y)$ is a monotonically decreasing function and $0 \le \mu(y) \le 1$.
(N3) For any $z \in [0, +\infty)$,

$$\theta(z) = \{y \in [0, +\infty) : \nu(y, z; x, A) = 1\}$$

is either an empty set or a closed interval $[0, \zeta(z)]$ where $\zeta(z) = \max\{z' : z' \in \theta(z)\}$.
(N4) For all $z \in [0, +\infty)$, $\zeta(z)$ is a monotonically decreasing function and $0 \le \zeta(z) < +\infty$.

Roughly, a generalized bag is characterized by the pair of monotonically decreasing functions $(\mu(y), \zeta(z))$ for each $x \in X$. These two functions are expressed as $\mu(y; x, A)$ and $\zeta(z; x, A)$, respectively, to explicitly show its dependence to x and A. An example of $\mathbf{N}(x, A)$ is shown in Figure 1. The inclusion and equality of G-bags are defined by
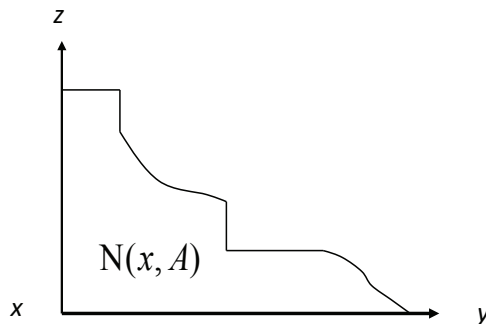
Fig. 1 An illustration of $\mathbf{N}(x, A)$ as the count function for a G-bag.

$$A \subseteq B \iff C_A(x) \subseteq C_B(x), \quad \forall x \in X, \tag{18}$$

$$A = B \iff C_A(x) = C_B(x), \quad \forall x \in X, \tag{19}$$

and the union and intersection are

$$C_{A \cup B}(x) = C_A(x) \cup C_B(x), \quad \forall x \in X, \tag{20}$$

$$C_{A \cap B}(x) = C_A(x) \cap C_B(x), \quad \forall x \in X. \tag{21}$$

G-bags are justified as a minimal extension of fuzzy bags and real-valued bags by the following consideration. Let B_j is a 'very thin' real-valued bag:

$$C_{B_j}(x) = \epsilon_j(x), \quad 0 < \epsilon_j(x) \ll 1.$$

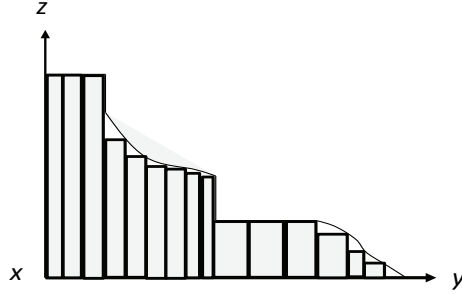Let F_j is a fuzzy bag with the membership sequence

$$(\beta_j(x), 0, 0, \ldots).$$

Thus, only one member is nonzero in the membership sequence. Put $D_j = B_j \cap F_j$: we then have a membership of a 'step function' with the width $\epsilon_j(x)$ and the height $\beta_j(x)$.

Let us add many D_j of such step functions:

$$D = D_1 + D_2 + \cdots + D_j + \cdots$$

The resulting D approximates $\mathbf{N}(x, A)$ as in Figure 2.

Fig. 2 Approximation of a region by the addition of step functions.

On the other hand, we reduce a G-bag to a fuzzy set and a real-valued bag using an α-cut and a ν-cut.

Definition 4 Given a G-bag A of X and $\alpha \in [0,1]$, an α-cut $[A]_\alpha$ of A is a real-valued bag characterized by:

$$C_{[A]_\alpha}(x) = \zeta(\alpha; x, A).$$

Given $\nu \in [0, +\infty)$, a ν-cut $_\nu[A]$ of A is a fuzzy set characterized by

$$\mu_{\nu[A]}(x) = \mu(\nu; x, A).$$

Note 1 Why we can say that G-bags are a 'minimal extension' which include both fuzzy bags and real-valued bags is that $D = D_1 + D_2 + \cdots + D_j + \cdots$ as defined above is easily derived by operations of fuzzy bags and real-valued bags. It approximates and approaches $\mathbf{N}(x, A)$ using some natural topology of which the details are omitted. Hence to use $\mathbf{N}(x, A)$ is easier than to handle the collection of all $D = D_1 + D_2 + \cdots + D_j + \cdots$.

5 Bag relations

As bags are contrasted with fuzzy sets, we can develop a theory of bag relations in contrast to fuzzy relations. In a sense bag relations are more natural, having more application possibilities.

We consider bag relations for real-valued bags for simplicity, but generalization to G-bag relations are straightforward by using the α-cuts.

To this end, we first introduce a max-s algebra and a max-t algebra:

5.1 max-s and max-t algebra

Let us use a particular notation of $\boxplus$ and $\boxdot$ for

$$a \boxplus b = \max\{a, b\}, \quad a \boxdot b = s(a, b) \tag{22}$$

where $s(a, b)$ is an s-norm for real-valued bags. We call this algebra as *max-s algebra*.

It is easy to see that the following properties hold.

$$a \boxplus b = b \boxplus a, \tag{23}$$

$$a \boxplus (b \boxplus c) = (a \boxplus b) \boxplus c, \tag{24}$$

$$a \boxplus 0 = a, \tag{25}$$

$$a \boxdot b = b \boxdot a, \tag{26}$$

$$a \boxdot (b \boxdot c) = (a \boxdot b) \boxdot c, \tag{27}$$

$$a \boxdot 0 = a. \tag{28}$$

Alternatively, we can define $\boxplus$ and $\boxdot$ for

$$a \boxplus b = \max\{a, b\}, \quad a \boxdot b = t(a, b) \tag{29}$$

where $t(a, b)$ is a t-norm for real-valued bags. We call this algebra as *max-t algebra*. We see that (23)–(27) hold, while (28) should be replaced by

$$a \boxdot +\infty = a. \tag{30}$$

Note the next lemma of which the proof is easy and omitted.

Lemma 1 *Let a, b, c be real numbers. Then*

$$a \boxdot (b \boxplus c) = (a \boxdot b) \boxplus (a \boxdot c). \tag{31}$$

The commutative, associative, and distributive properties thus hold for the max-s and max-t algebras. We observe that the max-t algebra includes the max-min algebra and the max-s algebra includes the max-plus algebra [3] as particular cases.

We now define bag relations.

Definition 5 An real-valued bag relation R on $X \times Y$ is a real-valued bag R of $X \times Y$. The count function is denoted by $R(x, y)$ instead of $C_R(x, y)$ for simplicity. A real-valued bag relation is also called bag relation for simplicity.

The reason why we call such a real-valued bag $R(x, y)$ a bag relation is that we can define composition operation.

Let X, Y, and Z be three universes. Assume R is a bag relation of $X \times Y$ and S is a bag relation of $Y \times Z$. Then the max-s composition $R \circ S$ is defined as follows.

$$(R \circ S)(x, z) = \boxplus_{y \in Y} \{R(x, y) \boxdot S(y, z)\} \tag{32}$$

Moreover the max-t composition is defined by the same equation (32) except that $\boxdot$ uses a t-norm.

Note 2 The addition is defined in a straightforward manner:

$$(R_1 \boxplus R_2)(x, y) = R_1(x, y) \boxplus R_2(x, y), \tag{33}$$

for bag relations on $X \times Y$. It is clear that the addition satisfies the commutative property and the associative property:

$$R_1 \boxplus R_2 = R_2 \boxplus R_1, \tag{34}$$

$$(R_1 \boxplus R_2) \boxplus R_3 = R_1 \boxplus (R_2 \boxplus R_3). \tag{35}$$

Note that

$$\boxplus_{y \in \{a_1, \ldots, a_L\}} A(y) = A(a_1) \boxplus A(a_2) \boxplus \cdots \boxplus A(a_L). \tag{36}$$

Hence we can write (32).

We have the following proposition.

Proposition 13 *The composition satisfies the associative property*

$$(R \circ S) \circ T = R \circ (S \circ T), \tag{37}$$

and the distributive property

$$(R_1 \boxplus R_2) \circ S = (R_1 \circ S) \boxplus (R_2 \circ S), \tag{38}$$

$$R \circ (S_1 \boxplus S_2) = (R \circ S_1) \boxplus (R \circ S_2). \tag{39}$$

We introduce the unit relations for the max-s and max-t compositions. Define O_X and Ω_X on $X \times X$:

$$O_X(x,y) = 0, \quad \forall x,y \in X, \tag{40}$$

$$\Omega_X(x,y) = +\infty, \quad \forall x,y \in X. \tag{41}$$

We define O_{XY} and Ω_{XY} on $X \times Y$ in the same way:

$$O_{XY}(x,y) = 0, \quad \forall (x,y) \in X \times Y, \tag{42}$$

$$\Omega_{XY}(x,y) = +\infty, \quad \forall (x,y) \in X \times Y. \tag{43}$$

We frequently omit the subscripts like O and Ω when we have no ambiguity.

The following two propositions are easily derived.

Proposition 14 *Assume that the max-s algebra is used. For an arbitrary bag relation R on $X \times Y$,*

$$R \boxplus O = O \boxplus R = R, \tag{44}$$

$$R \circ O = O \circ R = R. \tag{45}$$

Proposition 15 *Assume that the max-t algebra is used. For an arbitrary bag relation R on $X \times Y$,*

$$R \boxplus O = O \boxplus R = R, \tag{46}$$

$$R \circ \Omega = \Omega \circ R = R. \tag{47}$$

Note 3 The detailed proof of Proposition 13 is lengthy, but the way to prove this proposition is just the same as the proof of the matrix properties for ordinary algebra, since those properties for ordinary algebra uses the commutative, associative, and distributive laws which also are valid for the present algebra, since (23)–(28) (or (30)) and (31) hold.

Note 4 The bags relations can also be defined for G-bags by using α-cuts. In this case a G-bag relation is a G-bag of the product space. The composition operation is defined by α-cuts:

$$[R \circ S]_\alpha(x,z) = \boxplus_{y \in Y} \{ [R(x,y)]_\alpha \boxdot [S(y,z)]_\alpha \}$$

and then the membership is defined by:

$$(R \circ S)(x,z) = \bigcup_\alpha ([0, [R \circ S]_\alpha(x,z)] \times [0,\alpha]).$$

6 Possibilities for Real Applications

While the fourth generalization of G-bag is still on the stage of theoretical interests and having no direct applications, there are a number of possibilities to apply real-valued bags and fuzzy bags to real-world problems. We briefly mention two applications to graphs and data analysis.

Graphs and networks

As the classical problem of Euler graph is described by multigraphs [2], there are many problems described by bags ad real-valued bags. Indeed, max-min and max-plus algebra have applications to network problems [3]. Thus, real-valued bags and fuzzy bags as well as bag relations will generalize current results in graphs and networks.

Data analysis using bag models

Bag models are widely used in data analysis and information retrieval. Moreover, flexible querying using fuzzy bags has also been studied [21]. As web data retrieval results in multiple data with relevance degrees, they are modeled by fuzzy bags. Analysis of such data requires methods for bags and fuzzy bags. Clustering techniques using fuzzy bag models have been proposed [12,17] and text analysis based on bag relations are being studied using kernel functions [16].

7 Conclusions

We have overviewed four different generalizations of bags. More precisely, one of the two fuzzy bags do not generalize fuzzy sets, while the other really does. We note that most studies of fuzzy bags are using Yager's definition that is a mapping of fuzzy sets and not a generalization of them. One reason why they do not use the latter generalization is that they rarely use the union and intersection of fuzzy bags but mainly use the addition; the addition is the same between the two different systems of fuzzy bags.

Real-valued bags are simpler and convenient for introducing complementation, since the infinity point is included as a membership value. Consequently s-norms and t-norms were defined and used in bag operations. These norms were then used to define compositions of bag relations.

Direct applications of G-bags are not yet studied but G-bags are closely related to real-valued bags with fuzzy numbers [13]. This suggests that G-bags can be used in applications instead of real-valued bags with fuzzy numbers.

Although this paper is mainly theoretical, possible applications were suggested, which should further be investigated in near future. Real-world applications such as Kansei evaluation [24] using measures on fuzzy bags should be studied.

References

1. C.S. Calude, G. Păun, G. Rozenberg, A. Salomaa, eds., *Multiset Processing*, Lecture Notes in Computer Science, LNCS 2235, Springer, Berlin (2001).
2. G. Chartrand, L. Lesniak, *Graphs & Digraphs, 2nd Ed.*, Wadsworth & Brooks/Cole, Monterey, CA (1986).
3. B. Heidergott, G.J. Olsder, J. van der Woude, *Max Plus at Work*, Princeton University Press, Princeton, NJ (2006).
4. D. E. Knuth, *The Art of Computer Programming, Vol. 2 / Seminumerical Algorithms*, Addison-Wesley, Reading, Massachusetts (1969).

5. B. Li, W. Peizhang, and L. Xihui, Fuzzy bags with set-valued statistics, *Comput. Math. Applic.*, Vol. 15, pp. 811–818 (1988).
6. B. Li, Fuzzy bags and applications, *Fuzzy Sets and Systems*, Vol. 34, pp. 61–71 (1990).
7. Z. Manna and R. Waldinger, *The Logical Basis for Computer Programming, Vol. 1: Deductive Reasoning*, Addison-Wesley, Reading, Massachusetts (1985).
8. S. Miyamoto, Basic operations of fuzzy multisets, *Journal of Japan Society for Fuzzy Theory and Systems*, Vol. 8, No. 4, pp. 639–645 (1996) (in Japanese).
9. S. Miyamoto, Fuzzy multisets with infinite collections of memberships, *Proc. of the 7th International Fuzzy Systems Association World Congress (IFSA'97)*, June 25-30, 1997, Prague, Czech, Vol.1, pp.61-66 (1997).
10. S. Miyamoto, K. S. Kim, An image of fuzzy multisets by one variable function and its application, *J. Japan Society for Fuzzy Theory and Systems*, Vol. 10, No. 1, pp. 157–167 (1998) (in Japanese).
11. S. Miyamoto, K. S. Kim, Multiset-valued images of fuzzy sets, *Proceedings of the Third Asian Fuzzy Systems Symposium*, June 18-21, 1998, Masan, Korea, pp.543-548 (1998).
12. S. Miyamoto, Information clustering based on fuzzy multisets, *Information Processing and Management*, Vol.39, No.2, pp. 195–213 (2003).
13. S. Miyamoto, Generalizations of multisets and rough approximations, *International Journal of Intelligent Systems*, Vol.19, No.7, pp. 639–652 (2004).
14. S. Miyamoto, Remarks on basics of fuzzy sets and fuzzy multisets, *Fuzzy Sets and Systems*, Vol.156, No.3, pp. 427–431 (2005).
15. S. Miyamoto, Operations for Real-Valued Bags and Bag Relations, Proc. of 2009 Intern. Fuzzy Systems Assoc.World Congress (IFSA-EUSFLAT 2009), Lisbon, Portugal, July 20-24, 2009, pp.612-617 (2009).
16. S. Miyamoto, Y. Kawasaki, K. Sawazaki, An Explicit Mapping for Kernel Data Analysis and Application to Text Analysis, *Proc. of 2009 Intern. Fuzzy Systems Assoc.World Congress (IFSA-EUSFLAT 2009)*, Lisbon, Portugal, July 20-24, 2009, pp.618-623 (2009).
17. K. Mizutani, R. Inokuchi, S. Miyamoto, Algorithms of nonlinear document clustering based on fuzzy multiset model, *International Journal of Intelligent Systems*, Vo.23, No.2, pp.176-198 (2008).
18. A. Ramer and C.-C. Wang, Fuzzy multisets, *Proc. of 1996 Asian Fuzzy Systems Symposium*, Dec. 11-14, 1996, Kenting, Taiwan, pp. 429–434 (1996).
19. A. Rebai, Canonical fuzzy bags and bag fuzzy measures as a basis for MADM with mixed non cardinal data, *European J. of Operational Res.*, Vol. 78, pp. 34–48 (1994).
20. A. Rebai, J.-M. Martel, A fuzzy bag approach to choosing the "best" multiattributed potential actions in a multiple judgment and non cardinal data context, *Fuzzy Sets and Systems*, Vol. 87, pp. 159–166 (1997).
21. D. Rocacher, On fuzzy bags and their application to flexible querying, *Fuzzy Sets and Systems*, Vol. 140(1), pp.93-110 (2003).
22. R. R. Yager, On the theory of bags, *Int. J. General Systems*, Vol. 13, pp. 23–37 (1986).
23. R. R. Yager, Cardinality of fuzzy sets via bags, *Mathl. Modelling*, Vol. 9, No. 6, pp. 441–446 (1987).
24. H.-B. Yan, V.-N. Huynh, Y. Nakamori, A group nonadditive multiattribute consumer-oriented Kansei evaluation model with an application to traditional crafts, Annals of Operations Research, DOI 10.1007/s10479-010-0826-7 (2011).