ANNALS OF COMMUNICATIONS IN MATHEMATICS
Volume 4, Number 3 (2021), 284-292
ISSN: 2582-0818
© http://www.technoskypub.com

ON DYNAMIC MULTISETS AND THEIR OPERATIONS

T. O. WILLIAM-WEST, P. A. EJEGWA* AND A. U. AMAONYEIRO

ABSTRACT. The concept of multisets emerged by violating a principle of distinct collection of object in crisp sets. In some practical situations, multisets with multiplicity of their objects varying overtime are frequently encountered, such multisets are called dynamic multisets. However, there has been no formal mathematical study on dynamic multisets. Dynamic multiset is a special kind of multiset with time varying multiplicity of elements. The importance of dynamic multisets stems from their potential usefulness in resolving a task of finding duplicate records within large databases. In this paper, we vividly explore the concept of dynamic multisets and present some of its properties. We observe that, the operations on dynamic multisets are the same as that of static multisets, with the time parameter as the only distinction. Finally, some application-driven examples of dynamic multisets are presented.

1. INTRODUCTION

In many branches of computer science such as database systems, membrane systems, information retrieval systems etc., duplicates or multiplicities of data are encountered. Multiset theory initiated by [12] has the capacity to implement the structural analysis of such real-world data duplication. A multiset is a collection of objects in which objects are allowed to repeat [5]. Multiset theory is an established area of research in non-classical set theory with vast applications in philosophy, logic, linguistics, physics, mathematics, computer science, etc. [6, 14]. However, till date most of the research in multiset theory has considered only static multisets, i.e., multisets that do not change with time. A wealth of such literature has been developed for static multiset theory [1, 4, 7, 10, 15, 16, 17]. Multiset has been studied in the framework of rough sets [9].

The concept of dynamic multiset was first addressed in [13] to underpin a concept of cardinality queries and lookups in constant time. Pagh et al. [13] investigated the problem of storing dynamic multisets succinctly. Recently, dynamic dictionaries for multisets and counting filters with constant time operations has been discussed [3]. However, the mathematical construct of dynamic multisets have not been formalized. Our attempt to deal with situations involving data duplication mentioned above lead us to mathematically formalize dynamic multisets, i.e., multisets that change with time. Dynamic multisets appear in many

---

2010 *Mathematics Subject Classification.* 54A40, 03E72, 20N25, 06D72.
Key words and phrases. Multisets; Dynamic multisets; Data duplication.
Received: November 21, 2021. Accepted: December 21, 2021. Published: December 31, 2021.
*Corresponding author.

contexts. This is especially true in the World Wide Web (WWW), where almost always, based on different web search at different time intervals, varying number of multiplicity of web documents are returned due to the availability or non-availability of a host server, or overloaded database storage memory; which may lead to deletion of certain duplicates of copies or otherwise.

Very little is known about the properties of dynamic multisets. Given a wide range of their potential application and the frequency with which they appear in most areas (biological systems), their study is of considerable importance. There are several main avenues supporting the introduction of dynamic multisets.

- Large databases, especially big data system; when handling data duplication records within large databases, the notion of dynamic multisets is an important topic. The importance stems from resolving a task of finding and analyzing duplicate records in large database systems.
- Real time stock computation in Mega stores, such as Malls and Shops exhibit another position as to the introduction of dynamic multisets. In the event whereby per time update of available stock of various items sold in a Mega store is to be made after sales, a concept of dynamic multisets plays a significant role. Given such a situation, the constructs of dynamic multisets can be deployed and exploited for qualitative representation and update of information.

There is a strong motivation to introduce the mathematical formalism of dynamic multisets, no matter which way one decides to proceed:

- Dynamic multisets help us to perceive in far more meaningful structure, the rate of data duplication in ubiquitous systems.
- Dynamic multisets allow us to compute and update instantaneous multiplicity of files and objects at various time intervals in both biological and or reactive systems.

While dynamic multisets are subsets of static multisets, in general, they constitute a useful concept and powerful tool to study the computing of duplicate records in ubiquitous and biological systems, and their construct could be a useful model in relevant structures to a great extent.

2. Basic Notion of Dynamic Multisets: Main Concepts and Their Realization

For completeness, we will review the concept of multisets.

2.1. **Multisets.** Here, we recall the concept of multisets.

Definition 2.1. [18] Let $X = \{x_1, x_2, \cdots, x_n\}$ be a set. A multiset M drawn from X is a cardinal-valued function; i.e., $M\colon X \to V = \{0, 1, 2\}$ such that for $x \in \mathrm{Dom}(M)$ implies $M(x)$ is a cardinal and $M(x) = C_M(x) \geq 0$, where $C_M(x)$ denotes the number of times (i.e., the count of $x \in X$) an object x occurs in M. The set X is called the ground set of the classes of all multisets containing objects from M.

A multiset M can be represented by the set as follows [15]:

$$M = \left\{\frac{C_M(x_1)}{x_1}, \frac{C_M(x_2)}{x_2}, \cdots, \frac{C_M(x_n)}{x_n}\right\}. \tag{2.1}$$

Example 2.2. Let $X = \{a, b, c, d\}$ be a domain of discourse, then a multiset M drawn from X is represented thus;

$$M = \{a, a, a, b, b, c, c, c, c\} = \left\{\frac{3}{a}, \frac{2}{b}, \frac{4}{c}\right\},$$

where $C_M(a) = 3$, $C_M = 2$ and $C_M(c) = 4$. Noticeable is the cardinality of M, denoted by card(M) which in this case;

$$\begin{aligned}\text{card}(M) &= C_M(a) + C_M(b) + C_M(c)\\ &= 9.\end{aligned}$$

The reader is referred to Singh et al. [15] for detailed examples and operations on multisets.

2.2. **Dynamic multisets: the idea.** The underlying idea of dynamic multisets is to provide time-varying and computationally appealing representation of multisets.

The definition of (static) multiset involves the following entities: X (a ground set of objects), $C_M(x)$ (a count of the number of multiplicities of x in M). Informally, a *dynamic* multiset is obtained when any of these entities change over time $t \in T \subseteq \mathbb{R}^+$, where $\mathbb{R}^+ = [0, +\infty)$. Thus there are two basic main kinds of dynamic multisets.

- In a ground dynamic multiset, the membership of objects in X vary with time $t \in T$; some objects may be inserted into X or removed from X. When objects are inserted in X and/or removed from X, their corresponding count (i.e., the number of multiplicity of an object) respectively increase and/or decrease.

Example 2.3. Let $X = \{a, b, c, d\}$ be a domain of discourse representing the types of items sold in a certain shopping mall. Suppose the following multiset represent the various multiplicity of this items in stock $M_{t_0} = \left\{\frac{4}{a}, \frac{6}{b}, \frac{3}{c}, \frac{2}{d}\right\}$ drawn from X at time t_0.

At $t_1 > t_0$, suppose $X_t = \{a, b, d\}$, is the ground set of available stock in the Mall such that the total items available is represented by M_{t_1}. Since the item c is no longer in stock, we need not have a multiset which contains the element c.

In a similar fashion, a new item (i.e., an element e) could be inserted in X_{t_2} at some time t_2. As a result, the multiset M_{t_2}, its cardinality as well as its count of its objects may be time-varying.

- In a count dynamic multiset, the count $C_M(\cdot)$ varies with respect to time. Thus, multiple copies of various objects may be inserted in a dynamic multiset $M_{t\in T}$ or removed from a dynamic multiset $M_{t\in T}$ at different time intervals $t_i \in T$ without necessarily interfering with the elements of the domain of discourse. Hence in a count dynamic multiset the cardinality of the universe of discourse may be fixed even when that of its induced multiset is time-varying.

Example 2.4. Let $X = \{l, s\}$ be a set of some cells in the human body such that l represent a liver cell and s represent a cell lining the stomach. The two types of cells are well known to multiply in the human body at different times. Hence a count dynamic multiset

$$\overrightarrow{M}_{t_l} = \left\{\frac{\overrightarrow{C}_{\overrightarrow{M}_{t_i}}(l)}{l}, \frac{\overrightarrow{C}_{\overrightarrow{M}_{t_i}}(s)}{s}\right\},\ t_i \in T \text{ and } i = 0, 1, 2, \cdots \tag{2.2}$$

is realized. Following the nature of $\overrightarrow{M}_{t_l}$, consistent information about the multiplicities of these cells at different times can be reported.

In passing, all combination of the above types of dynamic multisets may occur.

2.3. **Dynamic multisets.** Let a mapping (see Eq. 2.3) be defined in the domain of discourse X of n-objects. A dynamic multiset M_D is characterized by the triple

$$(M, \overleftarrow{M}, \overrightarrow{M}) : (X, \overleftarrow{X}, \overrightarrow{X})_t \longrightarrow \{0, 1, 2, \cdots\} \times \{\cdot, \leftarrow, \rightarrow\}_t \tag{2.3}$$

defined by

$$(x, \overleftarrow{x}, \overrightarrow{x})_t \mapsto (C_M(x), C_M(\overleftarrow{x}), C_M(\overrightarrow{x}))_t \tag{2.4}$$

or by

$$(x, \overleftarrow{x}, \overrightarrow{x})_t \mapsto (C_M(x), \overleftarrow{C}_{\overleftarrow{M}}(x), \overrightarrow{C}_{\overrightarrow{M}}(x))_t \tag{2.5}$$

such that $(M, \overleftarrow{M}, \overrightarrow{M}) = M$ or $\overleftarrow{M}$ or $\overrightarrow{M}$ at some given interval, then we call $\mathcal{M} = (M, \overleftarrow{M}, \overrightarrow{M})$ the ground dynamic multiset of $(X, \overleftarrow{X}, \overrightarrow{X})_T$ if Eq. 2.3 is defined by Eq. 2.4. The multiset $\mathcal{M} = (M, \overleftarrow{M}, \overrightarrow{M})$ is called a count multiset of $(X, \overleftarrow{X}, \overrightarrow{X})$ if Eq. 2.3 is defined by Eq. 2.5. Moreover, the last expression in Eqs. 2.4 and 2.5 represent the count of $x \in M_D$ at time $t \in T$.

Let us elaborate on the main components of M_D (or simply M) for a fixed time $t \in T$. The elements of X with $C_M(x)$ and $C_M(\overleftarrow{x})$, $C_M(\overrightarrow{x})$ or $\overleftarrow{C}_{\overleftarrow{M}}(x)$, $\overrightarrow{C}_{\overrightarrow{M}}(x)$ respectively constitute the counts of the static and dynamic multisets discussed so far about the construct. Subsequently, the elements with count $C_M(x)$ are both in X and M. The elements with count $C_M(\overleftarrow{x})$ have been removed from X and need not appear in M, hence this removal initiates $\overleftarrow{X}$.

A similar instance holds for newly inserted elements with count $C_M(\overrightarrow{x})$, as a result, $\overrightarrow{X}$ is realized. The elements with counts $\overleftarrow{C}_{\overleftarrow{M}}(x)$ and $\overrightarrow{C}_{\overrightarrow{M}}(x)$ have been respectively removed from M and inserted in M. Consequently, the dynamic multisets $\overleftarrow{M}$ and $\overrightarrow{M}$ are realized. At this point we stress that the three-operation set $\{\cdot, \leftarrow, \rightarrow\}_T$ induces neutral, remove and insert operations on X and M_D, with the neutral operation taken to mean do nothing or static. Finally, the numbers $\overleftarrow{C}_{\overleftarrow{M}}(x)$ and $\overrightarrow{C}_{\overrightarrow{M}}(x)$ may be encountered as a result of deploying the remove and insert operations simultaneously performed in X and M_D. In view of the aforesaid transformations in X and M_D, some properties of multisets with time varying membership and count can be envisioned.

Alluding the formal definition of dynamic multisets M_D in Eq. 2.3, it is worth underlining that M_D is described by an operation-valued set rather than merely cardinal-valued count function. The co-domain (or range) of M_D consists of three main kinds of counts components; $C_M(x)$ and $C_M(\overleftarrow{x}), C_M(\overrightarrow{x})$ and $\overleftarrow{C}_{\overleftarrow{M}}(x), \overrightarrow{C}_{\overrightarrow{M}}(x)$. They can be respectively treated as count of a static multiset, count of a ground dynamic multiset and count of a count dynamic multiset. These three kinds of quantificational levels come in an apparent interpretation. All elements with count $C_M(x)$ cannot be removed from the construct or dataset X and M at time $t \geq 0$. The elements with count $C_M(\overleftarrow{x})$ and $C_M(\overrightarrow{x})$ may be respectively removed from the construct, list or dataset X, M_D at any time, t, and inserted in the construct, list or dataset X and M_D at any time t. At anytime $t \in T$, the elements with counts $\overleftarrow{C}_{\overleftarrow{M}}(x)$ and $\overrightarrow{C}_{\overrightarrow{M}}(x)$ may be respectively removed from the set of duplicates M_D only and inserted in the set of duplicates M_D only.

3. Operations on Dynamic Multisets

Throughout this paper we set

$$S_t = \left(X, \overleftarrow{X}, \overrightarrow{X}\right)_t,\ s_t = \left(x, \overleftarrow{x}, \overrightarrow{x}\right)_t,\ M_t = \left(M, \overleftarrow{M}, \overrightarrow{M}\right)_t \text{ and } C_{M_t}(\cdot)$$

to represent the corresponding count function of M_t at time $t \in T \subseteq \mathbb{R}^+$. We begin by presenting some constructs that can be gainfully exploited to describe the operations under dynamic multisets.

3.1. **Maximum multiset of a family of dynamic multisets.** Let $\mathcal{F} = \{M_t, M_t', \cdots\}$ be a family of dynamic multisets drawn from S_t. The maximum dynamic multiset $\hat{M}_t$ is defined by $C_{\hat{M}_t}(s_t) = \max_{M_t \in \mathcal{F}} C_{M_t}(s_t)$ for all $s_t \in S_t$ and all $M_t \in \mathcal{F}$, for each $t \in T$. This idea is relevant for defining the complement of a dynamic multiset. A similar variant for the maximum of a family of static multisets have be defined in [11].

It is important to underline that apart from the time varying parameter of a dynamic multiset, its operations are similar to that of a static multiset. Therefore, the following operations are defined for a fixed time t.

- **Union**. The union of two dynamic multisets M_t and M_t' denoted by $M_t \cup M_t'$, is defined as the largest dynamic multiset containing both M_t and M_t'. From this we get the count of $M_t \cup M_t'$;
$$C_{M_t}(x) \vee C_{M_t'}(x) = \max\Big[C_{M_t}(x), C_{M_t'}(x)\Big] \text{ for all } x \in S_t.$$
- **Intersection**. The intersection of two dynamic multisets M_t and M_t', denoted by $M_t \cap M_t'$, is defined as the smallest dynamic multiset contained in both M_t and M_t'. We know the count of $M_t \cap M_t'$ to be
$$\min\Big[C_{M_t}(x), C_{M_t'}(x)\Big] \text{ for all } x \in S_t.$$
- **Complementation**. A dynamic multiset M_t^c is the complement of a dynamic multiset M_t if and only if $M_t^c = \hat{M}_t - M_t$ is
$$C_{M_t^c}(x) = \Big[C_{\hat{M}_t'}(s_t) - C_{M_t}(x), \text{ for all } \quad x \in S_t\Big].$$
- **Equality**. Two dynamic multisets M_t and M_t' are equal, $M_t = M_t'$ if and only if $C_{M_t}(x) = C_{M_t'}$, for all $x \in S_t$.
- **Inclusion**. A dynamic multiset M_t' is contained in a dynamic multiset M_t if and only if $C_{M_t'}(x) \leq C_{M_t}$, for all $x \in S_t$.

It could be easily seen that the aforesaid operations are well-suited for dynamic multisets occurring at different instants of time. Further, the notion of monotone increasing or decreasing sets can be deployed in the construct to analyse the structure of a family of dynamic multisets.

4. Application Examples

In what follows, some application-driven examples of our construct are outlined.

4.1. **Biological Systems.** The use of dynamic multisets can be also substantial in biological systems especially to model the behaviour of such systems. In comparison to static multisets, dynamic multisets may be found to be relatively more apt to model multiplicities of reagents in dynamics of biochemical reactions (see Fig. 1), particularly for reporting consistent information about the multiplicities of reagents (see [2] for details about the static multiset variant of such application). A miniature system which mimics distinct transitions that may occur in the multiset of such a biological or biochemical system is envisioned as shown in Fig. 1.

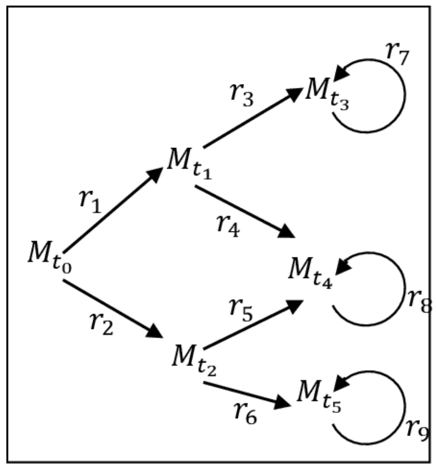

Fig. 1 A miniature system describing dynamic multisets M_{t_i}'s in a biochemical system

In Fig. 1, r_is represent the rate of reaction with respect to time, M'_{t_i}s are dynamic multisets and $t_i \in T$ represent different period time. Noticeable is the fact that the rate of reaction is a time dependent parameter.

4.2. **Ubiquitous systems.** Data duplication is ubiquitous. In certain online systems, as users upload the same files on the web or download the same files from the web into their storage facility, multiple copies of data are stored or accumulated over time. As said earlier, dynamic multisets have substantial potential in handling data duplication records within large databases. Their application can be envisaged for the case of a fast growing system $\mathcal{F}_T$ with growing rate (increasing or decreasing) of data duplication, r. Due to storing and transferring of duplicate data, this system can be conveniently analysed by deploying r which provides us with consistent information about copies of duplicate data at various periods within the time intervals T:

$$r = \sum_{\substack{1 \leq i \leq n \\ M_t \in \mathcal{F}_T}} \max\Big\{0, (C_{M_{t_n}}(s) - C_{M_{t_m}}(s))\Big\}_i, t \in T, n > m, \forall \quad s \in S_t. \tag{4.1}$$

In the right hand side of Eq. 4.1,

$$\max\Big\{0, (C_{M_{t_n}}(s) - C_{M_{t_m}}(s))\Big\}_i$$

represents the count of the difference of consecutive dynamic multisets M_{t_m} and M_{t_n} at various transitional periods i, $1 \leq i \leq n$.

The difference of two dynamic multisets is defined by

$$C_{M_t - M'_t}(s) = \Big\{\frac{\max\{0, (C_{M_t}(s) - C_{M'_t}(s))\}}{s}\Big| s \in S_t \quad \text{and} \quad \text{fixed time} \quad t \in T\Big\}. \tag{4.2}$$

The transition periods $i, 1 \leq i \leq n$, depict the time-varying change in copies of duplicate data at different time intervals.

5. Discussions

Very frequently, duplicate data copies are made for the purpose of reliability, preservation, and performance (i.e., duplicate data copies are made against inadvertent deletion or corruption of file data). However, in a fast growing system, since duplicate data can occupy a substantial portion of a storage system, its optimization may address the problem of having more efficient storage resources versus making duplicate data copies against inadvertent deletion or corruption of file data. In general, a problem of finding a solution to the following optimization problem is at hand:

$$\arg\min_{(\leftarrow,\rightarrow)} \sum_{\substack{1 \leq i \leq n \\ M_t \in \mathcal{F}_T}} \max\Big\{0, (C_{M_{t_n}}(s) - C_{M_{t_m}}(s))\Big\}_i, t \in T, n > m, \forall \quad s \in S_t, \tag{5.1}$$

where arg denotes the argument, namely, a pair of insert and remove operations, $(\leftarrow, \rightarrow)$, that minimizes the rate r at which files are added or deleted from the system.

In view of the above mentioned problem, there are several optimization methods that come in handy in an attempt to resolve it. However, this problem remains open. We hope that resolving the aforesaid problem may be useful in system (i.e., biological or computational) control. We will not concern ourselves to the discussion of its usefulness in system control in this paper. However, since the construct has potential usefulness in modelling the dynamics of biochemical reactions [2], we envision its application.

6. CONCLUSION

We have elaborately studied the concept of dynamic multisets and discussed its operations. The concept may be viewed as being induced by time-dependent change in copies of duplicate data in the course of *storing* and *transferring* various files in ubiquitous systems. This study provides a starting point for which it is possible to construct a consistent model which captures the transition of multiplicity of duplicate data in various storage tanks or systems. Dynamic multisets could be especially important in all situations where data (object) duplication or deduplication records become necessary for analyzing system performance or for other update record purposes. Nonetheless, it is important to note that, further study is needed to tackle the problem posed in section 5. The notion of dynamic multiset could be extended to soft computing.

7. ACKNOWLEDGEMENTS

The authors would like to thank Professor Emeritus D. Singh of the Department of Mathematics, Ahmadu Bello University, Zaria, Nigeria for his technical comments on the first draft of this paper. Also the first author sends special thanks to Patience Bankong for her encouragement throughout this work.

REFERENCES

[1] A. Alexandru and G. Ciobanu. Mathematics of multisets in the Fraenkel-Mostowski framework, Bull. Math. Soc. Sci. Math. Roumanie, 58(1) (2015), 3-18.

[2] C. Brink. Some backgrounds on multisets, TR- ARP- 2/87, ANU, Canbera, Australia, 1987.

[3] I. O. Bercea and G. Even. Dynamic dictionaries for multisets and counting filters with constant time operations. In: Lubiw A., Salavatipour M. (eds) Algorithms and Data Structures. WADS 2021. Lecture Notes in Computer Science, vol 12808. Springer, Cham, 2021.

[4] A. Bronselaer, A. Hallez and G. De Tre′. A possibilistic view on set and multiset comparison, Control Cybernet., 38(2) (2009), 341-366.

[5] A. Coletta, R. Gori and L. Francessca. Approximating probabilistic behaviours of biological systems using abstract interpretation, Elect. Note Theoret. Comput. Sci., 229 (2009), 165-182.

[6] G. F. Clement. On multisets k-families, Discrete Math., 69 (1988), 153-164.

[7] P. A. Ejegwa. Synopsis of the notions of multisets and fuzzy multisets, Ann. Commun. Math., 2(2) (2019), 101-120.

[8] S. Ghilezan, J. Pantovic and G. Vojvodic. Binary relations and algebras on multisets, Publicat. De Linstitut Mathmatique, 95(109) (2014), 11-117.

[9] K. P. Girish and S. J. John. Rough Multisets and Information Multisystems, Adv. Decision Sci., (2011), https://doi.org/10.1155/2011/495392.

[10] S. Hoskova-Mayerova and B. O. Onasanya. Results on functions on Dedekind multisets, Symmetry, 11(1125) (2019), https://doi.org/10.3390/sym11091125.

[11] A. M. Ibrahim, D. Singh and J. N. Singh. An outline of multiset space algebra, Int. J. Algebra, 5(31) (2011), 1515-1525.

[12] D. E. Knuth. The art of computer programming, Vol. 2 (Second Edition), Seminumerical Algorithm, Addison-Wesley, Reading, Massachusetts, 1981.

[13] A. Pagh, R. Pagh and S. Rao. An optimal bloom filter replacement, Seminar on Data Structures, 2004, p 1-7.

[14] A. B. Petrovesky. Structuring techniques in multiset spaces, Springer, 1997, p 174-184.

[15] D. Singh, A. M. Ibrahim, T. Yohanna and J. N. Singh. An overview of the application of multisets. Novi Sad J. Math., 37 (2007), 73-92.

[16] D. Singh and C. M. Peter. Multiset-based tree model for membrane Computing, Comput. Sci. J. Moldova, 19(1) (2011), 3-28.

[17] D. Singh and J. N. Singh. Some combinatorics of multisets, Int. J. Math. Edu. Sci. Tech., 34(4) (2003), 489–499.
[18] N. J. Wildberger. A new look at multiset, School of Mathematics, UNSW Sydney 2052, Australia, 2003.

T. O. William-West
Department of Mathematics, Ahmadu Bello University, Zaria, Nigeria
Email address: westtamuno@gmail.com

P. A. Ejegwa
Department of Mathematics, University of Agriculture, P.M.B. 2373, Makurdi, Nigeria
Email address: ejegwa.augustine@uam.edu.ng

A. U. Amaonyeiro
Department of Mathematics, University of Agriculture, P.M.B. 2373, Makurdi, Nigeria
Email address: ucheanslem@gmail.com