# FRÉCHET METRIC FOR SPACE OF BINARY CODED SOFTWARE

## Renáta MASÁROVÁ

RNDr. Renáta Masárová, PhD., Slovak University of Technology, Faculty of Materials Science and Technology in Trnava, Institute of Applied Informatics, Automation and Mathematics, Hajdóczyho 1, 917 24 Trnava, Slovakia
renata.masarova@stuba.sk

### Abstract

*As stated in (7), binary coded computer programs can be shown as a metric space. Therefore, they can be measured by metric in a sense of metric space theory. This paper presents the proof that Fréchet metric is a metric on the space of all sequences of elements $M = \{0,1,t\}$. Therefore, it is usable to build a system of software metrics based on the metric space theory.*

### Key words

*software metric, metric space, Fréchet metric*

### Introduction

Software measurement is one of the key tools used in software development. Different approaches to this topic have been described in e.g. (2, 4, 5, 8, 9). The new requirements necessitated the search for new instruments which could be used to measure the software. In recent years, attention was directed to the field of metric spaces as well. Research has produced some promising results. The metric spaces are used in such areas as proximity (11, 10) or searching (3, 6).
We divide the metrics into the following two groups:
- Syntactic – based on the syntactic or structural characteristics of particular parts of software;
- Semantic – based on the knowledge about concept on different levels of abstraction.

The syntactic metrics are based on investigating a particular property or structure of the software on particular level.

We classify as semantic those metrics which are based on some knowledge about meaning of the measured elements. The same metric can be seen as semantic and as syntactic from various points of view. An example of such situation is the measurement of procedures. If the procedure is understood only as a group of statements, we refer to the structure of the

code and we see all the procedures as the same type of element with different values of its properties. However, if it represents the implemented function, we talk about its semantics and each procedure represents different function. This view combines the concept on the code level (group of statements) and on the functional level (behaviour of the procedure). The more detailed view of these issues is provided in (7); the article discusses Baire metric.

This paper is focused on the Fréchet metric. The metric attaches greater importance to the beginning of sequences. This is a desirable feature if we examine the instructions for the processor. Instructions of the particular processor are formed in tuple with fixed number of bits. It is not quite accurate, but in terms of metrics, we can supplement the shorter instructions to the desired length. When expressed in binary form, operation is usually at the beginning of that instruction and the parameters follow it. However, they cannot be neglected.

Thus, we can assess whether the properties of Fréchet metric are sufficient for measuring software and the possibility of its use. We must first show that it is the metric in this particular instance.

Let $X$ be a set, $x, y, z \in X$ and $d(x, y)$ a function $d : X \times X \rightarrow R$ which satisfy these properties:

a) $d(x, y) \geq 0$ and $d(x, y) = 0 \Leftrightarrow x = y$

b) $d(x, y) = d(y, x)$                                                    [1]

c) $d(x, z) \leq d(x, y) + d(y, z)$

then the set $X$ with function $d(x, y)$ is called a metric space and is denoted as $(X, d)$. The function $d(x, y)$ is called a metric of this space.

If we replace a) by a*), where

a*) $d(x, y) \geq 0$ and if $x = y$, then $d(x, y) = 0$,

this space is called pseudo-metric.

**Fréchet metric for space of binary coded software**

Let $S$ be the set of all software solutions. Let $S_p$ be a set of all software which can be binary coded for a particular binary processor $p$ and let $M = \{0, 1, t\}$. We create a function $F$ which represents software $f : S_p \rightarrow M$.

Let $M = \{0, 1, t\}$, where $t = \dfrac{1}{2}$. A set all sequences of elements of set $M$ is denoted as $X$.

On $X \times X$, we define a function $d(x, y) = \sum\limits_{n=1}^{\infty} \dfrac{1}{2^n} \cdot \dfrac{|x_n - y_n|}{1 + |x_n - y_n|}$ (1, 4). We will prove that this function is a metric of $X \times X$, and therefore it has the following properties [1].

a) From the definition, it is evident that $d(x, y) \geq 0$.

If $d(x, y) = \sum\limits_{n=1}^{\infty} \dfrac{1}{2^n} \cdot \dfrac{|x_n - y_n|}{1 + |x_n - y_n|} = 0$, then $|x_n - y_n| = 0$ (because $\forall n \in N$ applies

$\dfrac{1}{2^n} \cdot \dfrac{|x_n - y_n|}{1 + |x_n - y_n|} \geq 0$). It follows $x_n = y_n$, and therefore $x = y$.

b) Obviously because $\left|x_n - y_n\right| = \left|y_n - x_n\right|$.

c) Let $x, y, z \in X$, $\quad x = (x_1, x_2,..., x_n,...)$, $y = (y_1, y_2,..., y_n,...)$, $z = (z_1, z_2,..., z_n,...)$. We

denote $\quad n_x = \min\{n; x_n = t\}, n_y = \min\{n; y_n = t\}, n_z = \min\{n; z_n = t\}$ and

$k = \max\{n_x, n_y, n_z\}$.

It is sufficient to prove the

inequality $\displaystyle\sum_{n=1}^{k} \frac{1}{2^n} \cdot \frac{\left|x_n - z_n\right|}{1 + \left|x_n - z_n\right|} \le \sum_{n=1}^{k} \frac{1}{2^n} \cdot \frac{\left|x_n - y_n\right|}{1 + \left|x_n - y_n\right|} + \sum_{n=1}^{k} \frac{1}{2^n} \cdot \frac{\left|y_n - z_n\right|}{1 + \left|y_n - z_n\right|}$, t. j.

$$\sum_{n=1}^{k} \frac{\left|x_n - z_n\right|}{1 + \left|x_n - z_n\right|} \le \sum_{n=1}^{k} \left( \frac{\left|x_n - y_n\right|}{1 + \left|x_n - y_n\right|} + \frac{\left|y_n - z_n\right|}{1 + \left|y_n - z_n\right|} \right).$$

We denote $A_n = \dfrac{\left|x_n - z_n\right|}{1 + \left|x_n - z_n\right|}$, $B_n = \dfrac{\left|x_n - y_n\right|}{1 + \left|x_n - y_n\right|}$, $C_n = \dfrac{\left|y_n - z_n\right|}{1 + \left|y_n - z_n\right|}$.

We will prove that for every $n = 1, 2,..., k$, it is true that $A_n \le B_n + C_n$.

We will break it down to individual cases.

$x_i = y_i = z_i$ ... obvious

$x_i = z_i$, then $A_n = 0$, $B_n + C_n \ge 0$ obvious

We only have to test the rest of the cases

$x_i = 1, y_i = 0, z_i = 0$ ... $\dfrac{1}{1+1} \le \dfrac{1}{1+1} + \dfrac{0}{1+0}$

$x_i = 1, y_i = 1, z_i = 0$ ... $\dfrac{1}{1+1} \le \dfrac{0}{1+0} + \dfrac{1}{1+1}$

$x_i = 1, y_i = t, z_i = 0$ ... $\dfrac{1}{1+1} \le \dfrac{\frac{1}{2}}{1 + \frac{1}{2}} + \dfrac{\frac{1}{2}}{1 + \frac{1}{2}} \Rightarrow \dfrac{1}{2} \le \dfrac{2}{3}$

$x_i = 1, y_i = 0, z_i = t$ ... $\dfrac{\frac{1}{2}}{1 + \frac{1}{2}} \le \dfrac{1}{1+1} + \dfrac{\frac{1}{2}}{1 + \frac{1}{2}} \Rightarrow \dfrac{1}{3} \le \dfrac{5}{6}$

$x_i = 1, y_i = 1, z_i = t$ ... $\dfrac{\frac{1}{2}}{1 + \frac{1}{2}} \le \dfrac{0}{1+0} + \dfrac{\frac{1}{2}}{1 + \frac{1}{2}}$

$x_i = 1, y_i = t, z_i = t$ ... $\dfrac{\frac{1}{2}}{1 + \frac{1}{2}} \le \dfrac{\frac{1}{2}}{1 + \frac{1}{2}} + \dfrac{0}{1+0}$

The rest can be proved by using similar methods.

19

The third property is satisfied as well – triangular inequality that is the function of

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{|x_n - y_n|}{1 + |x_n - y_n|}$$ on a given space is a metric.

## Discussion

The first two properties of a metric are evidently satisfied for every constant $t$. While proving the third property of a metric – triangular inequality – we simplified it using $t = \frac{1}{2}$. Given that this symbol only serves for transforming a finite sequence of ones and zeroes to an infinite one, it is not a limitation.

The proof allows the use of Fréchet metric under the conditions that were described in (7) as well.

## Conclusion

We showed that Fréchet metric is a metric for binary sequences tailed by value $t = \frac{1}{2}$. Therefore, it can be used to measure software. The mentioned metric is usable during a work on the low level activities of the software engineering. If we compare the signatures of routines (7) and their body (function), they have a similar relationship as operations and attributes. Therefore, we assume that the metric with similar characteristics can also be used on this level of software examination.

## Acknowledgement

**References:**

1. AGNEW, R. P. 1940. On rearrangement of series. *Bull. Amer. Soc.*, Vol. 46, pp. 797–799.
2. BRIAND, L.C., MORASCA, S., BASILI, V. R. 1996. Property-based software engineering measurement. *Software Engineering, IEEE Transactions on*, Vol. 22, issue 1, pp. 68-86. ISSN 0098-5589
3. CHAVEZ, E., NAVARRO, G., BAEZA-YATES, R., MARROQUIN, J. 2001. Searching in metric spaces. *ACM Comp. Surv.*, Vol. 33, No. 3, pp. 273–321.
4. EDGAR, G. 2008. *Measure, Topology and Fractal Geometry*. Springer.
5. S. HENRY, D. KAFURA. 1981. Software structure metrics based on information flow. *IEEE Transactions on Software Engineering*, Vol. 7, No. 5, pp. 510-518.
6. KLETTE, R., ROSENFELD, A. 2004. *Digital Geometry*. San Francisco, Morgan Kaufmann.

7. MASÁROVÁ, R., MASÁR, A. 2012. Available Metrics for Space of Binary Coded Software. In: *Mechatronika 2012*. Prague: Czech Technical University in Prague, pp. 373-376. ISBN 978-80-01-04985-3

8. MILLS, E. Software Metrics. *DTIC Online*. Defense Technical Information Center, December 1998, Web. 8 September 2012.

9. ROSENBERG, L.H., THEODORE, P., HAMMER, P. F. HUFFMAN, L. L. 1998. Requirements, Testing and Metrics. In: *Proceedings of the 15th Annual Pacific Northwest Software Quality Conference*.

10. WHALEN, M. W., RAJAN, A., HEIMDAHL, M. P. E., MILLER, S. P. 2006. Coverage metrics for requirements-based testing. In: *Proceedings of the 2006 international symposium on software testing and analysis (ISSTA '06)*. New York: ACM, pp. 25-36.

11. ZEZULA, P., AMATO, G., DOHNAL, V., BATKO, M. 2006. *Similarity Search: The Metric Space Approach, ser. Advances in Database Systems.* Springer, Vol. 32.

**Reviewers:**

doc. Mgr. Elena Pivarčiová, PhD.
doc. RNDr. Oleg Palubíny, CSc.