Contents

Methodological papers

ANNALES MATHEMATICAE ET INFORMATICAE 46. (2016)

ANNALES MATHEMATICAE ET INFORMATICAE

TOMUS 46. (2016)

HUNGARIA, EGER

ANNALES MATHEMATICAE ET INFORMATICAE

International journal for mathematics and computer science

Referred by
Zentralblatt für Mathematik
and
Mathematical Reviews

The journal of the Institute of Mathematics and Informatics of Eszterházy Károly University of Applied Sciences is open for scientific publications in mathematics and computer science, where the field of number theory, group theory, constructive and computer aided geometry as well as theoretical and practical aspects of programming languages receive particular emphasis. Methodological papers are also welcome. Papers submitted to the journal should be written in English. Only new and unpublished material can be accepted.

Authors are kindly asked to write the final form of their manuscript in LaTeX. If you have any problems or questions, please write an e-mail to the managing editor Miklós Hoffmann: `hofi@ektf.hu`

The volumes are available at `http://ami.ektf.hu`

ANNALES MATHEMATICAE ET INFORMATICAE

VOLUME 46. (2016)

INSTITUTE OF MATHEMATICS AND INFORMATICS
ESZTERHÁZY KÁROLY UNIVERSITY OF APPLIED SCIENCES
HUNGARY, EGER

HU ISSN 1787-5021 (Print)
HU ISSN 1787-6117 (Online)

A kiadásért felelős az
Eszterházy Károly Egyetem rektora
Megjelent a Líceum Kiadó gondozásában
Kiadóvezető: Grebely Gergely
Felelős szerkesztő: Zimányi Árpád
Műszaki szerkesztő: Tómács Tibor
Megjelent: 2016. december   Példányszám: 30

Készítette az
Eszterházy Károly Egyetem nyomdája
Felelős vezető: Kérészy László

Annales Mathematicae et Informaticae
46 *(2016) pp. 3–12*
http://ami.ektf.hu

Statistics in words and partitions of a set

Walaa Asakly

Department of Mathematics, University of Haifa, Haifa, Israel
walaa_asakly@hotmail.com

Submitted December 30, 2015 — Accepted September 12, 2016

Abstract

Let $[k] = \{1, 2, \ldots, k\}$ be an alphabet over k letters. A *word* ω of length n over alphabet $[k]$ is an element of $[k]^n$ and is also called a k-ary word of length n. We say that ω contains an ℓ-peak, if it exists an i such that $2 \leq i \leq n - \ell$ where $\omega_i = \omega_{i+1} = \cdots = \omega_{i+\ell-1}$ and $\omega_{i-1} < \omega_i$ and $\omega_{i+\ell-1} > \omega_{i+\ell}$. A *partition* Π of set $[n]$ of size k is a collection $\{B_1, B_2, \ldots, B_k\}$ of non empty disjoint subsets of $[n]$, called *blocks*, whose union equals $[n]$. In this paper, we find an explicit formula for the generating function for the number of words of length n over alphabet $[k]$ according to the number of ℓ-peaks in terms of Chebyshev polynomials of the second kind. As a consequence of the results obtained for words, we finally find the number of ℓ-peaks in set partitions of $[n]$ with exactly k blocks.

Keywords: Set partitions, words, ℓ-peak, Chebyshev polynomials of the second kind

MSC: 05A05

1. Introduction

Words

Let $[k] = \{1, 2, \ldots, k\}$ be an alphabet over k letters. A *word* ω of length n over alphabet $[k]$ is an element of $[k]^n$ and is also called a word of length n on k letters or a k-ary word of length n. The number of the words of length n over alphabet $[k]$ is k^n. Similar statistics in patterns of subwords have been widely studied in the literature (see [2]). For example, Kitaev, Mansour and Remmel [3] enumerated the number of rises (respectively, levels and falls) which are subword patterns 12, (respectively, 11 and 21) in words that have a prescribe first element. Heubach and Mansour [2] enumerated the number of words of length n over alphabet $[k]$ that contain the subword pattern 111 and the subword pattern 112 exactly r times. Burstein and Mansour [1] generalized the result to subword pattern of length ℓ. More recently, Mansour [4] enumerated the number of peaks (subword patterns 121, 132 or 231) and valleys (subword patterns 212, 213 or 312) in words of length n over alphabet $[k]$. Our aim is to extend this result to patterns of arbitrary length. We say that ω contains an *ℓ-peak*, if exists $2 \leq i \leq n-\ell$ such that $\omega_i = \omega_{i+1} = \cdots = \omega_{i+\ell-1}$ and $\omega_{i-1} < \omega_i$ and $\omega_{i+\ell-1} > \omega_{i+\ell}$. For example, the word $12^4 13^4 2 = 12222133332$ in $[3]^{11}$ contains two 4-peaks, namely 122221 and 133332.

Set partitions

A *partition* Π of set $[n]$ with exactly k blocks is a collection $\{B_1, B_2, \ldots, B_k\}$ of non empty disjoint subsets of $[n]$ whose union is equal to $[n]$. We assume that blocks are listed in increasing order of their minimal elements, that is, $minB_1 < minB_2 < \cdots < minB_k$. We denote the set of all partitions of $[n]$ with exactly k blocks to be $P_{n,k}$. The number of all partitions of $[n]$ with k blocks is $S(n,k)$, these are the Stirling numbers of the second kind [9]. We denote the set of all partitions of $[n]$ to be P_n, namely $P_n = \cup_{k=0}^{n} P_{n,k}$. The number of all partitions of $[n]$ is $B_n = \sum_{k=0}^{n} S_{n,k}$, which is the n-th Bell number. Any partition Π can be written as $\pi_1\pi_2\cdots\pi_n$, where $i \in B_{\pi_i}$ for all i, and this form is called the *canonical sequential form*. For example $\Pi = \{\{12\},\{3\},\{4\}\}$ is a partition of [4], the canonical sequential form is $\pi = 1123$. Several authors have studied different statistics on P_n (see [4]). For instance, Mansour and Munagi [6] found the generating function for the number of partitions of $[n]$ according to rises, descents and levels, they also computed the total number of t-rises (respectively, t-descents and t-levels), this is a increasing subword pattern of size t (respectively, decreasing subword pattern of size t, fixed subword pattern of size t), see [5]. A lot of attention has been given to the statistics on $P_{n,k}$ (see [4]). For example, Shattuck [8] counted the rises, descents and levels in the set partition of $[n]$ with exactly k blocks. In addition, Mansour [4] found an explicit formula for the generating functions for the number of set partition of $[n]$ with exactly k blocks according to the statistics ℓ-rise (respectively, ℓ-descent and ℓ-level). Mansour and Shattuck [7] found an explicit formula for the generating function of set partitions of n with exactly $[k]$ blocks according to the number of peaks (valleys). Our aim is to extend this result for the set $P_{n,k}$ according to the number of ℓ-peaks.

In this paper, we find the generating function of the words of length n over alphabet $[k]$ according to the number of ℓ-peaks. We also compute the total number of ℓ-peaks in the words of length n over alphabet $[k]$. As a consequence of these results, we find the number of ℓ-peaks in set partitions of $[n]$ with exactly k blocks.

2. Words and partitions of a set according to multi statistics ℓ-peaks

Let $W_k(x, q_1, \dots, q_\ell)$ be the generating function for the number of words of length n over alphabet $[k]$ according to the number of ℓ-peaks, namely,

$$W_k(x, q_1, \dots, q_\ell) = \sum_{n \geq 0} x^n \sum_{\omega \in [k]^n} \prod_{i=1}^{\ell} q_i^{i-peak(\omega)}.$$

Lemma 2.1. *The generating function $W_k(x, q_1, \dots, q_\ell)$ satisfies the recurrence relation*

$$W_k(x, q_1, \dots, q_\ell) = \frac{A_\ell - xB_\ell + W_{k-1}(x, q_1, \dots, q_\ell)(B_{\ell+1} - A_\ell)}{(1-x)(1+A_\ell) + x^{\ell+1} - W_{k-1}(x, q_1, \dots, q_\ell)\big((1-x)A_\ell + x^{\ell+1}\big)},$$

where $A_\ell = \sum_{i=1}^{\ell} x^i q_i$ and $B_\ell = \frac{1-x^\ell}{1-x}$.

Proof. It is obvious

$$W_k(x, q_1, \dots, q_\ell) = W_{k-1}(x, q_1, \dots, q_\ell) + W_k^\dagger(x, q_1, \dots, q_\ell), \tag{2.1}$$

where $W_k^\dagger(x, q_1, \dots, q_\ell)$ is the generating function for the number of words ω of length n over alphabet $[k]$ according to the number of ℓ-peaks such that ω contains at least one occurrence of the letter k. A word ω that contains a letter k can be decomposed as either

(1) k;

(2) $k\omega'$, where ω' is a non empty word over $[k]$;

(3) $\omega'' k^i \omega'''$, where k^i denotes a word $kk \cdots k$ with exactly i letters, ω'' is a non empty word over $[k-1]$ and ω''' is a non empty word over $[k]$ which starts with a letter $a \neq k$, for $1 \leq i \leq \ell$;

(4) $\omega'' k^i$, for $1 \leq i \leq \ell$; or

(5) $\omega'' k^{\ell+1} \omega''''$, where ω'''' is a word over $[k]$.

The corresponding generating functions of these decomposition are

(1) x;

(2) $x(W_k(x, q_1, \dots, q_\ell) - 1)$;

(3) $q_i x^i (W_{k-1}(x, q_1, \dots, q_\ell) - 1)(W_k(x, q_1, \dots, q_\ell)(1-x) - 1)$, for $1 \leq i \leq \ell$;

(4) $x^i (W_{k-1}(x, q_1, \dots, q_\ell) - 1)$, for $1 \leq i \leq \ell$; or

(5) $x^{\ell+1}(W_{k-1}(x,q_1,\dots,q_\ell)-1)W_k(x,q_1,\dots,q_\ell)$,

respectively. Hence, by (2.1), we obtain

$$\begin{aligned}
&W_k(x,q_1,\dots,q_\ell)\\
&=W_{k-1}(x,q_1,\dots,q_\ell)+x+x(W_k(x,q_1,\dots,q_\ell)-1)\\
&\quad+\sum_{i=1}^{\ell}q_ix^i(W_{k-1}(x,q_1,\dots,q_\ell)-1)(W_k(x,q_1,\dots,q_\ell)(1-x)-1)\\
&\quad+\sum_{i=1}^{\ell}x^i(W_{k-1}(x,q_1,\dots,q_\ell)-1)\\
&\quad+x^{\ell+1}(W_{k-1}(x,q_1,\dots,q_\ell)-1)W_k(x,q_1,\dots,q_\ell),
\end{aligned}$$

which equivalent to

$$\begin{aligned}
&W_k(x,q_1,\dots,q_\ell)\\
&=\frac{A_\ell-xB_\ell+W_{k-1}(x,q_1,\dots,q_\ell)(B_{\ell+1}-A_\ell)}{(1-x)(1+A_\ell)+x^{\ell+1}-W_{k-1}(x,q_1,\dots,q_\ell)\Big((1-x)A_\ell+x^{\ell+1}\Big)},
\end{aligned}\tag{2.2}$$

where $A_\ell=\sum_{i=1}^{\ell}x^iq_i$ and $B_\ell=\frac{1-x^\ell}{1-x}$. □

We plan to find an explicit formula for the generating function $P_k(x,q_1,\dots,q_\ell)$ for the number of partitions of n with exactly k blocks according to the number of ℓ-peaks.

$$P_k(x,q_1,\dots,q_\ell)=\sum_{n\geq 0}x^n\sum_{\pi\in P_{n,k}}\prod_{i=1}^{\ell}q^{i-peak(\pi)}.$$

To do that we will use Lemma 2.1.

Theorem 2.2. *For all $k\geq 1$,*

$$\begin{aligned}
&P_k(x,q_1,\dots,q_\ell)\\
&=\prod_{j=1}^{k}\left(\sum_{i=1}^{\ell}x^i(q_i(W_j(x,q_1,\dots,q_\ell)(1-x)-1)+1)+x^{\ell+1}W_j(x,q_1,\dots,q_\ell)\right).
\end{aligned}$$

Proof. Any partition π of $[n]$ with exactly k blocks can be decomposed either

(1) $\pi k^i\pi'$, πk^i, for $1\leq i\leq \ell$, where π is a set partition with exactly $k-1$ blocks, π' is a non empty word over alphabet $[k]$ which starts with a letter $a<k$; or

(2) $\pi k^{\ell+1}\pi''$, where π'' is a word over alphabet $[k]$.

The corresponding generating functions are

(1)

$$q_i x^i P_{k-1}(x, q_1, \ldots, q_\ell)(W_k(x, q_1, \ldots, q_\ell) \\ - xW_k(x, q_1, \ldots, q_\ell) - 1) + x^i P_{k-1}(x, q_1, \ldots, q_\ell),$$

for $1 \le i \le \ell$;

(2) $x^{\ell+1} P_{k-1}(x, q_1, \ldots, q_\ell) W_k(x, q_1, \ldots, q_\ell)$,

respectively. By summing all the last terms we obtain

$$\begin{aligned} &P_k(x, q_1, \ldots, q_\ell) \\ &= \sum_{i=1}^{\ell} q_i x^i P_{k-1}(x, q_1, \ldots, q_\ell)(W_k(x, q_1, \ldots, q_\ell) - xW_k(x, q_1, \ldots, q_\ell) - 1) \\ &\quad + \sum_{i=1}^{\ell} x^i P_{k-1}(x, q_1, \ldots, q_\ell) + x^{\ell+1} P_{k-1}(x, q_1, \ldots, q_\ell) W_k(x, q_1, \ldots, q_\ell) \\ &= P_{k-1}(x, q_1, \ldots, q_\ell) \cdot \\ &\quad \cdot \left(\sum_{i=1}^{\ell} x^i (q_i (W_j(x, q_1, \ldots, q_\ell)(1-x) - 1) + 1) + x^{\ell+1} W_j(x, q_1, \ldots, q_\ell) \right). \end{aligned}$$

Thus, by induction on k together with the initial condition $P_0(x, q) = 1$, we complete the proof. □

Example 2.3. Using the recursion given in Theorem 2.2, we may obtain the generating function for the number of partitions of $[n]$ with exactly k blocks,

$$\begin{aligned} P_k(x, 1, \ldots, 1) &= \prod_{j=1}^{k} \sum_{i=1}^{\ell} x^i W_j(x, 1, \ldots, 1)(1-x) + x^{\ell+1} W_j(x, 1, \ldots, 1) \\ &= \prod_{j=1}^{k} \left(x \frac{1-x^\ell}{1-x} \frac{1}{1-jx}(1-x) + x^{\ell+1} \frac{1}{1-jx} \right) \\ &= x^k \prod_{j=1}^{k} \frac{1}{1-jx}, \end{aligned}$$

which is in accord with the well-known the generating function for the number of partitions of $[n]$ with exactly k blocks.

Example 2.4. By substituting $\ell = 1$ and $q_1 = q$ in Lemma 2.1, we get $W_k(x, q)$ the generating function for the number of words of length n over the alphabet $[k]$ according to the number of peaks (peak of length one), which gives the following recursion

$$W_k(x, q) = \frac{x(q-1) + (1 - x(q-1))W_{k-1}(x, q)}{1 - x(1-q)(1-x) - x(x + q(1-x))W_{k-1}(x, q)}.$$

By using the same substitution in Theorem 2.2, we obtain the recurrence relation for the generating function for the number of set partitions $P_{n,k}$ according to the number of peaks (peak of length one), which gives the following recursion

$$P_k(x,q) = x^k \prod_{j=1}^{k} (1 + x(1-q)W_j(x,q) + q(W_j(x,q)-1)),$$

where the two above results agree with the results of Mansour and Shattuck (see [7]).

2.1. Counting ℓ-peaks in words and partitions of a set

Let $W_k(x,q)$ be the generating function for the number of words of length n over alphabet $[k]$ according to the number of ℓ-peaks.

$$W_k(x,q) = \sum_{n\geq 0} x^n \left(\sum_{\omega\in[k]^n} q^{\ell-peak(\omega)} \right).$$

Corollary 2.5. *The generating function $W_k(x,q)$ for the number of words of length n over alphabet $[k]$ according to the number of ℓ-peaks is*

$$W_k(x,q) = \frac{A + (1-A)W_{k-1}(x,q)}{a_\ell - xW_{k-1}(x,q)a_{\ell-1}} \tag{2.3}$$

where $A = x^\ell(q-1)$ and $a_\ell = 1 + x^\ell(q-1)(1-x)$, which is equivalent to

$$W_k(x,q) = \frac{x^\ell(q-1)(U_{k-1}(t) - U_{k-2}(t))}{U_k(t) - U_{k-1}(t) - (1 - x^\ell(q-1))(U_{k-1}(t) - U_{k-2}(t))}, \tag{2.4}$$

where $t = 1 + \frac{x^{\ell+1}}{2}(1-q)$ and U_m is the m-th Chebyshev polynomial of the second kind.

Proof. By substituting $q_i = 1$ for $i \neq \ell$, and $q_\ell = q$ in (2.2) we obtain (2.3). Then, by applying [Appendix D] [4] for (2.2), we obtain (2.4). □

Now, our aim is to find the total number of ℓ-peaks in all words of length n over alphabet $[k]$.

Lemma 2.6. *For all $k \geq 1$,*

$$\frac{d}{dq} W_k(x,q) \mid_{q=1} = \frac{x^{\ell+2}}{(1-kx)^2} \left(2\binom{k}{3} + \binom{k}{2} \right).$$

Proof. We compute the number of ℓ-peaks in all the words of length n over alphabet $[k]$. By differentiating (2.3) with respect to q, we obtain

$$V_k(x) = \frac{d}{dq} W_k(x,q) \mid_{q=1}$$

$$= \frac{(x^\ell(1-W_{k-1}(x,1))+V_{k-1}(x))(1-xW_{k-1}(x,1))}{(1-xW_{k-1}(x,1))^2}$$
$$- \frac{W_{k-1}(x,1)(x^\ell(1-x)(1-W_{k-1}(x,1))-xV_{k-1}(x))}{(1-xW_{k-1}(x,1))^2},$$

and using $W_k(x,1)=\frac{1}{1-kx}$ (easy to prove by induction), we obtain

$$\frac{d}{dq}W_k(x,q)\mid_{q=1}=\frac{x^{\ell+2}}{(1-kx)^2}\left(2\binom{k}{3}+\binom{k}{2}\right), \tag{2.5}$$

as claimed. □

By finding the coefficient of x^n in (2.5) we get the following result

Corollary 2.7. *The total number of ℓ-peaks in all the words of length n over alphabet $[k]$ is given by*

$$(n-1-\ell)k^{n-2-\ell}\left(2\binom{k}{3}+\binom{k}{2}\right).$$

We plan to find the explicit formula for the generating function $P_k(x,q)$ for the number of $P_{n,k}$ according to the number of ℓ-peaks.

$$P_k(x,q)=\sum_{n\geq 0}x^n\left(\sum_{\pi\in P_{n,k}}q^{\ell-peak(\pi)}\right).$$

Corollary 2.8. *For all $k\geq 1$, the generating function $P_k(x,q)$ is given by*

$$x^k\prod_{j=1}^{k}\big(W_j(x,q)(1+x^{\ell-1}(x-1))+x^{\ell-1}+qx^{\ell-1}(W_j(x,q)(1-x)-1)\big).$$

Proof. By substituting $q_i=1$ for $i\neq\ell$, and $q_\ell=q$ in Theorem (2.2). □

Lemma 2.9. *For all $k\geq 3$,*

$$\frac{d}{dq}P_k(x,q)\mid_{q=1}=\frac{x^{k+\ell}\binom{k}{2}}{(1-x)\cdots(1-kx)}+\frac{x^{k+\ell+2}}{(1-x)\cdots(1-kx)}\sum_{j=3}^{k}\frac{2\binom{j}{3}+\binom{j}{2}}{(1-jx)}.$$

Proof. By Corollary 2.8, we have

$$\frac{d}{dq}P_k(x,q)\mid_{q=1}=P_k(x,1)\sum_{j=1}^{k}\lim_{q\to 1}\left(\frac{\frac{d}{dq}L_j(q)}{L_j(q)}\right), \tag{2.6}$$

where

$$L_j(q)=(W_j(x,q)(1+x^{\ell-1}(x-1))+x^{\ell-1}+qx^{\ell-1}(W_j(x,q)(1-x)-1).$$

Note that

$$\lim_{q\to1}\frac{d}{dq}L_j(q)=\lim_{q\to1}\left(\frac{d}{dq}W_j(x,q)+x^{\ell-1}(W_j(x,q)(1-x)-1)\right)$$
$$=\frac{x^\ell\left((j-1)-(j-1)jx+(2\binom{j}{3}+\binom{j}{2})x^2\right)}{(1-jx)^2}=x^\ell\left(\frac{j-1}{1-jx}+\frac{(2\binom{j}{3}+\binom{j}{2})x^2}{(1-jx)^2}\right).$$

Hence, by using (2.6) we obtain

$$\frac{d}{dq}P_k(x,q)\mid_{q=1}=\frac{x^{k+\ell}}{(1-x)\cdots(1-kx)}\sum_{j=1}^{k}\left(j-1+\frac{(2\binom{j}{3}+\binom{j}{2})x^2}{(1-jx)}\right)$$
$$=\frac{x^{k+\ell}\binom{k}{2}}{(1-x)\cdots(1-kx)}+\frac{x^{k+\ell+2}}{(1-x)\cdots(1-kx)}\sum_{j=3}^{k}\frac{2\binom{j}{3}+\binom{j}{2}}{(1-jx)},$$

as required. □

By using the facts that $P_k(x,1)=\sum_{n\geq1}S_{n,k}x^n$ and $\sum_{j=1}^{k}(j-1)x^\ell=x^\ell\binom{k}{2}$, together with Lemma 2.9 we get the following corollary.

Corollary 2.10. *The total number of the ℓ-peaks in all set partitions $P_{n,k}$ is given by*

$$\binom{k}{2}S_{n-\ell,k}+\sum_{i=\ell+2}^{n-k}S_{n-i,k}\sum_{j=3}^{k}j^{i-\ell-2}\left(2\binom{j}{3}+\binom{j}{2}\right).$$

2.2. Applications

By substituting $\ell=2$ in Corollary 2.7, we obtain the following result

Corollary 2.11. *The total number of the 2-peaks in all the words of length n over alphabet $[k]$ is given by*

$$(n-3)k^{n-4}\left(2\binom{k}{3}+\binom{k}{2}\right).$$

By substituting $\ell=2$ in Corollary 2.10, this leads to

Corollary 2.12. *The total number of the 2-peaks in all set partitions $P_{n,k}$ is given by*

$$\binom{k}{2}S_{n-2,k}+\sum_{i=4}^{n-k}S_{n-i,k}\sum_{j=3}^{k}j^{i-4}\left(2\binom{j}{3}+\binom{j}{2}\right).$$

By substituting $q=0$ in (2.4), we obtain that the generating function for the number of words of length n over alphabet $[k]$ without ℓ-peaks is given by

$$W_k(x,0)=\frac{-x^\ell(U_{k-1}(t)-U_{k-2}(t))}{U_k(t)-U_{k-1}(t)-(1+x^\ell)(U_{k-1}(t)-U_{k-2}(t))},\tag{2.7}$$

where $t = 1 + \frac{x^{\ell+1}}{2}$ and U_m is m-th Chebyshev polynomial of the second kind. By substituting $q = 0$ in Corollary 2.8, we get

$$P_k(x,0) = x^k \prod_{j=1}^{k} \left(W_j(x,0)(1 + x^{\ell-1}(x-1)) + x^{\ell-1}\right), \tag{2.8}$$

by substituting (2.7) into (2.8), and using the relation $U_{j+1}(t) = 2tU_j(t) - U_{j-1}(t)$, we get

$$P_k(x,0) = x^k \prod_{j=1}^{k} \frac{U_{j-1}(t) - (1 + x^{\ell})U_{j-2}(t)}{(1-x)U_{j-1}(t) - U_{j-2}(t)},$$

where $t = 1 + \frac{x^{\ell+1}}{2}$, which is the generating function of $P_{n,k}$ without ℓ-peaks. By using the above result with $\ell = 1$, we obtain the same result of Mansour and Shattuck (see [7]).

Corollary 2.13. *The generating function for the number of set partitions of P_n without ℓ-peaks is given by*

$$1 + \sum_{k\geq 1} P_k(x,0) = \sum_{k\geq 0} x^k \prod_{j=1}^{k} \frac{U_{j-1}(t) - (1 + x^{\ell})U_{j-2}(t)}{(1-x)U_{j-1}(t) - U_{j-2}(t)},$$

where $t = 1 + \frac{x^{\ell+1}}{2}$ and U_m is the m-th Chebyshev polynomial of the second kind.

2.3. Conclusion

In the present paper, we determined the generating function for the number of k-ary words of length n according to the number of ℓ-peaks. Also, we determined the generating function for the number of set partitions of $[n]$ with exactly k blocks according to the number of ℓ-peaks. Seems our techniques can be extended to the case of compositions of n (a composition of n is a word $\sigma_1\sigma_2\cdots\sigma_m$ such that $\sum_{i=1}^{m} \sigma_i = n$), where we leave it to the interest reader.

Acknowledgment. The author expresses her appreciation to the referee for his/her careful reading of the manuscript.

References

[1] Burstein, A., Mansour, T., Counting occurrences of some subword patterns, *Discret Math. Theor. Comput. Sci.* 6(1) (2003) 1–11.

[2] Heubach S., Mansour, T., Combinatorics of compositions and words (Boca Raton), CRC Press, Boca Raton, 2010.

[3] Kitave, S., Mansour, T., Remmel, J.B., Counting descents, rises, and levels, with prescribe first element, in words, *Discrete Math. Theor. Comput. Sci.* 10(3) (2008) 1–22.

[4] Mansour, T., Combinatorics of set partitions (Boca Raton), CRC Press, Boca Raton, 2013.

[5] Mansour, T., Munagi, A.O., Enumeration of partitions by long rises, levels, and descents, *J. Integer Seq.* 12 (2009) Art. 9.1.8.

[6] Mansour, T., Munagi, A.O., Enumeration of partitions by rises, levels, and descents, in Permutation Patterns: London Mathematical Society, *Lect. Note Ser.* 376, Cambridge University Press, 2010.

[7] Mansour, T., Shattuck, M., Counting peaks and valleys in a partition of a set, *J. Integer Seq.* 13 (2010) Art. 10.6.8.

[8] Shattuck, M., Recounting the number of rises, levels and descentes in finite set partitions, *Integers* 10:2 (2010) 179–185.

[9] Stanley, R.P., Enumerative Combinatorics, Vol. 1, Cambridge University Press, Cambridge, UK, 1996.

Annales Mathematicae et Informaticae
46 (2016) pp. 13–36
http://ami.ektf.hu

Avatar-Based Sport Science Soccer Simulations

Norbert Bátfai[ab], András Mamenyák[ab], Péter Jeszenszky[ab],

Gergely Kövér[ab], Máté Smajda[ab], Renátó Besenczi[ab],

Béla Halász[a], György Terdik[ab], Márton Ispány[ab]

[a]SziMe3D Ltd., Debrecen, Hungary
[b]Department of Information Technology, University of Debrecen

Submitted November 2, 2015 — Accepted October 11, 2016

Abstract

In this paper, we propose a general framework for sport science simulations that we refer to as Simulation Oriented Architecture (SimOA). As a concrete implementation of the framework we present a collection of sport science simulators that we developed in the experimental industrial research and development project called "Football Avatar" for simulating soccer matches. The practical goal of performing such simulations is to help soccer teams with providing an effective tool for supporting tactical decision making. The paper also establishes a solid theoretical foundation for performing sport science simulations introducing the concept of avatars.

Keywords: Forecasting, Parallel Computing, Simulation, Soccer, Sports

MSC: 68U20

1. Introduction

Simulation techniques have been adopted in many fields of science for investigating the behavior of a real system by imitating it through connected artificial objects that exhibit a nearly identical behavior, at least in statistical sense, see the comprehensive book [20]. Simulation has also been contributed significantly to the progress of science, see [7] and provides an important methodological tool in information system research, see [31]. Simulation studies are considered particularly useful for complex systems which cannot be described by simple physical laws, mainly for understanding the behavior of humans on various domains of society, e.g., health care [24], economics and management [9], transportation [30], and sport [26]. The broad scope of the literature in which simulation is applied on these and other fields indicates its relevance.

The research gap to be addressed by this work is to provide a framework for sport science simulations and to establish solid theoretic foundations for that purpose. In order to realize these goals, in the framework of the FootballAvatar project [32] we have developed a novel collection of sport science soccer simulators. This paper presents some key aspects behind these simulators. The simulation algorithms and the technologies used in their implementations will be presented briefly. We call our approach for developing soccer simulators Simulation Oriented Architecture (SimOA) because the design of the developed software system is entirely organized around the logic of simulations. Although our work mainly focuses on soccer, we think that SimOA can be considered as a general approach for performing sport science simulations.

The main purpose of the research presented in this paper was to create a mathematical definition of avatars for sport simulations that was intuitively introduced in [1]. In the intuitive sense, football avatars are computerized abstractions of soccer players, coaches, and referees. However, we will see that the concept of avatars is not limited to be used for soccer only, it can be applied in the case of other sports, among others, ping-pong and tennis. The most important restriction on the simulations is that relevant probability properties both in the simulations and in reality must be the same. One of the most important results of the research is that we were able to develop such simulators.

The paper is organized as follows. The next section gives a general overview of soccer simulation algorithms. The third section introduces definitions for the concept of avatar for soccer simulations and illustrates them with examples. The fourth section discusses our soccer simulation algorithms in detail. The fifth section is dedicated to avatar transformations, biological and behavioural models incorporated into our simulations are also discussed here. Simulation computation results are presented in the sixth section. Finally, the seventh section concludes the paper.

2. Soccer Simulation Computations

2.1. Types of Simulation Models

Paper [4] presents a review of existing soccer simulation models and classifies them in the following three categories: (1) non-realistic, (2) quasi-realistic, (3) and realistic models. Simulation models classified in the first two classes are used in our research on sport science simulations as well as in software components of the FootballAvatar project. In the case of non-realistic models, we have typically used statistical simulation computations. Our core simulation computations are quasi-realistic where statistical methods are mainly used for validating simulation algorithms.

2.2. FootballAvatar Soccer Simulator Collection

The FootballAvatar Soccer Simulation Collection is based on the Simulation Oriented Architecture. The following three levels (or speeds) can be distinguished in the FootballAvatar Soccer Simulation Collection: (1) Level "-1" uses only publicly available or estimated data based on objective and/or subjective observations. (2) Level "0" uses dedicated equipment such as video cameras and sensors to gather data (this infrastructure is provided and operated by our partners). (3) Level "+" is built on the lower levels (see Sections 5.2. and 6.2.).

Software elements on each level can operate in the following three modes: (1) The standalone mode does not require any input at all. It is mainly used to generate test data. (2) The analyzer mode serves for analyzing test data or real soccer data. (3) The avatar simulation mode is the basis of the comparison of real and simulated soccer matches, it highly depends on real soccer data.

2.2.1. Statistical Simulation Computations

On level "-1", prediction tasks and related techniques that emerge in football science can be classified into the following classes:

- Predicting the outcomes of an indicator event or an event with a very small number of outcomes. The related techniques are logit, probit and other binary or ordered discrete regression models that contain different explanatory variables, see [14] and [12]. These models can be applied, e.g., to restrict forecasting directly to the match result, i.e., win, draw, or lose.

- Predicting a count type event like the number of faults, goals, or corners. The related techniques are Poisson distribution based models, general bivariate discrete distribution models using copulas, and different algorithms from the field of machine learning.

- Predicting a continuous variable such as the distance covered by players or the time of the possession of the ball by a team during a match. The related techniques are the standard methods of supervised learning to be used for a continuous target such as nonlinear regression and neural networks.

These are typically used in non-realistic simulations.

In Poisson distribution based data analysis, the dependent variable has one- or bivariate Poisson distribution, see [21] and [18]. This framework has been extended in [29] for the time-varying case. A possible application of these models is to forecast the number of goals at a match by bivariate Poisson regression. Poisson distribution based models were criticized in [10] from a football betting market perspective. More general models which use general bivariate discrete distributions generated by copulas have been developed in [23]. In [13], a comparison of goal-driven and ordered regression models can be found.

Machine learning techniques have also been proposed for the prediction of the outcomes of soccer events. In [34], a genetic programming based technique is applied and compared to other two methods based on fuzzy models and neural networks. Applicability of fuzzy rules is also investigated in [28] where the rules are generated by a combination of genetic and neural optimization techniques. More recently, a novel technique, the Bayesian network, which is a graphical probabilistic model, was introduced into soccer science in [8]. A Bayesian network represents the conditional dependencies among uncertain variables, which can be both objective and subjective.

In our competitive programming setup these statistical and machine learning models are competing with each other and they are compared by assessing the quality of their forecasting performance. In the literature, there are various ways for doing assessment, for example, different types of indicators can be considered such as accuracy and profitability. One of the most popular scores is the Rank Probability Score (RPS), see [11] and [8]. In the FootballAvatar system, several objective and subjective goodness-of-fit indicators can be used for assessing the accuracy of the forecasts derived by models mentioned above. In particular, it is also possible to compare our prediction with bookmakers' ones.

2.2.2. Core Simulation Computations

From the viewpoint of implementations, we distinguish the following two main types of simulators: the MABSA ones and the FANM ones. While MABSA (Multi-Agent-Based Server Architecture) is used for research purposes only, commercial software components of the FootballAvatar project are based on our FANM (FANM is Not MABSA) platform. FANM is the antithesis of MABSA. For example, the heart of the MABSA platform is an asynchronous I/O multiplexed TCP/IP proxy server written in C++11. In the MABSA platform teams, players, coaches, 2D, 3D and mobile display programs and the simulation algorithms themselves are implemented as clients that communicate with the server via TCP/IP using Google's Protobuf [15]. Conversely, FANM programs are standalone monolithic applications that do not use networking at all.

On higher levels, MABSA and FANM simulations are typically quasi-realistic ones. In contrast to realistic simulations (such as 2D robot soccer [19], simplified 2D robot soccer [3], or Simple Soccer [6, p. 133–193]), our quasi-realistic computations are organized around a few key features of the soccer game, such as passing graphs or lineups. The base algorithms of MABSA and FANM as well as their software infrastructure are presented in Section 4. More advanced simulation models such as the ones that use cellular networks or Bayesian networks will be discussed in a further paper.

2.3. Competitive Programming

During the development of the FootballAvatar project we have developed a new software process methodology that we call Competitive Programming (or CP for short) [4]. This methodology is based on a combination of eXtreme Programming (XP) ([5], [36]) and Rapid Application Development (RAD) [22]. Simulation programs presented in this paper were developed according to CP.

3. Definitions of Avatars

As the main result of this paper, the mathematical and the information technological definitions of avatars are presented in this section.

3.1. Mathematical Definition

The spirit of the following statistics-based definition may be deduced from the hypothesis testing of [2]. In addition, a heuristic version of this definition can be found in the paper [4].

First, let us select n number of properties that will be observed in simulations. These quantities are arranged in an n-dimensional random variable $\mathbf{X} = (X_1, \ldots, X_n)$. The realizations of this vector $\mathbf{X}$ will be called a-priori observations with a-priori distribution function F. For example, X_1 may be the number of goals scored by a given team in a given soccer match. This, of course, depends on chance but it has well-known realizations from the past. Simulations will give further realizations of $\mathbf{X}$ and all we have to do is to compare these realizations to a-priori observations. This approach is supported by the following definition of avatars.

Definition 3.1 (Avatars)**.** Let $\mathbf{X} = (X_1, \ldots, X_n)$ be the selected properties and let $\mathcal{S}\colon U \to \mathbb{R}^n$ symbolize a simulation algorithm, where U denotes an arbitrary (possibly empty) set of inputs of this algorithm. The pair $(\mathbf{X}, S)$ is referred to as an avatar (with significance level α) if the null hypothesis $H_0 : F = F_S$ is not being rejected, where F_S denotes the distribution function of the random outputs of the simulation algorithm.

It is possible to weaken this definition in the form that the hypotheses $H_0 : F_i = F_{S,i}$ are investigated for all $i = 1, \ldots, n$, where F_i and $F_{S,i}$ denote the distribution functions of the ith avatar property and the ith output coordinate of the simulation algorithm, respectively. The difference between the two definitions is that in the first case the joint distributions are compared, while in the second case only the marginal ones.

If the simulation has no input at all or its input is not decisive then it is referred to as standalone, otherwise it is called avatar-based. Here we present some trivial thought experiments with the definition, and a few non-trivial ones can be found in Section 6.2.

Example 3.2 (A trivial ping-pong avatar). Suppose that we observed two table tennis players during their last 10 matches. In each game, the relative frequencies of their successful serves and returns were determined. Let $y^{1,1}$, $y^{1,2}$, $y^{2,1}$, and $y^{2,2}$ denote the vectors of these relative frequencies. To be more precise, the sequence $y^{p,j}\colon \{\,1,\dots,10\,\} \to [0,1]$ tells us the relative frequency of successful serves ($j=1$) and returns ($j=2$) of player p ($p=1,2$) in the ith game ($i=1,\dots,10$).

Let assume that we have the following observations:

$$\begin{aligned}
y^{11} &= (21/41, 10/36, 23/40, 15/28, 19/32, 10/33, 24/42, 12/41, 12/30, 29/33),\\
y^{12} &= (11/25, 16/21, 6/21, 15/16, 12/29, 14/23, 10/19, 7/15, 6/14, 6/16),\\
y^{21} &= (11/41, 17/30, 14/23, 11/31, 14/22, 8/21, 14/27, 8/17, 11/14, 12/20),\\
y^{22} &= (19/39, 17/28, 12/33, 8/24, 14/24, 10/18, 8/27, 20/30, 11/29, 11/26).
\end{aligned}$$

Now, for example, let us consider the last 5 matches only. Accordingly, let

$$\begin{aligned}
p_{11} &= \sum_{i=6}^{10} y_i^{11}/5 = 0.4891859, \quad & p_{12} &= \sum_{i=6}^{10} y_i^{12}/5 = 0.4810499,\\
p_{21} &= \sum_{i=6}^{10} y_i^{21}/5 = 0.5511547, \quad & p_{22} &= \sum_{i=6}^{10} y_i^{22}/5 = 0.4641812.
\end{aligned}$$

Then the a-priori probabilities can be estimated as

$$\mathbf{X} = \begin{pmatrix} \text{player 1's serves} \\ \text{player 1's returns} \\ \text{player 2's serves} \\ \text{player 2's returns} \end{pmatrix} = \begin{pmatrix} \frac{p_{11}}{p_{11}+p_{22}} \\ \frac{p_{12}}{p_{12}+p_{21}} \\ \frac{p_{21}}{p_{21}+p_{12}} \\ \frac{p_{22}}{p_{22}+p_{11}} \end{pmatrix} = \begin{pmatrix} 0.5131139 \\ 0.4660412 \\ 0.5339588 \\ 0.4868861 \end{pmatrix}.$$

The components of this vector denote the estimated probabilities of successful serves and returns. To be more precise, they indicate whether the serving or the returning player gets the point. We define the avatar data transformation function as

$$\mathcal{A}(y^{11}, y^{12}, y^{21}, y^{22}) = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

$\mathcal{A}$ maps the observations to the input parameters of the algorithm S. The IT implementation of avatar transformations is discussed in Section 5.

For simulations we have used the algorithm S shown in Fig. 2(a). Our practical experience shows that $(\mathbf{X}, S \circ \mathcal{A})$ is an avatar.

Example 3.3 (Tennis avatar). In this example we consider data about the two finalists of the Australian Open 2014 tennis championship, Rafael Nadal and Stanislas Wawrinka available from [33]. In each game of the tournament the relative frequencies of the points won after the 1st serve and the points won after the opponent's 1st serve were collected. Let $y^{1,1}$, $y^{1,2}$, $y^{2,1}$, $y^{2,2}$ denote the vectors of these measurements. To be more precise, the sequence $y^{p,j}\colon \{\,1,\dots,7\,\} \to [0,1]$

tells us the relative frequencies of the points won after the 1st serve ($j = 1$) and the points won after the opponent's first serve ($j = 2$) of the pth player ($p = 1, 2$) in the ith game ($i = 1, \ldots, 7$). The first player is Stanislas Wawrinka, the second one is Rafael Nadal.

In our case we have the following:

$$\begin{aligned}
y^{11} &= (20/22, 57/69, N/A, 54/60, 71/98, 71/87, 50/84),\\
y^{12} &= (18/36, 25/81, N/A, 23/76, 30/105, 16/88, 7/53),\\
y^{21} &= (14/16, 45/55, 38/53, 67/88, 66/91, 41/56, 46/53),\\
y^{22} &= (5/20, 19/52, 29/68, 25/76, 26/90, 24/69, 34/84).
\end{aligned}$$

Note that the 3rd match of Wawrinka was canceled. In the following we consider only the values of the last 4 matches. Accordingly, let

$$\begin{aligned}
p_{11} &= \sum_{i=4}^{7} y_i^{11}/4 = 0.85152475, \quad p_{12} = \sum_{i=4}^{7} y_i^{12}/4 = 0.22555988,\\
p_{21} &= \sum_{i=4}^{7} y_i^{21}/4 = 0.76058081, \quad p_{22} = \sum_{i=4}^{7} y_i^{22}/4 = 0.34260606.
\end{aligned}$$

Then the a-priori probabilities can be estimated as

$$\mathbf{X} = \begin{pmatrix} \text{Wawrinka's serves} \\ \text{Wawrinka's returns} \\ \text{Nadal's serves} \\ \text{Nadal's returns} \end{pmatrix} = \begin{pmatrix} \frac{p_{11}}{p_{11}+p_{22}} \\ \frac{p_{12}}{p_{12}+p_{21}} \\ \frac{p_{21}}{p_{21}+p_{12}} \\ \frac{p_{22}}{p_{22}+p_{11}} \end{pmatrix} = \begin{pmatrix} 0.71309168 \\ 0.22872992 \\ 0.77127008 \\ 0.28690832 \end{pmatrix}.$$

The components of this vector denote the estimated probabilities of the points won after the 1st serve and the points won after the opponent's first serve. To be more precise, they indicate whether the serving or the returning player gets the point. We define the avatar data transformation function as

$$\mathcal{A}(y^{11}, y^{12}, y^{21}, y^{22}) = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

For simulations we have used the algorithm S shown in Fig. 2(b) that is analogous with the one used for the ping-pong avatar. Our practical experience shows that $(\mathbf{X}, S \circ \mathcal{A})$ is an avatar.

Example 3.4 (Trivial dribbling-tackling or one-dimensional football)**.** Let us consider two soccer players A and B as it is shown in Fig. 1. During their last 5 matches, these two players were observed in order to determine the relative frequencies of their successful dribbles and tackles. Let y^{AT}, y^{AD}, y^{BT}, and y^{BD} denote the vectors of these relative frequencies. With the notation of Definition 3.1, we have the following:

$$y^{AT} = (3/9, 2/8, 3/10, 4/10, 2/8), \quad p_{AT} = \sum y_i^{AT}/5 = .3067,$$

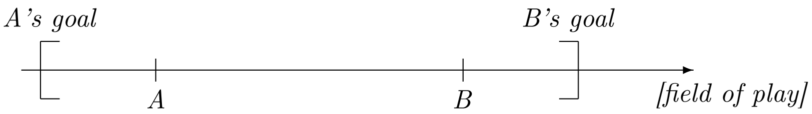


Figure 1: The informal interpretation of the one-dimensional football, where A and B are the two investigated players. For example, a situation $A_t > B_t$ may denote that the attacker A dribbles the defender B (and accordingly, B cannot tackle A), or vice versa, at time t.

$$\begin{aligned}
y^{AD} &= (6/8, 9/10, 6/6, 4/8, 6/7), \quad & p_{AD} &= \sum y_i^{AD}/5 = .8014,\\
y^{BT} &= (8/10, 6/7, 6/6, 8/9, 5/6), \quad & p_{BT} &= \sum y_i^{BT}/5 = .8759,\\
y^{BD} &= (3/10, 2/8, 4/9, 3/5, 1/4), \quad & p_{BD} &= \sum y_i^{BD}/5 = .3689.
\end{aligned}$$

Let's estimate the a-priori probabilities for the experiment by the following:

$$\mathbf{x} = \begin{pmatrix} 1 - \frac{p_{BD}}{p_{BD}+p_{AT}} \\ \frac{p_{AD}}{p_{AD}+p_{BT}} \\ 1 - \frac{p_{AD}}{p_{AD}+p_{BT}} \\ \frac{p_{BD}}{p_{BD}+p_{AT}} \end{pmatrix} = \begin{pmatrix} 0.4539668 \\ 0.4777917 \\ 0.5222083 \\ 0.5460332 \end{pmatrix}.$$

We define the avatar data transformation function as

$$\mathcal{A}(y_t^{AT}, y_t^{AD}, y_t^{BT}, y_t^{BD}) = \begin{pmatrix} p_{AT} & p_{BT} \\ p_{AD} & p_{BD} \end{pmatrix}.$$

The simulation S is implemented by the R code shown in Fig. 2(b).

(a) The R code of the ping-pong avatar example.

```
serve <- function(p) {
  u <- runif(1)
  if(u < p)
    return(1)
  else
    return(0)
}

p_1s <- 0.5131
p_1r <- 0.4660
p_2s <- 0.5339
p_2r <- 0.4868

nofmatches <- 10000
c_1s <- 0
c_1r <- 0
c_2s <- 0
c_2r <- 0

for(i in 1:nofmatches) {
  for(j in 1:10) {
    if(serve(p_1s)) {
      c_1s <- c_1s + 1
    } else {
      c_2r <- c_2r + 1
    }
    if(serve(p_2s)) {
      c_2s <- c_2s + 1
    } else {
      c_1r <- c_1r + 1
    }
  }
}

cat(c_1s / (nofmatches * 10),
  c_1r / (nofmatches * 10),
  c_2s / (nofmatches * 10),
  c_2r / (nofmatches * 10),
  "\n"
)
```

(b) The R code of the one-dimensional football.

```
attack <- function(p_d, p_t) {
  p <- p_d * (1 / (p_d + p_t))
  u <- runif(1)
  if(u < p)
    return(1)
  else
    return(0)
}

p_AT <- 0.3067
p_AD <- 0.8014
p_BT <- 0.8759
p_BD <- 0.3689

nofmatches <- 10000
AT <- 0
AD <- 0
BT <- 0
BD <- 0

for(i in 1:nofmatches) {
  for(j in 1:10) {
    if(attack(p_AD, p_BT)) {
      AD <- AD + 1
    } else {
      BT <- BT + 1
    }
    if(attack (p_BD, p_AT)) {
      BD <- BD + 1
    } else {
      AT <- AT + 1
    }
  }
}

cat(AT / (nofmatches * 10),
  AD / (nofmatches * 10),
  BT / (nofmatches * 10),
  BD / (nofmatches * 10),
  "\n"
)
```

Figure 2: The two trivial simulation algorithms of the introductory examples.

The results from running the simulation algorithms in Fig. 2 shows that $(\mathbf{x}, S \circ \mathcal{A})$ is an avatar.

Definition 3.5 (Football avatar)**.** Let $\mathbf{X} = (X_1, \ldots, X_n)$ be the investigated avatar properties. An avatar $(\mathbf{X}, S \circ \mathcal{A})$ is referred to as a passing distribution or lineup-based football avatar if $\mathcal{A} : U \to \mathbb{R}^{(11+k)\times 11}$ and $S : \mathbb{R}^{(11+k)\times 11} \to \mathbb{R}^n$, where k denotes the number of control avatar parameters.

The number 11 corresponds to the starting 11, however, the whole team must be used in practical applications.

Definition 3.6 (Deep avatar)**.** A football avatar $(\mathbf{X}, S \circ \mathcal{A})$ is referred to as a deep avatar with depth v if $\mathcal{A} = \mathcal{A}_v \circ \mathcal{A}_{v-1} \circ \cdots \circ \mathcal{A}_0$, where $\mathcal{A}_0 \colon U \to \mathbb{R}^{(11+k)\times 11}$ and $\mathcal{A}_i \colon \mathbb{R}^{(11+k)\times 11} \to \mathbb{R}^{(11+k)\times 11}$, for $i = 1, \ldots, v$.

Remark 3.7 (Simulational, molecular, physiological, mechanical, and psychological avatar data transformations). Let $(\mathbf{X}, S \circ \mathcal{A})$ be a deep avatar with $\mathcal{A} = \mathcal{A}_4 \circ \mathcal{A}_3 \circ \mathcal{A}_2 \circ \mathcal{A}_1 \circ \mathcal{A}_0$, where $\mathcal{A}_0$ is called a simulational (or tactical), $\mathcal{A}_1$ is called a molecular, $\mathcal{A}_2$ is called a physiological, $\mathcal{A}_3$ is called a mechanical, and finally, $\mathcal{A}_4$ is called a psychological avatar data transformation function. It is interesting to notice that the simulational avatar transformation $\mathcal{A}_0$ is a strongly distinguished function in the definition of deep avatars.

3.2. Information Technological Definition

In the sense of information technology, football avatars are simply cross cutting concepts between tactical avatars and higher order avatars. For example, it is obvious that players' passing probabilities or shooting accuracies are influenced by their physiological and psychological states. The basic data of tactical avatars are based on the passing matrices, lineups, and some other avatar control properties. These empirical and/or predicted quantities are used by our simulation software by default. Therefore, it is natural to use aspects (in the sense of Aspect Oriented Programming) to implement the influences of several internal and external factors to the tactical avatars. For example, external factors to be taken into consideration include environmental aspects (such as weather conditions and properties of the pitch) or the referee.

Example 3.8 (A top goalscorer aspect)**.** Suppose that we have tactical avatars for both the next round's home and away teams, but the home team has already won the national championship, so it can play without stakes. However, let us assume that it is possible for a striker of the home side to win the top goalscorer title. Accordingly, the tactic of the home team has been changed to achieve this goal. It is an interesting question how this change impacts the tactical avatars for the simulation of the next match. For example, it is necessary to modify the passing distribution matrix and the selfishness control property. It can be done quickly and easily by writing an appropriate aspect.

4. Soccer Simulation Algorithms

4.1. Statistical Simulation Computations

This kind of simulation computation is typically used on level "-1". In this paper, the bivariate Poisson regression model is described in more detail. In this model, the result of a match, which is a pair of the number of goals scored by the home and the away teams (denoted by X and Y, respectively), has a bivariate Poisson distribution. This distribution is a linear transformation of three independent Poisson distributions. Namely, (X, Y) follows the bivariate Poisson distribution with parameters $\lambda_1, \lambda_2, \lambda_3$, if $X = Z_1 + Z_3$ and $Y = Z_2 + Z_3$, where Z_1, Z_2, Z_3 follow independent Poisson distributions with parameters $\lambda_1, \lambda_2, \lambda_3$, respectively. Clearly, X and Y are correlated with the correlation coefficient λ_3. In our case, X and Y denote the number of goals scored by the home and away teams, respectively, while Z_3 denotes the number of goals scored by both teams, Z_1 and Z_2 are the goal differences by which the match was won and lost, respectively, from the viewpoint of the home team. In the bivariate Poisson regression model, the intensity parameters $\lambda_1, \lambda_2, \lambda_3$ depend on various covariates, e.g., the team indicators or the home playing indicator. The parameters incorporated by these covariates into the model are able to measure the impact of the home pitch or the team strength at home or away.

Algorithm 1 The EM algorithm for computing team scores.

Require: *match_results* : array of (*home*, *away*) team results and *design_matrices* : X_w, X_l, X_d; ▷ The win, lose, and draw design matrices are defined by team-indicators and additional covariates.
Ensure: *parameters* : $\beta_w, \beta_l, \beta_d$; ▷ The parameters according to winning, losing, and drawing of the bivariate Poisson distribution.
1: **procedure** TEAMSCORE(*match_results*, *design_matrices*, *intensities*)
▷ **Initialization**
2: $g_d \leftarrow \min\{home, away\}$; ▷ Goals by both teams.
3: $g_w \leftarrow home - g_d$; ▷ Winning goals.
4: $g_l \leftarrow away - g_d$; ▷ Losing goals.
5: $\beta_w \leftarrow \mathrm{glm}(X_w, g_w, \texttt{"poisson"})$; ▷ Fitting Poisson regression for winning goals by using the glm procedure.
6: $\beta_l \leftarrow \mathrm{glm}(X_l, g_l, \texttt{"poisson"})$; ▷ Same for losing goals.
7: $\beta_d \leftarrow \mathrm{glm}(X_d, g_d, \texttt{"poisson"})$; ▷ Same for draw.
8: $\lambda_w \leftarrow \exp\{X_w * \beta_w\}$; ▷ Computing initial winning intensity.
9: $\lambda_l \leftarrow \exp\{X_l * \beta_l\}$; ▷ Computing initial losing intensity.
10: $\lambda_d \leftarrow \exp\{X_d * \beta_d\}$; ▷ Computing initial draw intensity.
11: **repeat**
▷ **Expectation step**
12: $\varrho \leftarrow \lambda_d / (\lambda_w * \lambda_l)$;
13: **if** $(home > 0) \& (away > 0)$ **then**
14: $ratio \leftarrow G[home, away]/G[home + 1, away + 1]$; ▷ G is computed by **CondExp**
15: $g_d \leftarrow home * away * \varrho * ratio$;
16: **else**
17: $g_d \leftarrow 0$;
18: **end if**
19: Repeat steps 3 and 4 with new g_d;
▷ **Maximization step**
20: Repeat steps 5, 6, and 7 with new g_w, g_l, g_d;
21: Repeat steps 8, 9, and 10 with new $\beta_w, \beta_l, \beta_d$;
22: **until** convergence;
23: **return** $\lambda_w, \lambda_l, \lambda_d$;
24: **end procedure**

The parameters of bivariate Poisson regression are estimated by an EM-type algorithm, see Algorithm 1. This algorithm consists of two steps: the expectation (E) and the maximization (M) steps. In the E-step, the independent components Z_1, Z_2, Z_3 are estimated from marginals X and Y using conditional expectation, see Algorithm 2. In the M-step, three separate Poisson regressions are fitted using the generalized linear model (GLM) framework. For example, this fitting can be done by using the `family = poisson` option in the `glm` function of R. The output of Algorithm 1 can be, for example, the home, away, and draw strength of the teams in a league which can be shown as in Fig. 4 in the case of the Hungarian

National Championship during season 2011/2012.

Algorithm 2 Computing the G matrix in Algorithm 1.

Require: ϱ : *real* correlation, *max_goal* : *integer* maximum number of goals;
Ensure: $G : (max\_goal + 1) \times (max\_goal + 1)$ array;

```
procedure CondExp(ϱ,max_goal,G)
    G ← 1;                                            ▷ Initialization
    for j = 1 to max_goal do
        for k = 1 to j do
            G[j + 1, k + 1] ← G[j + 1, k] + (j + 1) ∗ ϱ ∗ G[j, k];
            G[k + 1, j + 1] ← G[j + 1, k + 1];
        end for
    end for
    return G;
end procedure
```

4.2. Core and Massively Parallel Simulation Computations

Our core soccer simulation algorithms are implemented as quasi-realistic MABSA or FANM simulation computations. The two most successful MABSA soccer algorithms were *FBA One* and *TF*. Their pseudo-codes are shown in Algorithm 3 and Algorithm 4, respectively.

Algorithm 3 The default simulation algorithm used by the team FBA One.

Require: ActorRequest *actors*[]; ▷ The class ActorRequest is a Google Protocol Buffers object.
Ensure: SensoryInput *response*; ▷ The class SensoryInput is a Google Protocol Buffers object.

```
procedure A simulation step(actors[])                 ▷ Called at discrete time step t
    t ∈ ℕ is the discrete time index, Players players[], Ball ball;   ▷ Server side global objects.
    Player p = possession of the ball();              ▷ Who has the possession of the ball?
    if SHOOT or PASS ∈ actors[p] then
        start ball moving simulation();   ▷ Runge-Kutta to predict the future motion of the ball.
    end if
    moving(ball);
    for each q ∈ players do
        moving(q);
    end for
    Searching for actions or play modes();
    response = create response(ball, actors, players);
    return response;                      ▷ This will be the sensory input of the players.
end procedure
```

FANM algorithms were developed from the successful MABSA algorithms. The architecture of FANM teams are suitable for porting these to GPU in order to implement massively parallel simulation computations. All CUDA ported versions are shown in Fig. 3. Two of these ports will be presented in detail in Sect. 6.3.

5. Avatar Transformations

5.1. Soccer Analytics

In order to use a soccer simulator in practice it should be driven by real soccer data. Our simulation programs developed in the FootballAvatar project also provide such

Algorithm 4 The simulation algorithm used by the team TF.

```
Require: ActorRequest actorRequest[];
Ensure: SensoryInput sensoryInput;    ▷ The classes ActorRequest and SensoryInput are Google
    Protocol Buffers objects.
procedure SIMULATE(SensoryInput)
    for i = 0 to all_players_size do
        player ← mutable_all_players(i)
        switch state_of_play
            case BALL_IN_PLAY
                switch player.action_type
                    case MOVE
                        move_ball_in_play(player)
                    case KICK
                        kick_ball_in_play(player)
                    ...
            ...
    end for

    runge_kutta()                     ▷ Runge-Kutta to predict the motion of the ball.
    time_stamp ← time_stamp + 1

    sleep()
    write_server(SensoryInput)
end procedure
```

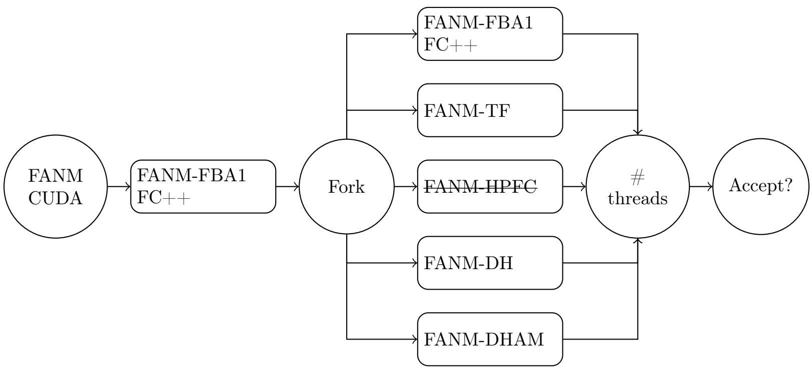


Figure 3: An arrangement of the CUDA ports of prototypes in our competition programming flow chart. The object of the contest is to maximize the number of parallel threads of soccer matches in a CUDA block.

functionality. We have developed file formats for representing soccer data that originate from wearable sensors and from video processing. Our partner companies provide and operate an infrastructure that can supply us with soccer data in our file formats. In order to be useful for our simulators real soccer data is subject to processing, that is referred to as Basic Avatar Transformation (or BAT for short).

The output of BAT serves as input to our simulators. Currently, BAT computes a passing matrix for both teams and another matrix called "heat map" for each player. Each element of a "heat map" matrix corresponds to a specific rectangular region of the pitch (the pitch is divided into disjoint regions by an equidistant grid).

A number in the "heat map" matrix represents the time spent by the player in the corresponding region.

The FANM-Debrecen Handsomes (or FANM-DH for short) simulation algorithm (mentioned later in Sect. 6.2) uses AspectJ crosscutting to manipulate certain parameters. One of these parameters is selfishness. Assume that one of our players, for some reason, shoots on goal with greater probability than he passes in some situations. The crosscutting made for the FANM-DH algoritm can manipulate this probability. By these AspectJ crosscuts we are able to manipulate more player parameters, constructing every players properties.

5.2. Biological and Behavioural Models

The Player Stamina Avatar Property (PSAP) is introduced at higher levels. Each player has a personalized fitness level based on real-life measured data. Basic physical parameters (e.g., weight, height), sport performance indices (e.g., 40 yard dash, fitness tests) and basic physiological/haematological levels (e.g., blood pressure, blood sugar) determine a measure in the range between 1 and 1000 PSL (the abbreviation PSL stands for Player Stamina Level). In simulations the value of this measure is continuously decreased based on the players' behaviour (e.g, running at a higher speed results in more decrease). Because it is practically not possible to obtain all these data before every match, this avatar property have to be estimated based on the matches since the last measurement. PSAP works like the 2DRCSS's stamina model [27], but is more realistic, because it is based on real data. In our plans, a more advanced PSAP will be able to indicate suspicious events like doping, not from measured data like the Union Cycliste Internationale's (UCI) Athlete Biological Passport [35], but from the difference between the reality and the simulations.

The Fouls Avatar Property (FAP) is also introduced at higher levels. Each player varies in how aggressive he/she is in the pitch. However, the level of aggression can be characterized by the number of yellow and red cards awarded, or the number of fouls committed during previous matches. FAP incorporates these factors, as well as some others. For example, aggression is influenced by mood: after making a successful tackle a player is likely to become more aggressive, that implies higher chance of making a foul. Furthermore, FAP also interacts with PSAP, e.g., the more tired a player is, the higher is the probability of committing a mistake.

6. Simulation Computation Results

6.1. Statistical Simulation Computations

The simulation computations used or investigated on level "-1" may typically give several alternative league tables. Now, we compare two families of such league tables, namely the alternative and the PCA-Poisson league tables that are shown in Table 1 together with the real league table for the Hungarian National Championship. The order in the alternative league table is based on Google's Page Rank algorithm [25]. In this case, there is an edge from team A to team B in the Google matrix if team B can score at least one point against team A. An example of an alternative league table can be seen in the middle part of Table 1. Uses of Page Rank in sport performance analysis can be found in the scientific literature, e.g., see [17] and [16]. One of the authors continuously maintain some Hungarian language web pages devoted to alternative league tables at `http://hu.wikipedia.org/wiki/Alternat%C3%ADv_tabella` and `http://nsza.blog.hu/`. This activity takes place independently from the FootballAvatar project and these pages contain alternative league tables for the Premier League, Serie A, Primera División, and the Bundesliga, too. On level "-1" of the FootballAvatar project Page Rank algorithm-based experiments are performed by two competing development teams (see Table 1).

Table 1: The first two columns show the official league table for the Hungarian National Championship for the 2011/2012 season. The third and fourth columns contain an alternative league table for the same season. The fifth and sixth columns contain a PCA-Poisson based league table for the same season as well.

Official		Alternative		PCA-Poisson	
Team	Points	Team	Rank	Team	Score
Debrecen	74	Debrecen	0.0992	Debrecen	1.2021
Videoton	66	Kaposvár	0.0755	Videoton	1.1666
Győr	63	Siófok	0.0719	Győr	0.9545
Honvéd	46	Honvéd	0.0702	Kecskemét	0.3926
Kecskemét	45	Videoton	0.0696	Haladás	0.2676
Paks	45	Győr	0.0682	Diósgyőr	0.0232
Diósgyőr	43	Paks	0.0679	Paks	0.0045
Haladás	38	Haladás	0.0646	Újpest	-0.1163
Siófok	36	Pécs	0.0631	Pécs	-0.1951
Kaposvár	35	Kecskemét	0.0615	Honvéd	-0.2376
Ferencváros	34	Vasas	0.0535	Kaposvár	-0.2856
Pécs	34	Pápa	0.0517	Siófok	-0.3628
Újpest	32	Diósgyőr	0.0499	Ferencváros	-0.4397
Pápa	30	Ferencváros	0.0498	Pápa	-0.4843
Vasas	22	Újpest	0.0473	Vasas	-0.6430
Zalaegerszeg	13	Zalaegerszeg	0.0353	Zalaegerszeg	-1.2467

The other league table called PCA-Poisson is based on the EM algorithm shown in Alg. 1 and Principal Component Analysis (PCA) applied for the three dimensional vector with coordinates home, away, and draw strength. The one-dimensional score given by the dimension reduction technique PCA is shown in the last column of Table 1.

6.2. Core Simulation Computations

Core simulation computations were introduced in [4]. We have collected the total number of points and goals during seasons 2004/2005–2012/2013 in the Hungarian National Championship, that are shown in Table 2. Note that $\overline{x}_g$ and $\overline{x}_p$ denote the mean of goals and points respectively, and s^*_{ng} and s^*_{np} denote the standard

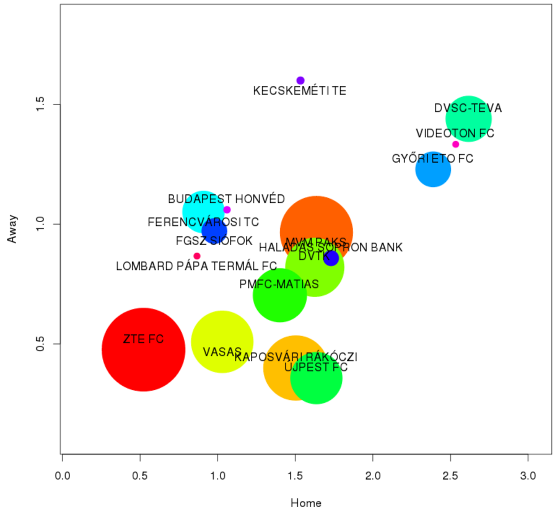


Figure 4: The scatterplot of home and away team strengths with draw strength in bubble size.

deviation of goals and points respectively.

Season	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13
Goals	681	707	677	746	710	707	690	648	639
Points	644	658	671	666	667	659	667	656	660

Table 2: The total number of goals and points scored during seasons 2004/2005–2012/2013 in the Hungarian National Championship ($\overline{x}_g = 689.44$, $s^*_{n\,g} = 33.02$, $\overline{x}_p = 660.89$, and $s^*_{n\,p} = 8.10$).

We have simulated nine seasons for this championship using the FANM-Debrecen HardAsMuscle and the FANM-Debrecen Handsomes algorithms, Tables 3 and 4 shows the total number of goals and points obtained from these simulations. These results were compared with the similar statistical results in [4]. The significance level α for all tests is chosen as $\alpha = 0.05$. Our findings show that both simulation algorithm passes the statistical tests applied.

6.3. Massively Parallel Simulation Computations

In this point, two CUDA ported FANM teams are presented briefly. All FANM and CUDA ported FANM teams have the same functionality because they were inherited from a common ancestor.

Season	1.	2.	3.	4.	5.	6.	7.	8.	9.	Test stat.
Goals	611	626	649	639	596	606	644	610	622	6/78
Points	663	670	672	658	665	667	671	685	669	10/21

Table 3: The total number of goals and points obtained from the simulation algorithm FANM-Debrecen HardAsMuscle ($\overline{x}_g = 622.56$, $s^*_{n\,g} = 18.41$, $\overline{x}_p = 668.89$, and $s^*_{n\,p} = 7.47$). The last column shows the values of test statistics in form of Wald-Wolfowitz/Mann-Whitney.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	Test stat.
Goals	704	674	703	701	714	665	694	679	682	11/42
Points	667	658	669	665	665	657	666	661	656	9/41

Table 4: The total number of goals and points obtained from the simulation algorithm FANM-Debrecen Handsomes ($\overline{x}_g = 690.67$, $s^*_{n\,g} = 15.42$, $\overline{x}_p = 662.67$, and $s^*_{n\,p} = 4.5$). The last column shows the values of test statistics in form of Wald-Wolfowitz/Mann-Whitney.

6.3.1. The FANM-TF CUDA Implementation

This algorithm simulates several thousand matches between two teams. By default, each team has 4 formations (as shown in Table 5) with a corresponding passing distribution and avatar property table. This gives us 16 matchups, each of which has its own CUDA block. The matches are played on separate threads but with common constant data. The algorithm has two kernel launches for each of the two halves with an in-between initialization of the second half.

Lineups	4-4-2	4-2-2-2	4-2-3-1	4-3-3
4-4-2	240/27/133	235/36/129	357/13/30	389/4/7
4-2-2-2	163/23/214	189/35/176	314/16/70	369/6/25
4-2-3-1	78/16/306	102/23/275	219/15/166	319/15/66
4-3-3	15/13/372	31/9/360	58/10/332	195/16/189

Table 5: FANM-TF's Hungarian flag notation.

6.3.2. The FANM-Debrecen Handsomes CUDA Implementation

The results are less predictable compared to FANM-TF CUDA implementation, given that the differences between the players (avatars) of the teams reflected by

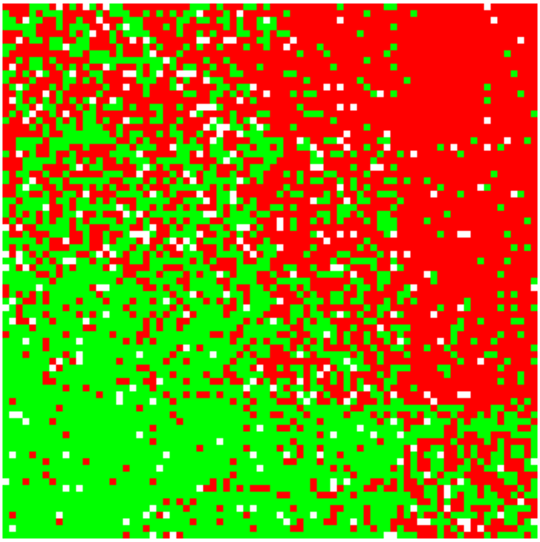


Figure 5: FANM-TF's Hungarian flag notation.

Lineups	4-4-2	4-2-2-2	4-2-3-1	4-3-3
4-4-2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 1 1 398	0 0 0 0 400	0 0 0 1 399	0 0 0 1 399
4-2-2-2	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 400	0 0 1 1 398	0 0 0 0 400	0 0 0 0 400
4-2-3-1	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 400	0 0 0 0 400	0 0 0 0 400	0 0 0 0 400
4-3-3	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 0	0 0 0 0 1	0 0 0 0 0	0 0 0 0 0
	0 0 0 0 400	0 0 0 0 399	0 0 0 0 400	0 0 0 0 400

Table 6: FANM-TF's crystal ball notation.

this algorithms are less significant and the players of a team have quite similar characteristics (see Fig. 6 and Table 8).

6.4. Avatar Simulation Computations

We have simulated nine seasons for the same championship as in Sect. 6.2 using the FANM-TF algorithm. The total number of goals and points obtained from the simulations are shown in Table 9. These results was compared with the similar statistical results in [4]. Our findings show that this simulation algorithm passes

Lineups	4-4-2	4-2-2-2	4-2-3-1	4-3-3
4-4-2	127/148/125	162/88/150	88/119/93	143/113/144
4-2-2-2	158/83/159	140/119/141	186/115/99	136/121/143
4-2-3-1	96/130/174	101/131/168	8/200/102	79/141/180
4-3-3	149/116/135	133/123/144	147/147/106	136/130/134

Table 7: FANM-Debrecen Handsomes' Hungarian flag notation.

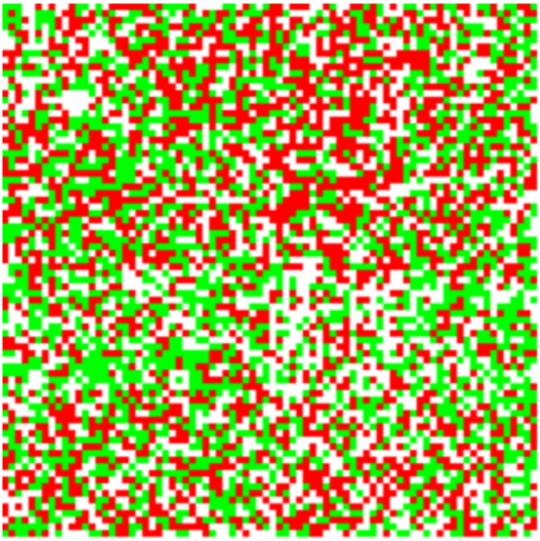


Figure 6: FANM-Debrecen Handsomes' Hungarian flag notation.

Lineups	4-4-2	4-2-2-2	4-2-3-1	4-3-3
4-4-2	82 29 18 13 7 16 20 18 9 10 14 35 34 10 7 7 20 6 14 3 6 14 4 6 4	29 24 26 10 4 31 34 33 20 12 25 32 21 11 8 6 18 17 4 2 7 13 9 4 0	60 32 10 5 0 68 44 26 6 6 28 29 14 5 3 24 13 9 1 0 6 4 7 0 0	26 43 20 15 6 38 70 18 17 8 13 30 11 13 4 10 24 11 5 0 2 9 3 2 2
4-2-2-2	31 21 27 15 9 23 24 45 9 11 24 31 19 14 7 13 17 14 6 1 8 11 12 5 3	41 21 20 10 6 29 47 37 17 7 27 35 22 9 10 12 13 5 9 4 5 7 5 2 0	59 34 19 4 0 68 45 23 7 1 35 33 7 8 3 14 20 4 3 0 3 6 1 2 1	57 40 20 10 8 38 43 27 14 4 23 29 16 18 2 6 12 10 4 0 4 9 4 1 1
4-2-3-1	74 55 40 11 2 33 39 31 16 6 28 18 16 10 3 0 0 0 0 0 0 0 0 0 400	68 58 40 11 6 49 42 26 9 7 15 16 17 7 2 4 6 5 4 2 0 2 3 0 0	158 64 18 4 1 62 37 12 2 1 19 4 5 0 0 2 6 1 0 0 1 1 1 1 0	88 74 34 11 4 42 46 29 15 3 12 16 6 5 4 1 3 3 1 1 1 1 0 0 0
4-3-3	46 36 17 9 5 40 43 30 14 8 23 29 19 11 2 10 15 5 7 3 4 13 5 4 2	59 43 35 8 8 50 41 23 12 2 15 23 5 3 0 13 12 5 3 0 5 6 3 1 0	100 55 18 7 1 63 41 10 9 0 26 21 5 4 1 12 7 6 1 1 2 4 4 2 0	61 51 20 5 5 38 51 28 9 6 18 37 10 4 2 7 9 6 7 3 1 8 1 2 2

Table 8: FANM-Debrecen Handsomes' crystal ball notation.

the statistical tests applied.

Season	1.	2.	3.	4.	5.	6.	7.	8.	9.	Test stat.
Goals	693	628	661	639	665	682	708	630	701	10/56
Points	668	655	659	659	665	658	666	660	669	11/38

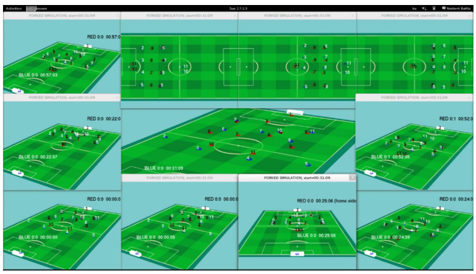
Table 9: The total number of goals and points obtained from the simulation algorithm FANM-TF ($\overline{x}_g = 667.44$, $s^*_{n\,g} = 30.51$, $\overline{x}_p = 662.11$, and $s^*_{n\,p} = 4.96$). The last column shows the values of test statistics in form of Wald-Wolfowitz/Mann-Whitney.

It should be noted that all these simulations have been computed on the level "-1" in avatar mode.

7. Conclusion and Future Work

In this work, we introduced a general framework for sport science simulations that we refer to as Simulation Oriented Architecture (SimOA). As a concrete implementation of the framework, the basic soccer simulation algorithms investigated or used in the FootballAvatar project have been shown. The concept of avatars provides a solid basis for our simulations. It should be noted that more advanced simulation models have also been developed and implemented in our simulators. For example, fork-join soccer computations (see Fig. 7) or avatar clouds based cellular automata simulations will be presented in a further work.

(a) The time moment of starting the forked simulations.



(b) Further (heuristic) investigation of some selected passes.

Figure 7: All possible passes are simulated in different forked computations that can be seen around the large central soccer window.

8. Acknowledgments

The authors would like to thank Elemér Kondás, Sándor Szilágyi, Péter Szakály, and Tamás Sándor for contributing their professional expertise in football to this study. Sándor Szilágyi provided indispensable help with the foul model presented in Sect. 5.2. Similarly, the comments of Péter Szakály, Elemér Kondás, and Tamás Sándor regarding the biological and behavioural model (see Sect. 5.2) were essential. The authors also would like to thank Ferenc Frida and Géza Róka for their support and help. Last but not least, the authors are grateful to all members of the "Nagyerdei Gerundium" working group and other project partner companies (namely, Esantu Ltd., U1 Research Ltd., IQRS Ltd., and Satrax Ltd.) for the meetings and discussions about soccer.

NEW SZÉCHENYI PLAN

The publication was supported by the GOP-1.2.1-11-2012-0005 (*SziMe3D – 3D-s technológiai innovációk a turizmus, oktatás és sport területén*, *SziMe3D–3D technological innovation in tourism, education and sport*) project. The project has been supported by the European Union.

References

[1] N. Bátfai. Footballer and football simulation markup language and related simulation software development. *Journal of Computer Science and Control Systems*, 3(1):13–18, 2010.

[2] N. Bátfai. The socceral force. *CoRR*, abs/1004.2003, 2010. URL `http://arxiv.org/abs/1004.2003`.

[3] N. Bátfai, R. Dóczi, J. Komzsik, A. Mamenyák, Cs. Székelyhídi, J. Zákány, M. Ispány, and Gy. Terdik. Applications of a simplified protocol of RoboCup 2D soccer simulation. *Infocommunications Journal*, 5(1):15–20, 2013.

[4] N. Bátfai, P. Jeszenszky, A. Mamenyák, B. Halász, R. Besenczi, J. Komzsik, B. Kóti, G. Kövér, M. Smajda, C. Székelyhídi, T. Takács, G. Róka, and M. Ispány. Competitive Programming: a Case Study for Developing a Simulation-based Decision Support System. *Infocommunications Journal*, 8(1):24–38, 2016.

[5] K. Beck and C. Andres. *Extreme Programming Explained: Embrace Change*. Addison-Wesley, second edition, 2004. ISBN 978-0-321-27865-4.

[6] M. Buckland. *Programming Game AI By Example*. Wordware Publishing Inc., 2005. ISBN 1556220782.

[7] J. L. Casti. *Would-Be Worlds: How Simulation Is Changing the Frontiers of Science*. John Wiley & Sons, 1997.

[8] A. C. Constantinou, N. E. Fenton, and M. Neil. pi-football: A Bayesian network model for forecasting association football match outcomes. *Knowledge-Based Systems*, 36:322–339, 2012. doi: 10.1016/j.knosys.2012.07.008.

[9] J. P. Davis, C. B. Bingham, and K. M. Eisenhardt. Developing theory through simulation methods. *The Academy of Management Review*, 32(2):480–499, 2007.

[10] M. Dixon and S. Coles. Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(2):265–280, 1997. doi: 10.1111/1467-9876.00065.

[11] E. Epstein. A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology*, 8:985–987, 1969. doi: 10.1175/1520-0450(1969)008<0985:ASSFPF>2.0.CO;2.

[12] D. Forrest, J. Goddard, and R. Simmons. Odds-setters as forecasters: the case of english football. *International Journal of Forecasting*, 21(3):551–564, 2005. doi: 10.1016/j.ijforecast.2005.03.003.

[13] J. Goddard. Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2):331–340, 2005. doi: 10.1016/j.ijforecast.2004.08.002.

[14] J. Goddard and I. Asimakopoulos. Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, 23(1):51–66, 2004. doi: 10.1002/for.877.

[15] Google Inc. Protocol Buffers, 2014. URL `https://developers.google.com/protocol-buffers/`.

[16] A. Y. Govan and C. D. Meyer. Ranking national football league teams using Google's PageRank. Technical Report CRSC-TR06-19, Center for Research in Scientific Computation at North Carolina State University, 2006. URL `http://www.ncsu.edu/crsc/reports/ftp/pdf/crsc-tr06-19.pdf`.

[17] A. Y. Govan, C. D. Meyer, and R. Albright. Generalizing Google's PageRank to rank national football league teams. In *Proceedings of SAS Global Forum*, 2008. URL `http://meyer.math.ncsu.edu/Meyer/PS_Files/SASGF08RankingPaper.pdf`.

[18] D. Karlis and I. Ntzoufras. Analysis of sports data by using bivariate Poisson models. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 52(3):381–393, 2003. doi: 10.1111/1467-9884.00366.

[19] H. Kitano, M. Asada, Y. Kuniyoshi, I. Noda, and E. Osawa. RoboCup: The robot world cup initiative. In *Proceedings of the first international conference on Autonomous agents*, AGENTS'97, pages 340–347. ACM, 1997. doi: 10.1145/267658.267738.

[20] A. M. Law. *Simulation Modeling and Analysis*. Mcgraw-Hill, fifth edition, 2015.

[21] A. J. Lee. Modeling scores in the Premier League: is Manchester United really the best? *Chance*, 10(1):15–19, 1997.

[22] J. Martin. *Rapid Application Development*. Macmillan Publishing Co., 1991. ISBN 0-02-376775-8.

[23] I. McHale and P. Scarf. Modelling soccer matches using bivariate discrete distributions with general dependence structure. *Statistica Neerlandica*, 61 (4):432–445, 2007. doi: 10.1111/j.1467-9574.2007.00368.x.

[24] B. Mielczarek and J. Uzialko-Mydlikowska. Application of computer simulation modeling in the health care sector: A survey. *Simulation: Transactions of The Society for Modeling and Simulation International*, 88(2):197–216, 2012.

[25] L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999. URL `http://ilpubs.stanford.edu:8090/422/`.

[26] M. Peters, editor. *Computational Fluid Dynamics for Sport Simulation*. Lecture Notes in Computational Science and Engineering. Springer, 2009.

[27] RCSS Stamina Model, 1999. URL `http://people.dsv.su.se/~johank/RoboCup/manual/ver5.1release/browsable/browsablenode54.html`.

[28] A. Rotshtein, M. Posner, and A. Rakityanskaya. Football predictions based on a fuzzy model with genetic and neural tuning. *Cybernetics and Systems Analysis*, 41(4):619–630, 2005. doi: 10.1007/s10559-005-0098-4.

[29] H. Rue and O. Salvesen. Prediction and retrospective analysis of soccer matches in a league. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 49(3):399–418, 2000. doi: 10.1111/1467-9884.00243.

[30] J. Sewall, D. Wilkie, and M. C. Lin. Interactive hybrid simulation of large-scale traffic. *ACM Transactions on Graphics*, 30(6):135:1–135:12, 2011.

[31] P. Spagnoletti, S. Za, and R. Winter. Exploring foundations for using simulations in is research. In *Proceedings of the 34th International Conference on Information Systems (ICIS)*, 2013.

[32] SziMe3D Inc. FootballAvatar project, 2014. URL `https://footballavatar.hu`.

[33] The Australian Open Tennis Championship, 2014. URL `http://www.ausopen.com/index.html`.

[34] A. Tsakonas, G. Dounias, S. Shtovba, and V. Vivdyuk. Soft computing-based result prediction of football games. In *Proc. of the First International Conference on Inductive Modelling (ICIM'2002), Lviv*, pages 15–21, 2002.

[35] Union Cycliste Internationale. The Athlete Biological Passport, 2014. URL `http://www.uci.ch/clean-sport/the-athlete-biological-passport-abp/`.

[36] D. Wells. XP flow chart. URL `http://www.extremeprogramming.org/map/project.html`.

Annales Mathematicae et Informaticae
46 *(2016) pp. 37–53*
http://ami.ektf.hu

On the Mersenne sequence

Paula Catarino, Helena Campos, Paulo Vasco*

Universidade de Trás-os-Montes e Alto Douro, UTAD
Quinta de Prados, 5000-801 Vila Real, Portugal
pcatarin@utad.pt, hcampos@utad.pt, pvasco@utad.pt

Submitted March 8, 2016 — Accepted May 21, 2016

Abstract

From Binet's formula of Mersenne sequence we get some properties for this sequence. Mersenne, Jacobsthal and Jacobsthal-Lucas sequences are considered in order to achieve some relations between them, sums and certain products involving terms of these sequences. We also present some results with matrices involving Mersenne numbers such as the generating matrix, tridiagonal matrices and circulant type matrices.

Keywords: Mersenne numbers, Jacobsthal numbers, Jacobsthal-Lucas numbers, *g-circulant* matrix.

MSC: 11B37, 15B36, 40C05.

1. Introduction

In this paper we consider one of the sequences of positive integers satisfying a recurrence relation and we give some well-known identities for this type of sequences. One of the sequences of positive integers (also defined recursively) that have been studied over several years is the well-known Fibonacci (and Lucas) sequence. Many papers are dedicated to Fibonacci sequence, such as the works of Hoggatt, in [15], Vorobiov, in [38], and Koshy, in [32], among others. Others sequences satisfying

*The first author is member of the research centres CMAT-Polo of UTAD inserted in the research centre of the University of Minho and LabDCT/CIDTFF-research centre in Didactics and Technology in Education of Trainers; the second author is member of the research centre LabDCT/CIDTFF-research centre in Didactics and Technology in Education of Trainers; the third author is member of the research centre CMAT-Polo of UTAD inserted in the research centre of the University of Minho.

a second-order recurrence relations are the main topic of the research for several authors, such as the studies of the sequences $\{J_n\}_n$ $(n \geq 0)$ and $\{j_n\}_n$ $(n \geq 0)$ of Jacobsthal and Jacobsthal-Lucas numbers, respectively. Recall that the second-order recurrence relation and the initial conditions for the Jacobsthal numbers, J_n, $n \geq 0$, and for the Jacobsthal-Lucas numbers, j_n, $n \geq 0$, are given by

$$J_{n+2} = J_{n+1} + 2J_n, \ J_0 = 0, \ J_1 = 1 \tag{1.1}$$

and

$$j_{n+2} = j_{n+1} + 2j_n, \ j_0 = 2, \ j_1 = 1, \tag{1.2}$$

respectively. The Binet's formulas for these sequences are given by

$$J_n = \frac{2^n - (-1)^n}{3} \tag{1.3}$$

and

$$j_n = 2^n + (-1)^n, \tag{1.4}$$

respectively, where -1 and 2 are the roots of the characteristic equation associated with the above recurrence relations (1.1) and (1.2).

More recently, some of this type of sequences were generalized for any positive real number k. The studies of k-Fibonacci, k-Lucas, k-Pell, k-Pell-Lucas, Modified k-Pell, k-Jacobhstal and k-Jacobhstal-Lucas sequences are examples of this generalization and some of these studies can be found, in [3–5, 7–9, 16]. More generalizations can be found in [20, 22], where the authors considered the generalized order-k Fibonacci-Pell sequences and the Fibonacci and Lucas p-numbers for any given natural number p.

In this paper we do not have such kind of generalization, but we will follow closely some of these studies. About the Mersenne sequence, also some studies about this sequence have been published, such as the work of Koshy and Gao in [34], where the authors investigate some divisibility properties of Catalan numbers with Mersenne numbers M_k as their subscripts, developing their work in [33]. In number theory, recall that a Mersenne number of order n, denoted by M_n, is a number of the form $2^n - 1$, where n is a nonnegative number. This identity is called as the Binet formula for Mersenne sequence and it comes from the fact that the Mersenne numbers can also be defined recursively by

$$M_{n+1} = 2M_n + 1, \tag{1.5}$$

with initial conditions $M_0 = 0$ and $M_1 = 1$. Since this recurrence is inhomogeneous, substituting n by $n + 1$, we obtain the new form

$$M_{n+2} = 2M_{n+1} + 1. \tag{1.6}$$

Subtracting (1.5) to (1.6), we have that $M_{n+2} - M_{n+1} = 2M_{n+1} + 1 - 2M_n - 1$ and then

$$M_{n+2} = 3M_{n+1} - 2M_n, \tag{1.7}$$

other form for the recurrence relation of Mersenne sequence, with initial conditions $M_0 = 0$ and $M_1 = 1$. The roots of the respective characteristic equation $r^2 - 3r + 2 = 0$ are $r_1 = 2$ and $r_2 = 1$ and we easily get the Binet formula

$$M_n = 2^n - 1, \tag{1.8}$$

already given before. Note that there are Mersenne numbers prime and not prime and the search for Mersenne primes is an active field in number theory and computer science. It is now known that for M_n to be prime, n must be a prime p, though not all M_p are prime.

There are large number of sequences indexed in *The Online Encyclopedia of Integer Sequences*, being in this case

$$\begin{aligned}
\{M_n\} &= \{0, 1, 3, 7, 15, 31, 63, 127, 255, \ldots\} : A000225 \\
\{J_n\} &= \{0, 1, 1, 3, 5, 11, 21, 43, 85, 171, \ldots\} : A001045 \\
\{j_n\} &= \{2, 1, 5, 7, 17, 31, 65, 127, 257, \ldots\} : A014551.
\end{aligned}$$

The main purpose of this paper is to present some results involving the Mersenne sequence as a consequence of the respective Binet formula. Besides some relations between Mersenne, Jacobsthal and Jacobsthal-Lucas numbers allows us to obtain some new properties involving sums and products of terms of these sequences. We also present some results with matrices involving Mersenne numbers such as the generating matrix, tridiagonal matrices and circulant type matrices.

2. Mersenne sequence and some identities

According with the Binet formula (1.8) for the sequence $\{M_n\}$ we get for this sequence the following interesting identity

Proposition 2.1 (Catalan's identity)**.** *For $n \geq r$ we have*

$$M_{n-r}M_{n+r} - M_n^2 = 2^{n+1} - 2^{n-r} - 2^{n+r}. \tag{2.1}$$

Proof. Using the Binet formula (1.8) we easily obtain

$$\begin{aligned}
M_{n-r}M_{n+r} - M_n^2 &= (2^{n-r} - 1)(2^{n+r} - 1) - (2^n - 1)^2 \\
&= 2^{2n} - 2^{n-r} - 2^{n+r} + 1 - 2^{2n} + 2^{n+1} - 1 \\
&= 2^{n+1} - 2^{n-r} - 2^{n+r},
\end{aligned}$$

as required. □

Note that for $r = 1$ in Catalan's identity obtained, we get the Cassini identity for this sequence. In fact, the equation (2.1), for $r = 1$, yields

$$M_{n-1}M_{n+1} - M_n^2 = 2^{n+1} - 2^{n-1} - 2^{n+1}$$

and we get the following result

Proposition 2.2 (Cassini's identity). *For the sequence $\{M_n\}_n$ we have*

$$M_{n-1}M_{n+1} - M_n^2 = -2^{n-1}. \tag{2.2}$$

The d'Ocagne identity for this sequence can also be obtained by the use of the respective Binet formula. Hence, in this case, we get

Proposition 2.3 (d'Ocagne's identity). *For the sequence $\{M_n\}_n$ if $m > n$, we have*

$$M_m M_{n+1} - M_{m+1}M_n = 2^n M_{m-n}. \tag{2.3}$$

Proof. Once more, using the Binet formula (1.8), we get

$$\begin{aligned} M_m M_{n+1} - M_{m+1}M_n &= (2^m - 1)(2^{n+1} - 1) - (2^{m+1} - 1)(2^n - 1) \\ &= (2^{m+1} - 2^{n+1}) - (2^m - 2^n) \\ &= (2 - 1)(2^m - 2^n) \\ &= 2^m - 2^n \end{aligned}$$

and the result follows. □

Now we establish some new identities for the common factors of Mersenne numbers and both Jacobsthal and Jacobsthal-Lucas numbers. Such identities are listed in the next propositions and some of these identities involve either Mersenne and Jacobsthal numbers or Mersenne and Jacobsthal-Lucas numbers or all sequences, sums of terms, products of terms, among others.

Proposition 2.4. *If M_k, J_k and j_k are the kth term of the Mersenne sequence, Jacobsthal sequence and Jacobsthal-Lucas sequence, respectively, then*

1. *for an even subscript k,*

$$3J_k = M_k, \tag{2.4}$$
$$j_k = M_k + 2, \tag{2.5}$$

2. *and for an odd subscript k*

$$3J_k = M_k + 2, \tag{2.6}$$
$$j_k = M_k. \tag{2.7}$$

Proposition 2.5. *If M_j is the jth term of the Mersenne sequence then*

1. $M_n^2 = 4^n - M_{n+1}$;
2. $\sum_{j=0}^{n} M_j = M_{n+1} - (n+1) = 2M_n - n$;
3. $$M_n + J_n = \begin{cases} 3J_n + J_n = 4J_n, & \text{if } n \text{ is an even number} \\ 3J_n - 2 + J_n = 4J_n - 2, & \text{if } n \text{ is an odd number} \end{cases}$$

4.

$$M_n + j_n = \begin{cases} j_n - 2 + j_n = 2j_n - 2, & \textit{if } n \textit{ is an even number} \\ 2j_n, & \textit{if } n \textit{ is an odd number} \end{cases}$$

5.

$$M_n \times J_n = \begin{cases} 3J_n^2, & \textit{if } n \textit{ is an even number} \\ (3J_n - 2)J_n = 3J_n^2 - 2J_n, & \textit{if } n \textit{ is an odd number} \end{cases}$$

6.

$$M_n \times j_n = \begin{cases} (j_n - 2)j_n = j_n^2 - 2j_n, & \textit{if } n \textit{ is an even number} \\ j_n \times j_n = j_n^2, & \textit{if } n \textit{ is an odd number} \end{cases}$$

Proof. Using the Binet Formula for the Mersenne numbers, the first identity easily follows. For the second one according the equation (1.5) we have

$$\begin{aligned} M_1 &= 2M_0 + 1 \\ M_2 &= 2M_1 + 1 \\ M_3 &= 2M_2 + 1 \\ &\cdots \\ M_{n+1} &= 2M_n + 1. \end{aligned}$$

Thus if we get the sum of both terms of these equations we obtain

$$\sum_{j=1}^{n+1} M_j = 2\sum_{j=0}^{n} M_j + n + 1,$$

which implies (using the fact that $M_0 = 0$) that

$$\begin{aligned} 2\sum_{j=0}^{n} M_j &= \sum_{j=1}^{n+1} M_j - n - 1 \\ &= \sum_{j=0}^{n} M_j - n - 1 - M_0 + M_{n+1} \\ &= \sum_{j=0}^{n} M_j - n - 1 + M_{n+1} \end{aligned}$$

Hence $\sum_{j=0}^{n} M_j = M_{n+1} - (n+1)$ and the result follows by equation (1.5).

About the others identities in this proposition, we decided to omit its proof here because they can be easily proven. □

Again using the Binet formula (1.8), we obtain other property of the Mersenne sequence which is stated in the following proposition.

Proposition 2.6. *If M_n is the nth term of the Mersenne sequence, then we have*

$$\lim_{n\to\infty} \frac{M_n}{M_{n-1}} = r_1. \tag{2.8}$$

Proof. We have that

$$\lim_{n\to\infty} \frac{M_n}{M_{n-1}} = \lim_{n\to\infty} \left(\frac{2^n - 1}{2^{n-1} - 1} \right) = \lim_{n\to\infty} \left(\frac{1 - \frac{1}{2^n}}{\frac{1}{2} - \frac{1}{2^n}} \right). \tag{2.9}$$

Since $\left|\frac{1}{2}\right| < 1$, $\lim_{n\to\infty} (\frac{1}{2})^n = 0$. Next we use this fact in (2.9) obtaining

$$\lim_{n\to\infty} \frac{M_n}{M_{n-1}} = \frac{1}{\frac{1}{2}} = r_1. \qquad \square$$

Also, we easily can show the following result using basic tools of calculus of limits and (2.8).

Corollary 2.7. *If M_n is the nth term of the Mersenne sequence, then*

$$\lim_{n\to\infty} \frac{M_{n-1}}{M_n} = \frac{1}{r_1}. \tag{2.10}$$

3. Matrices with Mersenne numbers

3.1. Generating Matrix

One of the usual methods for the study of the recurrence sequences is the use of a generating matrix. The Mersenne numbers form a sequence of such type which is defined recursively as a linear combination of the preceding p terms

$$a_{n+p} = c_0 a_n + c_1 a_{n+1} + \cdots + c_{p-1} a_{n+p-1}, \tag{3.1}$$

where $c_0, c_1, \ldots, c_{p-1}$ are real constants. Some authors used the matrix method for the case of the study of some recurrence sequences of natural numbers (see, for example, [2, 6, 19–21, 23, 25], among others). Consider a square matrix M of order $p \times p$ such that the least line is $c_0, c_1, \ldots, c_{p-1}$, the $(i, i+1)-$entries, for $i = 1, \ldots, p-1$, are equal to 1 and the remaining entries are zero. According to (1.7) and (3.1) we have that $p = 2$, $c_0 = -2$ and $c_1 = 3$. Hence the matrix M is given by

$$M = \begin{pmatrix} 0 & 1 \\ c_0 & c_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}, \tag{3.2}$$

with $|M| = 2$.

Proposition 3.1. *For $n \geq 1$ we have*

$$M^n = \begin{pmatrix} -2M_{n-1} & M_n \\ -2M_n & M_{n+1} \end{pmatrix}, \tag{3.3}$$

Proof. We use induction on n. For $n = 1$,

$$M = \begin{pmatrix} -2M_0 & M_1 \\ -2M_1 & M_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix},$$

which is verified by (1.7) taking into account the initial conditions. Suppose that (3.3) is valid for n. Then

$$\begin{aligned} M^{n+1} = MM^n &= \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} -2M_{n-1} & M_n \\ -2M_n & M_{n+1} \end{pmatrix} \\ &= \begin{pmatrix} -2M_n & M_{n+1} \\ 4M_{n-1} - 6M_n & -2M_n + 3M_{n+1} \end{pmatrix} \\ &= \begin{pmatrix} -2M_n & M_{n+1} \\ -2M_{n+1} & M_{n+2} \end{pmatrix}, \end{aligned}$$

as required. □

Using this proposition and some properties involving the determinant of the matrices M and M^n, we obtain Cassini's identity in a different way than Proposition 2.2.

3.2. Special kind of tridiagonal matrix

We use tridiagonal matrices in the same way that Falcon used in [12]. A tridiagonal matrix is a matrix that has nonzero elements only on the main diagonal and on the first diagonal below and above this. Let A_n a square matrix ($n \geq 1$) of order $n \geq 1$ defined by

$$A_n = \begin{pmatrix} a & b & 0 & 0 & \cdots & 0 & 0 & 0 \\ c & d & e & 0 & \cdots & 0 & 0 & 0 \\ 0 & c & d & e & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & c & d & e \\ 0 & 0 & 0 & 0 & \cdots & 0 & c & d \end{pmatrix}, a, b, c, d, e \in \mathbb{R}.$$

Using some properties of the determinant of a tridiagonal matrix of order n, we have

$$\begin{aligned} |A_1| &= a \\ |A_2| &= d|A_1| - bc \\ |A_3| &= d|A_2| - ce|A_1| \\ &\vdots \\ |A_{n+1}| &= d|A_n| - ce|A_{n-1}|. \end{aligned}$$

If $a = 3$, $b = 2$, $c = -1$, $d = 3$ and $e = -2$, the matrix A_n is the tridiagonal matrix:

$$N_n = \begin{pmatrix} 3 & 2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 3 & -2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 3 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 3 & -2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 3 \end{pmatrix}, \tag{3.4}$$

In this case,

$$\begin{aligned} |N_1| &= 3 = M_2 \\ |N_2| &= 3|N_1| - 2 \times (-1) = 7 = M_3 \\ |N_3| &= 3|N_2| - (-1) \times (-2)|N_1| = 15 = M_4 \\ &\vdots \\ |N_{n+1}| &= 3|N_n| - (-1) \times (-2)|N_{n-1}| = 3|N_n| - 2|N_{n-1}|. \end{aligned}$$

for $n \geq 1$. Then we have the following result involving the Mersenne number of order n in terms of the determinant of a tridiagonal matrix:

Proposition 3.2. *If N_n is the $n-$by$-n$ tridiagonal matrix considered in* (3.4), *then the nth Mersenne number is given by $M_n = |N_{n-1}|$.*

Now, using the tridiagonal matrix (2.3) considered in [31] for any recurrence sequence of order two, we have that $C = M_0$, $D = M_1$, $A = 3$ and $B = -2$. Note that A and B are such that

$$x_{n+1} = Ax_n + Bx_{n-1}, n \geq 1$$

with $C = x_0$, $D = x_1$ and $\{x_n\}$ a sequence defined by this recurrence relation of order 2. The corresponding tridiagonal matrix considered in [31] for the Mersenne sequence is

$$N'_n = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & -2 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 3 & -2 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 3 & -2 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 3 \end{pmatrix}, \tag{3.5}$$

where

$$\begin{aligned} |N'_0| &= 0 = M_0 \\ |N'_1| &= \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = 1 = M_1 \end{aligned}$$

$$|N'_2| = \begin{vmatrix} 0 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & -1 & 3 \end{vmatrix} = 3 = M_2$$

$$\vdots$$

and then

Proposition 3.3. *Let N'_n is the $(n+1)-by-(n+1)$ tridiagonal matrix considered in* (3.5)*, then the nth Mersenne number is given by $M_n = |N'_n|$.*

Some authors studied the relationships between determinant or permanent of certain tridiagonal matrices and terms of recursive sequences, such as the works of Kiliç and Stanica in [24], Kiliç and Taşci in [26], [27] and [30]. Next we derive relationships between the Mersenne numbers and determinant or permanent of certain Hessenberg matrices as complementary of the results that we have already obtained. First we stated some background related with it.

The *permanent* of an n-square matrix $A = (a_{ij})$ is defined by

$$perA = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)},$$

where the summation extends over all permutations σ of the symmetric group S_n. The permanent of a matrix is analogous to the determinant, where all of the signs used in the Laplace expansion of minors are positive.

A matrix is said to be a $(0,1)$-matrix if each of its entries are 0 or 1.

In [28], the authors consider the relationships between the sums of the Fibonacci and Lucas numbers and 1-factors of bipartite graphs. In this paper, the authors consider families of square matrices such that each matrix is the adjacency matrix of a bipartite graph and the permanent of the matrix is a sum of consecutive Fibonacci or Lucas numbers. For Fibonacci and Lucas sequences, they consider special n-square $(0,1)$-matrices for which the sum of two of these matrices - one tridiagonal matrix and the other not tridiagonal - is the adjacency matrix of a bipartite graph and its permanent is a sum of consecutive elements of these sequences. Also, in [22], the authors consider certain generalizations of the Fibonacci and Lucas p-numbers. In this paper are considered certain families of square matrices such that each matrix is the adjacency matrix of a bipartite graph which its permanent is the generalized Fibonacci or Lucas p-numbers and a sum of consecutive elements of these sequences. Also, the authors consider certain matrices whose determinants are related with the Fibonacci and Lucas p-numbers and their sums. In [29] some relationships between the sums of second order linear recurrences and permanent or determinants of certain Hessenberg matrices are derived. Motivated essentially by these works, next we present some more results involving the permanent, the determinant of certain Hessenberg matrices and the sums of Mersenne numbers. Following the work of Kiliç and Taşci in [29] and [30] and if we consider the matrix

$T_n = (t_{ij})$ defined by

$$T_n = \begin{pmatrix} r_1 + r_2 & r_2 & 0 & \cdots & 0 \\ r_1 & r_1 + r_2 & r_2 & \cdots & 0 \\ 0 & r_1 & r_1 + r_2 & r_2 & \vdots \\ \vdots & \vdots & \vdots & \ddots & r_2 \\ 0 & 0 & \cdots & r_1 & r_1 + r_2 \end{pmatrix}, \tag{3.6}$$

where r_1 and r_2 are the roots of the characteristic equation associated to identity (1.7). According to the values of r_1 and r_2 and the result (1.8) in [29] we obtain, without any proof, the relationship between the determinant of a certain Hessenberg matrix and Mersenne numbers.

Proposition 3.4. *If M_j is the jth term of the Mersenne sequence and R_n is the matrix T_n in (3.6) considering the values of r_1 and r_2 for Mersenne sequence, then*

$$det R_n = M_{n+1}.$$

Next, using Theorems 1, 2 and 3 of [29] applied to the Mersenne numbers, we immediately get the relationship between the permanent of a certain tridiagonal matrix and Mersenne numbers. Note that the last part of the following proposition relates the determinant of a certain matrix with the sum of Mersenne numbers and for the last two parts we use the result obtained in Section 2, namely in item 2 of Proposition 2.5.

Proposition 3.5. *Let M_j be the jth term of the Mersenne sequence and let n be a natural number. If*

1.

$$A_n = \begin{pmatrix} 3 & -2 & 0 & \cdots & 0 \\ 1 & 3 & -2 & \cdots & 0 \\ 0 & 1 & 3 & -2 & \vdots \\ \vdots & \vdots & \vdots & \ddots & -2 \\ 0 & 0 & \cdots & 1 & 3 \end{pmatrix}, \tag{3.7}$$

then $perm A_n = M_{n+1}$;

2.

$$H_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & & & & \\ 0 & & A_n & & \\ \vdots & & & & \end{pmatrix}, \tag{3.8}$$

then $perm H_n = \sum_{i=0}^{n} M_i = 2M_n - n$;

3.

$$G_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 3 & -2 & \cdots & 0 \\ 0 & -1 & 3 & -2 & \vdots \\ \vdots & \vdots & \vdots & \ddots & -2 \\ 0 & 0 & \cdots & -1 & 3 \end{pmatrix}, \tag{3.9}$$

then $detG_n = \sum_{i=0}^{n} M_i = 2M_n - n.$

3.3. Circulant type matrix

Circulant matrices have been a great topic of research and its history and applications are vast (see, for example, [10, 14, 17, 37, 40]). All types of circulant matrices arise in the study of periodic or multiply symmetric dynamical systems and they play a crucial role for solving various differential equations (see, for example, [1, 11, 35, 39]). These matrices have been exploited to obtain the transient solution in closed form for fractional order differential equations (see, for example, [1]). Wu and Zou in [39] discussed the existence and approximation of solutions of asymptotic or periodic boundary value problems of mixed functional differential equations. In the literature, papers on several types of circulant matrices have been published (see, for example, [36]). Some authors study these type of matrices whose entries are integers that belong to sequences defined recursively. This is, for example the case of [36] and [41] where the authors considered circulant matrices with the Fibonacci and Lucas numbers and in [13] where are considered circulant matrices whose entries are Jacobsthal and Jacobsthal-Lucas numbers.

For a natural number n and a nonnegative integer g, a *g-circulant* matrix is a square matrix of order n with the following form:

$$A_{g,n} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_{n-g+1} & a_{n-g+2} & \cdots & a_{n-g} \\ a_{n-2g+1} & a_{n-2g+2} & \cdots & a_{n-2g} \\ \vdots & \vdots & \ddots & \vdots \\ a_{g+1} & a_{g+2} & \cdots & a_g \end{pmatrix}, \tag{3.10}$$

where each of the subscripts is understood to be reduced modulo n. The first row of $A_{g,n}$ is $(a_1, a_2, \ldots, a_n)$ and its $(j+1)$th row is obtained by giving its jth row a right circular shift by g positions.

If $g = 1$ or $g = n + 1$ we obtain the standard *right circulant* matrix, or simply, *circulant* matrix. Thus a right circulant matrix is written as

$$RCirc(a_1, a_2, \ldots, a_n) = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_n & a_1 & \cdots & a_{n-1} \\ \vdots & \vdots & & \vdots \\ a_2 & a_3 & \cdots & a_1 \end{pmatrix}. \tag{3.11}$$

If $g = n-1$, we obtain the standard *left circulant* matrix, or *reverse circulant* matrix. In this case we write a left circulant matrix as

$$LCirc(a_1, a_2, \ldots, a_n) = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & a_3 & \cdots & a_1 \\ \vdots & \vdots & & \vdots \\ a_n & a_1 & \cdots & a_{n-1} \end{pmatrix}. \tag{3.12}$$

Let $A_n = RCirc(M_1, M_2, \ldots, M_n)$ be a *right circulant* matrix. Using the idea of Gong, Jiang and Gao in [13], we present a determinant formula for A_n.

Theorem 3.6. *For $n \geq 1$, let $A_n = RCirc(M_1, M_2, \ldots, M_n)$ be a right circulant matrix. Then we have*

$$detA_n = (1-M_{n+1})^{n-1} + (-2M_n)^{n-2} \sum_{k=1}^{n-1} \left(\frac{1-M_{n+1}}{-2M_n} \right)^{k-1} (-2M_k). \tag{3.13}$$

Proof. For $n = 1$, $detA_1 = M_1$ satisfies (3.13). In the case of $n \geq 2$, we consider the square matrices of order n of common use in the theory of circulant matrices

$$\Gamma = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -3 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 & \cdots & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -3 & 2 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & -3 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -3 & 2 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

and

$$\Pi = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \left(\frac{-2M_n}{1-M_{n+1}}\right)^{n-2} & 0 & \cdots & 0 & 1 \\ 0 & \left(\frac{-2M_n}{1-M_{n+1}}\right)^{n-3} & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \frac{-2M_n}{1-M_{n+1}} & 1 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Note that

$$det\Gamma = det\Pi = (-1)^{\frac{(n-1)(n-2)}{2}}. \tag{3.14}$$

Calculating the product $\Gamma A_n \Pi$ we obtain

$$C = \begin{pmatrix} M_1 & \delta_n & M_{n-1} & \cdots & M_3 & M_2 \\ 0 & \delta'_n & M_{n-2} & \cdots & -2M_2 & -2M_1 \\ 0 & 0 & M_1 - M_{n+1} & & & \\ 0 & 0 & 2M_n & \ddots & & 0 \\ \vdots & \vdots & & \ddots & \ddots & \\ 0 & 0 & & & & \\ 0 & 0 & 0 & & 2M_n & M_1 - M_{n+1} \end{pmatrix},$$

where

$$\delta_n = \sum_{k=1}^{n-1} \left(\frac{2M_n}{M_1 - M_{n+1}} \right)^{n-(k+1)} M_{k+1}$$

and

$$\delta'_n = (M_1 - M_{n+1}) + \sum_{k=1}^{n-1} \left(\frac{-2M_n}{M_1 - M_{n+1}} \right)^{n-(k+1)} (-2M_k). \tag{3.15}$$

Calculating the determinant of the matrix $C = \Gamma A_n \Pi$ we obtain

$$detC = M_1 \delta'_n (M_1 - M_{n+1})^{n-2}. \tag{3.16}$$

Using the identity (3.15), the recurrence relation (1.7) with the initial condition ($M_1 = 1$) and doing some calculations, a new expression for the determinant (3.16) is given by

$$detC = (1 - M_{n+1})^{n-1} + (-2M_n)^{n-2} \sum_{k=1}^{n-1} \left(\frac{1 - M_{n+1}}{-2M_n} \right)^{k-1} (-2M_k). \tag{3.17}$$

Using the property of the determinant of a product of matrices and the identity (3.14), we conclude that

$$detA_n = detC$$

and the result follows. □

Let $B_n = LCirc(M_1, M_2, \ldots, M_n)$ be a *left circulant* matrix whose entries are Mersenne numbers. We present a determinant formula for B_n using again the idea of Gong, Jiang and Gao in [13]. Lemma 5 in [18] will help us to obtain this formula. In Lemma 5 of [18] the authors define a matrix Δ which is an orthogonal cyclic shift matrix (and a left circulant matrix) of order n. They stated that

$$LCirc(a_1, a_2, \ldots, a_n) = \Delta RCirc(a_1, a_2, \ldots, a_n). \tag{3.18}$$

Using the fact that $det\Delta = (-1)^{\frac{(n-1)(n-2)}{2}}$, calculating the determinant in both sides of the identity (3.18) and according to the result obtained in Theorem 3.6, the following result is easily proved.

Theorem 3.7. *For $n \geq 1$, let $B_n = LCirc(M_1, M_2, \ldots, M_n)$ be a left circulant matrix. Then we have*

$$detB_n = (-1)^{\frac{(n-1)(n-2)}{2}} \left((1-M_{n+1})^{n-1} + (-2M_n)^{n-2} \sum_{k=1}^{n-1} \left(\frac{1-M_{n+1}}{-2M_n} \right)^{k-1} (-2M_k) \right).$$

Let us consider $C_n = A_{g,n}$ be a g-circulant matrix defined as in (3.10), whose entries are Mersenne numbers. We present the determinant formula of C_n and for that we use Lemma 6 and Lemma 7 of [18]. Thus, from these Lemmas and Theorem 3.6, we deduce the following result

Theorem 3.8. *Let $C_n = A_{g,n}$ be a g-circulant matrix defined as in (3.10), whose entries are Mersenne numbers. Then one has*

$$detC_n = det\mathbb{Q}_g[(1-M_{n+1})^{n-1} + (-2M_n)^{n-2} \sum_{k=1}^{n-1} \left(\frac{1-M_{n+1}}{-2M_n} \right)^{k-1} (-2M_k)],$$

where $\mathbb{Q}_g$ is a g-circulant matrix with the first row $e^* = [1, 0, \ldots, 0]$.

4. Conclusions

Sequences of integer numbers have been studied over several years, with emphasis on studies of the well known Fibonacci sequence (and then the Lucas sequence) that is related to the golden ratio and of the Pell sequence that is related to the silver ratio. In this paper we also contribute for the study of Mersenne sequence giving some identities which some of them involve Jacobsthal and Jacobsthal-Lucas numbers.

Several studies involving all types of circulant matrices and tridiagonal matrices can easily be found in the literature. Here we have considered the g-circulant, right and left circulant matrices whose entries are Mersenne numbers. For these cases we have provided the determinant of these matrices.

In the future, we intend to discuss the invertibility of these circulant type matrices associated with these sequence, such as Shen, in [36], did in the case of Fibonacci and Lucas numbers.

Acknowledgements. The authors would like to thank the referee for his/her constructive criticism, for pertinent comments and valuable suggestions, which significantly improve the manuscript. Also this research was financed by Portuguese Funds through FCT-Fundação para a Ciência e a Tecnologia, within the Projects UID/MAT/00013/2013 and PEst-C/CED/UI0194/2013.

References

[1] Ahmed, E., Elgazzar, A. S., On fractional order differential equations model for nonlocal epidemics, *Phys. A*, Vol. 379(2) (2007), 607–614.

[2] Borges, A., Catarino, P., Aires, A. P., Vasco, P., Campos, H., Two-by-two matrices involving k-Fibonacci and k-Lucas sequences, *Appl. Math. Sci. (Ruse)*, Vol. 8(34) (2014), 1659–1666.

[3] Campos, H., Catarino, P., Aires, A. P., Vasco, P., Borges, A., On some Identities of k-Jacobsthal-Lucas Numbers, *Int. J. Math. Anal. (Ruse)*, Vol. 8(10) (2014), 489–494.

[4] Catarino, P., On Some Identities and Generating Functions for k-Pell Numbers, *Int. J. Math. Anal. (Ruse)*, Vol. 7(38) (2013), 1877–1884.

[5] Catarino, P., On some identities for k-Fibonacci sequence, *Int. J. Contemp. Math. Sci.*, Vol. 9(1) (2014), 37–42.

[6] Catarino, P., On generating matrices of the k-Pell, k-Pell-Lucas and modified k-Pell sequences, *Pure Math. Sci.*, Vol. 3(2) (2014), 71–77.

[7] Catarino, P., Vasco, P., Modified k-Pell Sequence: Some Identities and Ordinary Generating Function, *Appl. Math. Sci. (Ruse)*, Vol. 7(121) (2013), 6031–6037.

[8] Catarino, P., Vasco, P., On some Identities and Generating Functions for k-Pell-Lucas sequence, *Appl. Math. Sci. (Ruse)*, Vol. 7(98) (2013), 4867–4873.

[9] Catarino, P., Vasco, P., Some basic properties and a two-by-two matrix involving the k-Pell Numbers, *Int. J. Math. Anal. (Ruse)*, Vol. 7(45) (2013), 2209–2215.

[10] Davis, P. J., *Circulant Matrices*, John Wiley & Sons, New York, 1979.

[11] Erbas, C., Tanik, M. M., Generating solutions to the N-Queens problems using 2-circulants, *Math. Mag.*, Vol. 68(5) (1995), 343–356.

[12] Falcon, S., On the generating matrices of the k-Fibonacci numbers, *Proyecciones*, Vol. 32(4) (2013), 347–357.

[13] Gong, Y., Jiang, Z., Gao, Y., On Jacobsthal and Jacobsthal-Lucas Circulant Type Matrices, *Abstr. Appl. Anal.*, (2014), Article ID 418293, 11 pages.

[14] Good, I. J., On the inversion of circulant matrices, *Biometrika*, Vol. 37(1-2) (1950), 185–186.

[15] Hoggatt, V. E., *Fibonacci and Lucas Numbers.* A publication of the Fibonacci Association. University of Santa Clara, Santa Clara, Houghton Mifflin Company, 1969.

[16] Jhala, D., Sisodiya, K., Rathore, G. P. S., On Some Identities for k-Jacobsthal Numbers, *Int. Journal of Math. Anal.*, Vol. 7(12) (2013), 551–556.

[17] Jiang, Z. L., Zhou, Z. X., *Circulant Matrices*, Chengdu Technology University Publishing Company, Chengdu, 1999.

[18] Jiang, Z., Gong, Y., Gao, Y., Circulant type matrices with the sum and product of Fibonacci and Lucas numbers, *Abstr. Appl. Anal.*, Vol. 2014, Article ID 375251, 12 pages, 2014.

[19] Kalman, D., Generalized Fibonacci numbers by matrix methods, *Fibonacci Quart.*, Vol. 20(1) (1982), 73–76.

[20] Kiliç, E., The generalized order-k Fibonacci-Pell sequence by matrix methods, *J. Comput. Appl. Math.*, Vol. 209(2) (2007), 133–145.

[21] Kiliç, E., Ömür, N., Ulutaş, Y., Matrix representation of the second order recurrence $\{u_{kn}\}$, *Ars Combin.*, Vol. 93 (2009), 181–190.

[22] Kiliç, E., Stakhov, A.P., On the Fibonacci and Lucas p-numbers, their sums, families of bipartite graphs and permanents of certain matrices, *Chaos Solitons Fractals*, Vol. 40(5) (2009), 2210–2221.

[23] Kiliç, E., Stanica, P., A matrix approach for general higher order linear recurrences, *Bull. Malays. Math. Sci. Soc.*, (2) Vol. 34(1) (2011), 51–67.

[24] Kiliç, E., Stanica, P., Factorizations and representations of binary polynomial recurrences by matrix methods, *Rocky Mountain J. Math.*, Vol. 41(4) (2011), 1247–1264.

[25] Kiliç, E., Taşci, D., The generalized Binet formula, representation and sums of the generalized order-k pell numbers, *Taiwanese J. Math.*, Vol. 10(6) (2006), 1661–1670.

[26] Kiliç, E., Taşci, D., Factorizations and representations of the backward second order linear recurrences, *J. Comput. Appl. Math.*, Vol. 201(1) (2007), 182–197.

[27] Kiliç, E., Taşci, D., On the permanents of some tridiagonal matrices with applications to the Fibonacci and Lucas numbers, *Rocky Mountain J. Math.*, Vol. 37(6) (2007), 1953–1969.

[28] Kiliç, E., Taşci, D., On families of bipartite graphs associated with sums of Fibonacci and Lucas numbers, *Ars Combin.*, Vol. 89 (2008), 31–40.

[29] Kiliç, E., Taşci, D., On sums of second order linear recurrences by hessenberg matrices, *Rocky Mountain J. Math.*, Vol. 38(2) (2008), 531–544.

[30] Kiliç, E., Taşci, D., On the second order linear recurrences by tridiagonal matrices, *Ars Combin.*, Vol. 91 (2009), 11–18.

[31] Kiliç, E., Taşci, D., Haukkanen, P., On the generalized Lucas sequences by Hessenberg matrices, *Ars Combin.*, Vol. 95 (2010), 383–395.

[32] Koshy, T., *Fibonacci and Lucas Numbers with Applications*, John Wiley, New York, 2001.

[33] Koshy, T., Gao, Z., Some divisibility properties of Catalan numbers, *Math. Gazette*, Vol. 95(532) (2011), 96–102.

[34] Koshy, T., Gao, Z., Catalan numbers with Mersenne subscripts, *Math. Scientist*, Vol. 38(2) (2013), 86–91.

[35] Qu, W., Lei, S. L., Vong, S. W., Circulant and skew-circulant splitting iteration for fractional advection-diffusion equations, *Int. J. Comput. Math.*, Vol. 91(10) (2014), 2232–2242.

[36] Shen, S. Q., Cen, J. M., Hao, Y., On the determinants and inverses of circulant matrices with Fibonacci and Lucas numbers, *Appl. Math. Comput.*, Vol. 217(23) (2011), 9790–9797.

[37] Stallings, W. T., Boullion, T. L., The pseudoinverse of an r-circulant matrix, *Proc. Amer. Math. Soc.*, Vol. 34(2) (1972), 385–388.

[38] Vorobiov, N. N., *Números de Fibonacci*, Editora MIR, URSS, 1974.

[39] Wu, J., Zou, X., Asymptotic and periodic boundary value problems of mixed FDEs and wave solutions of lattice differential equations, *J. Differential Equations*, Vol. 135(2) (1997), 315–357.

[40] Wu, Y. K., Jia, R. Z., Li, Q., g-circulant solutions to the $(0, 1)$ matrix equation $A^m = J_n$, *Linear Algebra Appl.*, Vol. 345(1-3) (2002), 195–224.

[41] Zhou, J., Jiang, Z., The spectral norms of g-circulant matrices with classical Fibonacci and Lucas numbers entries, *Appl. Math. Comput.*, Vol. 233 (2014), 582–587.

Annales Mathematicae et Informaticae
46 *(2016) pp. 55–75*
http://ami.ektf.hu

Case study for the vudc R Package

Csaba Faragó

Department of Software Engineering, University of Szeged, Hungary
farago@inf.u-szeged.hu

Submitted February 22, 2016 — Accepted October 11, 2016

Abstract

In this study we present the usage of Cumulative Characteristic Diagram and Quantile Difference Diagram – implemented in the vudc R package – using the results of our research on the connection between version control history data and the related maintainability.

With the help of these diagrams, we illustrate the results of five studies, in which we executed contingency Chi-Squared test, Wilcoxon rank tests and variance test. We were motivated by the question: how did these diagrams support the numeric results?

We found that the diagrams spectacularly supported the results of the statistic tests, furthermore, they revealed other important connections which were left hidden by the tests.

Keywords: univariate data, data visualization, maintainability, Cumulative Characteristic Diagram, Quantile Difference Diagram

1. Introduction

A researcher might run into trouble how to present the promising results. Just writing texts and presenting tables might not be convincing enough. We ran into similar problem in our research. Therefore, we created two diagram types – the *cumulative characteristic diagram* and the *quantile difference diagram* – and presented in paper *Visualization of Univariate Data for Comparison* [1].

In our research we discovered some connections between version controls history data and maintainability of the source code. First, we considered the number of version control operations, i.e. how many Java files have been added, updated and deleted in a certain commit. Then we considered the cumulative code churn of the files in every commit, i.e. the absolute sum of number of lines added and deleted from the beginning of the available revision history. After that, we checked how the code ownership affected maintainability. In this paper we apply the above mentioned diagrams on the already published results.

In Section 1.1, we present how we calculated the maintainability. In Section 1.2, we summarize the results we want to visualize. In Section 1.3, we form research questions. In Section 2, we present the work related to this one. In Section 3, we summarize the already presented `vudc` package, now with the emphasis on the extended project data. Section 4 is the core of this paper, containing all the diagrams raised in the research questions. In this section, we answer the research questions, one by one. Finally, in Section 5, we summarize the results, and conclude the paper.

1.1. Calculating maintainability

For the measurement of the maintainability of source code, we used the Columbus Quality Model. This model is based on the fact that source code metrics – like logical lines of code, complexity, coding rule violations, or the object-oriented metrics defined in article by Chidamber and Kemerer [2] – affect the maintainability. In their study Gyimóthy et al. [3] showed that the higher values of these metrics resulted in higher number of post-release bugs.

Bakota et al. [4] published the Columbus Quality Model. They referred to the ISO/IEC 9126 standard [5], which defines six high-level characteristics that determine the product quality of software: functionality, reliability, usability, efficiency, portability, and maintainability. The model determines relevant source code metrics, and compares them with the same metrics found in a benchmark. Then it aggregates the results of the comparisons, and as a final result it returns a number indicating the maintainability of the checked source code.

The model we use was implemented to analyze Java source code. Hegedűs [6] adopted the algorithm to C#. Hegedűs et al. [7] extended the base algorithm to be able to determine the relative maintainability index of every source code element, like class or function.

1.2. Version control history metrics and maintainability

We opened the black box of this research area in work [8]. In that study, we showed that connection between version control operations and maintainability change existed.

Then we analyzed the version control operations (Add, Update, Delete) one by one. First, in study [9] we checked how these version control operations affect the maintainability change. We concluded that file additions improved, or at least less eroded the maintainability than file modifications, the file updates mainly eroded them, while we could not establish the effect of the file deletion. Then in article [10], we checked the version control operations considering how they affect the variance of the maintainability change. We concluded that the file addition and file deletion increased, while the file update decreased the variance. As the amplitude was much bigger than the absolute change, as a final conclusion we stated that it was recommended to pay special attention to file additions.

In work [11], we checked how code ownership affected the maintainability change of the source code, caused by the actual commit. We concluded that the lack of clean code ownership indicated the future decrease of the maintainability.

In study [12], we checked the effect of the intensity of past code modification intensity on the maintainability change of future commits. As result we gained that intensive past code modifications was likely to cause further code decay, compared to modifying files which had been less intensively modified in the past.

In article [13], we defined the following six version control history metrics: cumulative code churn, number of modifications, ownership, ownership with tolerance, code age, and last modification time. For a certain version of the analyzed system we sorted the source files, based on every metrics. We determined the order of files on the relative maintainability index basis; furthermore, as a cross-check, we determined the order based on the number of post-release bugs as well. We determined the correlation between these orders with help of the Spearman's rank correlation test. As result we got that higher intensity of modifications, the higher number of code modifications and developers (without and with tolerance), the older code and the later last modification date resulted lower maintainability and higher number of post-release bugs.

Putting together the code ownership, code churn and metrics result, we concluded that what was already wrong, it was more likely to become even worse, compared to the source code of better maintainability.

1.3. Research questions

In this study, we try to answer the following research questions by visualization:

RQ1: *Illustrate the operation and maintainability change based contingency tables – as we presented in article [8] – using cumulative characteristic diagram. Do the created diagrams support the published results?*

RQ2: *Consider commits containing file addition on one hand, and commits not containing file additions on the other hand. Create cumulative characteristic diagrams and quantile difference diagrams using the related maintainability change values. Do these diagrams support the results we published in study [9], considering the value of maintainability change?*

RQ3: *Considering the same data as in case of RQ2, do the mentioned diagrams support the results we published in study [10], about the variance of maintainability change?*

RQ4: *Create quantile difference diagrams using the cumulative churn values, related to positive and negative maintainability changes, as we described in study [12]. How do these diagrams support the published results?*

RQ5: *Create quantile difference diagrams using the ownership values, related to positive and negative maintainability changes, as we described in study [11]. How these diagrams support the published results?*

2. Related Work

Several diagram types exists for illustrating univariate, bivariate and multivariate data. The book by Chambers et al. [14] provides a summary of the most important possibilities. The book by Murrel [15] focuses on the diagram creating possibilities in R statistical programming language [16].

One of the most frequent diagram type for illustrating univariate numeric data is the box plot. However, as it became very popular, researchers faced its shortcomings. Several proposals appeared to make it better.

McGill et al. [17] suggested variable width and notched box plots. Not to forget that those times the computers were expensive and slow, and the diagrams were mostly drawn by hand. Benjamini [18] exploited the capability of the computer. Frigge et al. [19] dealt mainly with the problem of outliers. Potter et al. [20] provided a summary of the variations of box plots.

Probably the most important problem with box plot is that it hides the local densities. To overcome this shortcoming in R, the density plot could be a good choice in several cases. Other popular diagram types handling this issue are violin plots (R function `vioplot()` in package `vioplot` [21], presented by Hintze and Nelson [22]) and bean plots (R function `beanplot()` in package `beanplot`, described by Kampstra [23]).

The problem of illustrating bivariate data is also very common. Goldberg and Iglewicz [24] presented an early proposal of a bivariate extension of boxplots. An interesting two dimensional extension of the box plots is the bag plot, as article by Rousseeuw et al. [25] suggests. For the implementation, see R function `bagplot()` in package `aplpack` [26].

Visualizing multivariate data is even harder. Hornik et al. [27] suggested a framework for visualizing multi-way contingency tables.

The presented R functions are mainly based on base package `graphics`. Another basic visualization related package in R is `grid`. The `lattice` package is based on `grid`; it was presented by Sarkar [28].

3. The `vudc` R package

To recall, in the study [1], we presented the `vudc` R package [29]. This package contains the implementation of two diagram types: the Cumulative Characteristic Diagram (CCD) and Quantile Difference Diagram (QDD), furthermore, data for the case study presented in this paper.

In this section, we summarize the two diagram types, and the data provided to the package. We have not changed the function since its first version, while we extended the data with information used in recent research.

3.1. Cumulative Characteristic Diagram (CCD)

The input of the base diagram is a set of numbers. In the first step, we sort these numbers non-ascending. Then we calculate cumulatives for every element: the series starts with 0, the next element will be the value of the first element of the sorted array, the second element will be the sum of the first 2 elements, and so on. In the diagrams, the x-coordinate represents the number of elements, and the y-coordinate represents the calculated cumulatives. Instead of drawing each point one by one, we connect these points with straight lines. If the number of elements is high enough, the result will look like a continuous line without bends.

The diagram type is mostly suitable for data of normal distribution with the mean close to 0. The diagram is applicable for quick comparison of several data sets: to illustrate the similarities and differences. It can be used to illustrate quickly two or more – seemingly similar – data sets if they are really similar or not.

A CCD which contains two or more characteristics on the same diagram we call *Composite Cumulative Characteristic Diagram.*

The R implementation of CCD is the function `ccdplot`. For detailed usage, we refer the help page of the function. We present examples for CCD in Section 4 in Figures 3, 4, and 5.

3.2. Quantile Difference Diagram (QDD)

The idea behind the Quantile Difference Diagrams is to compare the same quantiles of two sets of numbers. This means the first element of the first set should be compared to the first element of the second one, similarly the 10% to the 10%, the median to the medial, the 90% to the 90%, highest to the highest and so on.

Therefore, the input of the QDD is always two sets of numbers. We determine every centile in both subsets, i.e. the 0% (which is the lowest one), the 1% (e.g. if the set contains 1000 elements, this is the 10^{th}) etc. This results 101 values in every case, either by omitting values, or taking the same values several times. Then we calculate the differences at every centile. We display these differences on the diagram as a line.

The R implementation of QDD is the function `qddplot`, and the detailed usage can be found in the help page. We present examples for QDD in Section 4 in Figures 6, 7, and 8.

3.3. Project data

We attached the data necessary for reproducing the diagrams found in this case study into the `vudc` R package, to data structure `projectdata`. It contains information about the following software systems:

- **Ant** – a command line tool for building Java applications (`http://ant.apache.org`).
- **Gremon** – a greenhouse work-flow monitoring system (commercial; `http://www.gremonsystems.com`).

- **Struts 2** – a framework for creating enterprise-ready java web applications (`http://struts.apache.org/`).
- **Tomcat** – an implementation of the Java Servlet and Java Server Pages technologies (`http://tomcat.apache.org`).

For each project, we provided a *data frame* containing information of every available commit. The rows of the data frame represent commits, and there are the following columns:

- `Revision`: the original revision number in the version control system
- `MaintainabilityDiff`: maintainability difference of the actual and the previous commit
- `A`: number of added Java files in the commit
- `U`: number of updated Java files in the commit
- `D`: number of deleted Java files in the commit
- `Churn`: a real number representing the code churn value of the commit
- `Ownership`: a real number representing the ownership of the commit

We removed the commits not containing Java files.

In order to be able to identify the commit – especially for the open source systems – we added the revision number to the data, exactly as it is located in the version control system.

The `MaintainabilityDiff` is the difference of maintainability values of two subsequent revisions. We calculated the maintainability of every revision with the help of Columbus Quality Model [4]. Then we normalized these maintainability values and we calculated the difference as we described in paper [9]. The final result is a real number.

The number of added, updated and deleted files are non-negative integers, containing information about Java files.

Cumulative code churn of a file is the absolute sum of the number of added and removed lines of sources, from the beginning of the available revision history, up to the examined revision, but excluding that one. Cumulative code churn of a commit is the average of the actual cumulative code churn values of files appearing in the commit.

Ownership of a file is the number of different contributors, from the beginning of the available revision history, up to the examined revision, but now including. Ownership of a commit is the geometric mean of the actual ownership values of files appearing in the commit.

The `vudc` package contains revision related information only, and no file related information.

4. Case study

In this section, we illustrate the usage of the diagrams in our research. The source code of all the diagrams presented in this section, furthermore, the statistic tests are found in the help of the `vudc` package, as follows:

```
library(vudc)
?projectdata
```

4.1. Connection between version control operations and quality change of the source code

In the study [8], we divided the commits based on the number of operations into the following 4 disjoint subsets:

- **D** – commits containing at least one delete,
- **A** – commits containing at least one add but no delete,
- **U+** – commits containing neither add nor delete, and containing at least 2 updates,
- **U1** – commits consisting of exactly one update.

On the other hand, another dimension of the division we performed based on maintainability change values, into 3 subsets: positive (maintainability increase), zero (no traceable maintainability change) and negative (maintainability decrease).

This resulted a table of dimensions 4 and 3, with 12 cells. Each commit belongs to exactly one cell. We counted how many commits a cell contained. Then we performed the 2 dimensional Contingency Chi-Squared test with the null-hypothesis that the commits were proportionally distributed in the cells, using the `chisq.test()` R function.

In Table 1, we present the overall p-values of every analyzed systems.

Project	**p-value**
Ant	$1.60 \cdot 10^{-151}$
Gremon	$1.19 \cdot 10^{-52}$
Struts 2	$4.47 \cdot 10^{-64}$
Tomcat	$4.84 \cdot 10^{-33}$

Table 1: Overall p-values of the contingency Chi-squared tests

To summarize, the results were significant, i.e. there were hardly any cells with no significant deviation from the expected value; furthermore, the values in the same cells of different projects tended to deviate from the null-hypothesis in the

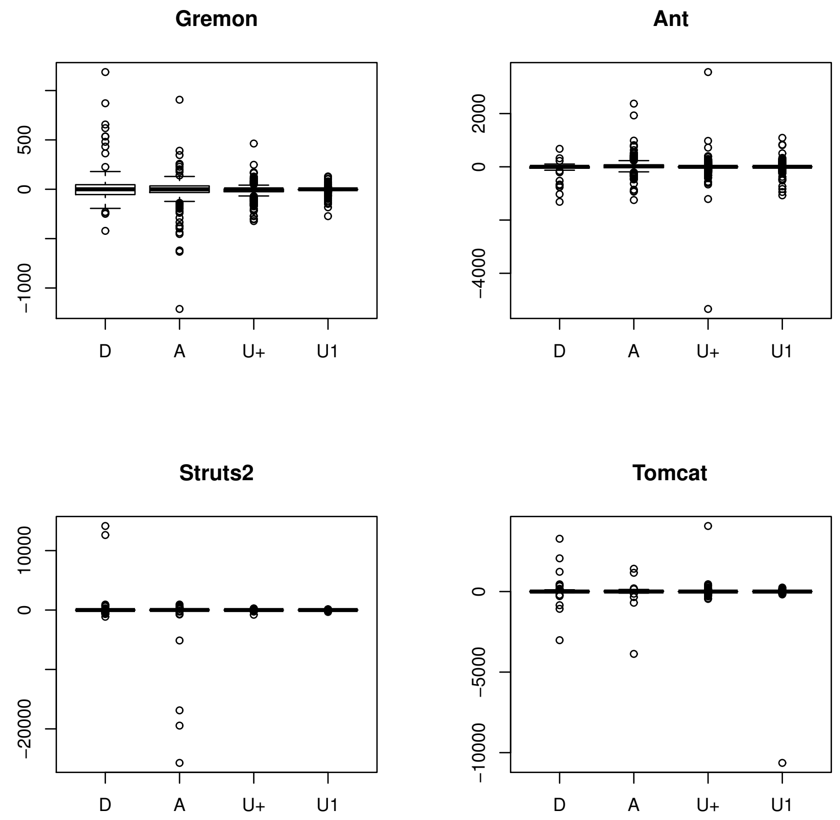


Figure 1: Research data using box plots

same direction. Therefore, we found clear connection between commit operations and maintainability changes.

We wanted to somehow visualize the input data of the tests to make the differences obvious. The most straightforward choice was the box plot diagram; however – as seen in Figure 1 – we found it not really useful.

We noticed that the outliers had significant bias on the diagrams. Some unusual commits, like merging a whole branch to the trunk, or renaming files in two steps (first remove, and then in another commit add again) resulted in huge outliers. We removed the effect of these extraordinary commits by removing the huge values (absolute values being higher than 1000.0). The results became slightly better (see Figure 2), but still not spectacular enough.

In Figure 3, we illustrate the values as already presented in Figure 1, but now using the Cumulative Characteristic Diagrams.

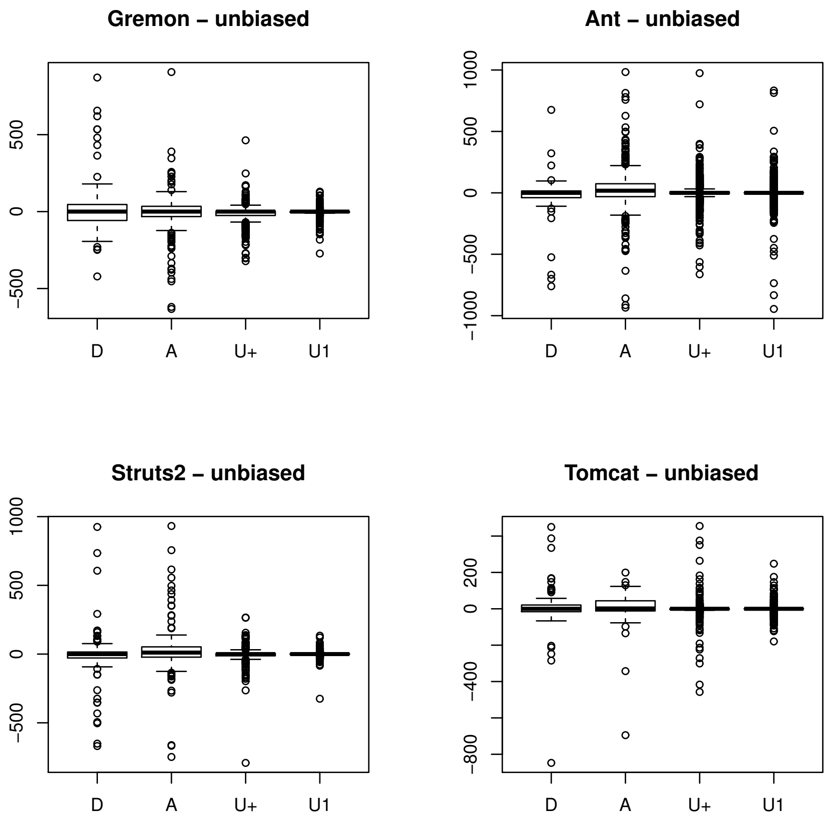


Figure 2: Research data using box plots, without outliers

Note that the outliers have significant bias on this diagram as well. See for example the characteristic of operation Add in case of Struts 2. By removing these values we receive more concise diagrams presented in Figure 4.

The curves within diagrams are obviously different, and there are similarities between the diagrams. The following can be deduced from these diagrams after a short analysis.

Overall characteristic

All the characteristics start with a precipitous rising, continuous with a relatively long horizontal part and ends with a slightly less precipitous slope. If the right end is located below 0, it means the net effect of all the commits was negative from maintainability point of view; if it is located above 0 then the opposite is true. Based on the difference in the slope of the left and the right part, we can

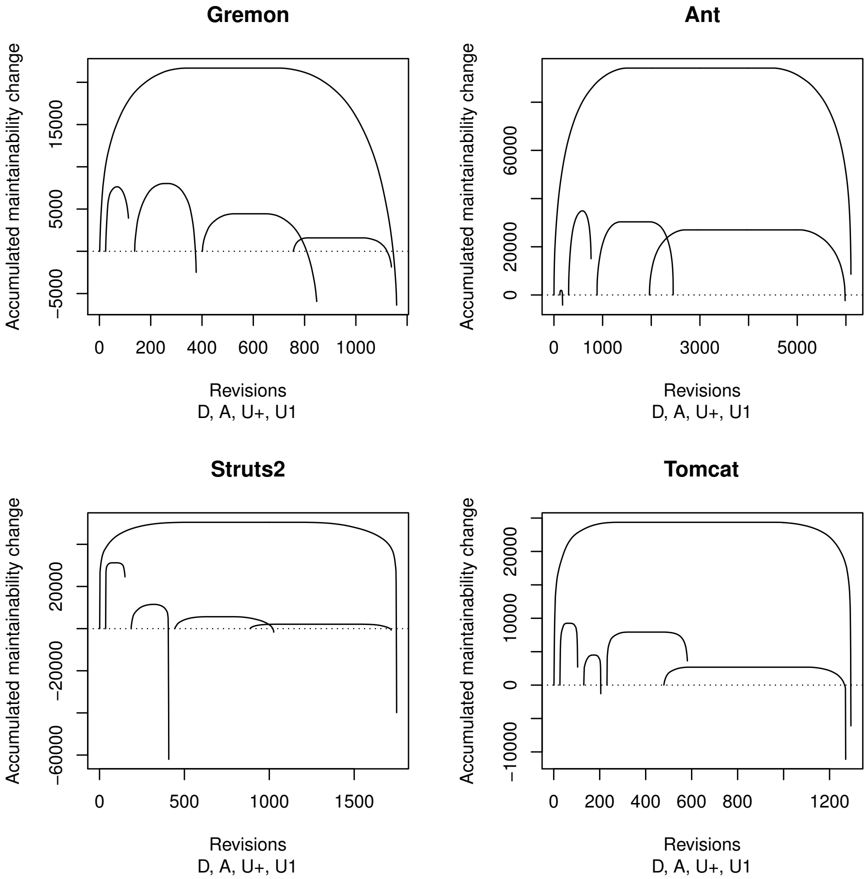


Figure 3: Composite cumulative characteristic diagrams about maintainability

conclude that the maintainability increase is rather caused by smaller number of bigger steps, while maintainability decrease is caused by a bit higher number of a bit less steps. Note that this result was not identified with the help of statistical tests.

Commits containing Delete

The number of elements of this type of commits is relatively small. In case of Ant, it is practically negligible. But the relative height is very big; the magnitude of its height on the CCD diagram is similar to those of other types with much higher number of commits. This indicates that the variance caused by operation delete is much higher than those commits not containing this operation, as shown later in Section 4.3. On the other hand, the right end seems to be hectic, therefore

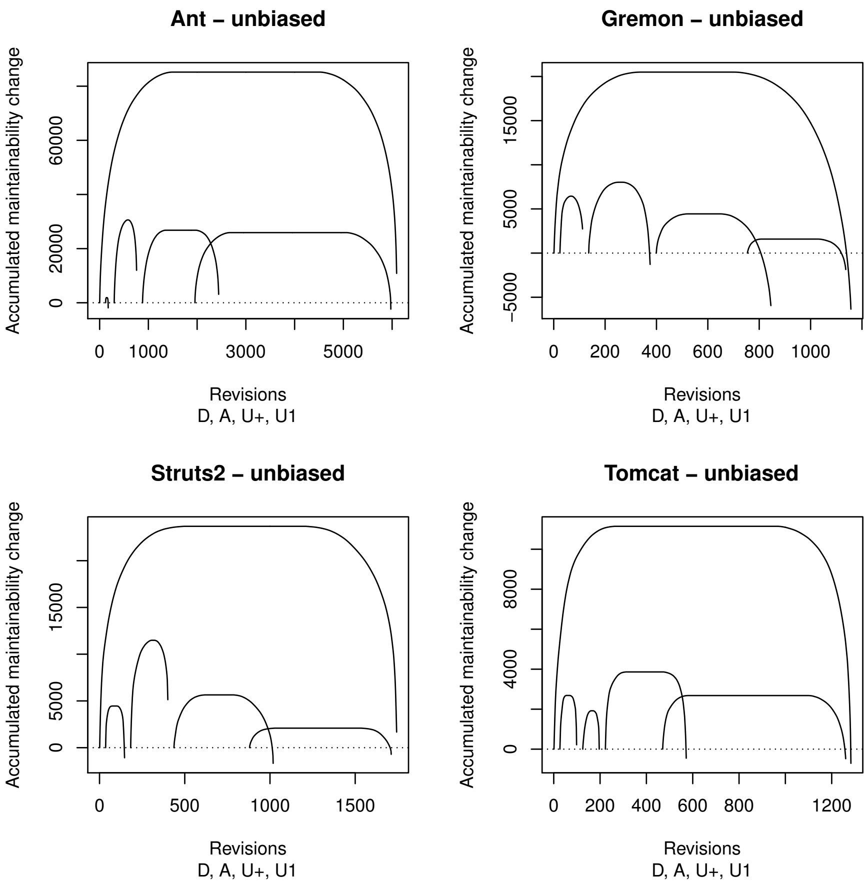


Figure 4: Composite cumulative characteristic with removed outliers

we cannot form clear statement about this operation.

Commits containing Add without any Delete

There are some spectacularly similar properties of the second characteristic in all projects. First of all – considering the characteristics without outliers – the right end of this characteristics is located above the composite one, or those containing exclusively file updates. In 3 out of the 4 cases it was positive as well. This implies that the operation Add has a good, or at least better effect on the maintainability than the others. The other spectacular property is – similarly to operation Delete – the relative height of the characteristic. Despite of its small width it is high; in 3 out of the 4 cases it is higher than the much wider Update related characteristic. Again, this visually represents the high variance of the maintainability caused by operation Add. Finally, the horizontal part in the middle is negligible in all of the cases, meaning that Add had some traceable effect on maintainability in most of the cases.

Commits consisting of several Updates

Commits consisting of several Updates (these are typically smaller feature developments or bigger bug fixes) have some typical characteristics. Probably the most obvious common attribute of them is that their width is relatively large; greater than the joint width of operation Delete and Add. The right end is always located lower than the right end of the common characteristic, meaning that this type of commit tends to decrease the overall maintainability (as was confirmed with the help of statistical tests). Also the horizontal part in the middle is significant, meaning the number of commits with no traceable maintainability change is relatively high in this category. Finally, the relative height is smaller than in the case of the first two curves, but bigger than the fourth one.

Commits consisting of exactly one Update

This is the most frequent commit type, this fact is very spectacular in 3 out of the 4 cases. These commits are typically smaller bug fixes. The relative height is small, i.e. the variance caused by this type of commit is low. The horizontal part in the middle is very long in all of the 4 cases, again meaning that the proportion of commits with no traceable maintainability change caused by this type of commit is high. It is also important that the right end is located below 0 in all of the cases, meaning that the net effect of this commit type is always negative.

Answer to RQ1

Figure 3 contains the cumulative characteristic diagram related to study [8]. This diagram lead to the conclusion that the data contain outliers which have drastic impact of the results. This fact led us to the decision later not to perform t-test but Wilcoxon-test. Figure 4 contains the CCD of the data without outliers. The cumulative lines related to subsets differ, and therefore this diagram support the published results. Furthermore, the curves related to the same category resemble to each other.

4.2. The impact of version control operations on the quality change of the source code

The results presented in study [8] and outlined in Section 4.1 convinced us that it was worth to perform detailed analysis on all the 3 occurring version control operation: Add, Update and Delete. We performed the analysis and presented the results [9]. In the analysis. we divided the commits based on the number or proportion of occurrences of the actually analyzed operation, in each analyzed system, and performed Wilcoxon test to compare the maintainability change values related to those commits of these two subsets.

Here we present the following check: we divided all the commits into two based on the existence of operation Add. The first subset contained commits where at least one new source file was added, and the second one contained the remaining commits. We considered the related maintainability changes, performed the Wilcoxon test on the mentioned sets of numbers, and concluded that the maintainability change of commits containing file addition was significantly higher than those of not containing file addition.

Table 2 contains the resulting p-values, Figure 5 contains the CCD of the same values, and Figure 6 contains the related QDD.

System	**p-value**
Ant	$3.54 \cdot 10^{-16}$
Gremon	$3.79 \cdot 10^{-2}$
Struts 2	$5.27 \cdot 10^{-7}$
Tomcat	$7.18 \cdot 10^{-5}$

Table 2: Difference in maintainability change values of commits containing and not containing file addition

In the Cumulative Characteristic Diagram the right end of the curves for commits containing Add are located spectacularly higher than these of the other curve, which supports the results of the Wilcoxon test.

The Quantile Difference Diagrams revealed some more important details. The territory above x-coordinate is spectacularly higher in case of Ant, Struts 2 and Tomcat. The Wilcoxon test resulted that it is true also in case of Gremon, but with weaker significance.

Indeed, the diagrams support the findings based just on the Wilcoxon test, furthermore, it revealed such details which we analyzed further and present in Section 4.3.

Answer to RQ2

We presented the cumulative characteristic diagram in Figure 5, and the quantile difference diagram in Figure 6, related to the impact of operation add on maintainability. On the CCDs the right end of the left hand side curve is located higher than the right end of the right hand side curve, and this supports the result presented in study [9]. It is important to note that the line is located below the x-axis for the leftmost 30-50%, and it means the operation add is present in the commits causing big code decay as well. On the QDDs the territory above the x-axis is higher than the territory below the x-axis which also supports the result presented in study [9].

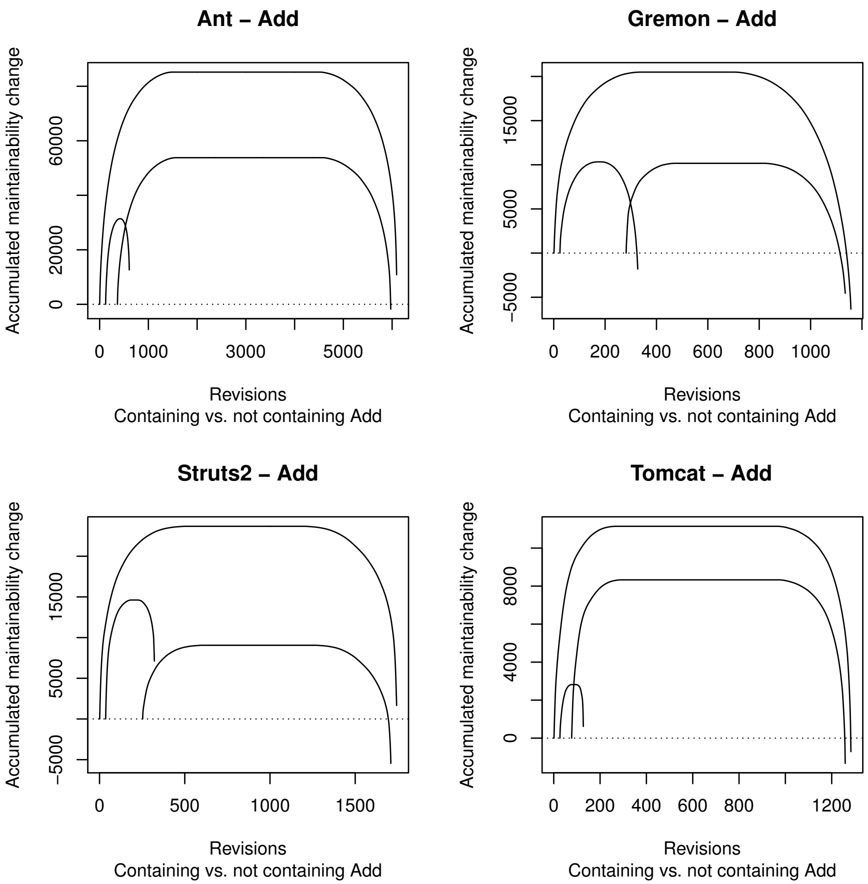


Figure 5: CCD about maintainability changes of commits with and without file additions

4.3. Variance of source code quality change caused by version control operations

One of the results presented in study [9] was that commits containing file additions have significantly higher maintainability change compared to those of not containing file additions at all. This result we might imagine as follows: this statement is true in all the magnitudes of maintainability changes. However, based on the diagrams in Figure 5 and especially in Figure 6, this is not true. In case of low values (high maintainability decreases), the values are much lower in case of commits containing file additions on the same quantile, compared to the commits not containing file additions.

The most important thing to see is it is not true – as one would expect just

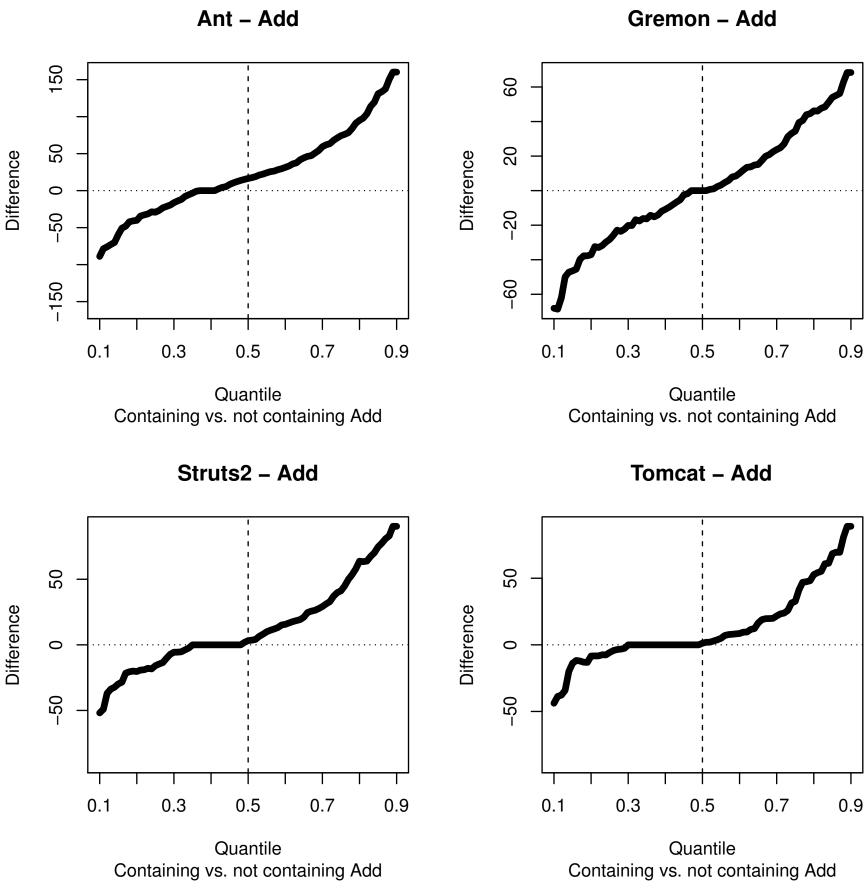


Figure 6: QDD about maintainability changes of commits with and without file additions

based on the preliminary test results – that the maintainability change values are higher in case of all magnitude of values. This means file additions really do their bits of the code erosion, and if it erodes, its erosion is much higher than the erosion caused by other commits. Without the QDD, just using the results of Wilcoxon test, this would not have been revealed.

We decided to analyze this phenomena further, and we presented our findings in paper [10]. In this study we took the similar divisions as in the earlier study, and performed the comparison of variances.

Table 3 contains the ratio of variances of maintainability change values, not containing the outliers. The p-values in these cases were so low, that the R rounded down to 0, meaning the value were less than 10^{-350} in all the cases.

Based on the CCD the difference in variance is indicated by the following: the

System	Ratio of variances
Ant	8.74
Gremon	3.13
Struts 2	10.01
Tomcat	4.18

Table 3: Ratio of variances of maintainability change values of commits containing and not containing file addition

ratio of the horizontal width (i.e. the number of observations) and the vertical width are different. This is especially apparent in Figure 5 at project Gremon: the heights of the two lower curves are similar, but the width of the left one is spectacularly lower than the width of the right one. Considering the QDD (Figure 6) it is apparent in all the four cases that the lines have significant slopes, and their main shapes are not horizontal.

Answer to RQ3

In Figure 5, the relative height of the left hand side curve is bigger than the relative height of the right hand side one, which supports the results published in article [10]. Considering the Figure 6, the fact that the tendency of the line is slant, also supports the results presented in article [10].

4.4. Cumulative code churn: impact on maintainability

In study [12], we presented how the intensity of past modifications influenced the maintainability change of the future commits.

We calculated for each file and revision from the very beginning, how many lines had been added and deleted all together. On a certain commit we averaged these values. We divided these values into two subsets, based on the maintainability change of the related commit, if it decreased or increased it; we omitted the commits related to neutral maintainability changes. Finally we compared the values using Wilcoxon test.

Table 4 contains the p-values of the Wilcoxon test results.

System	p-value
Ant	0.00235
Gremon	0.00436
Struts 2	0.00018
Tomcat	0.03616

Table 4: Cumulative code churn analysis results

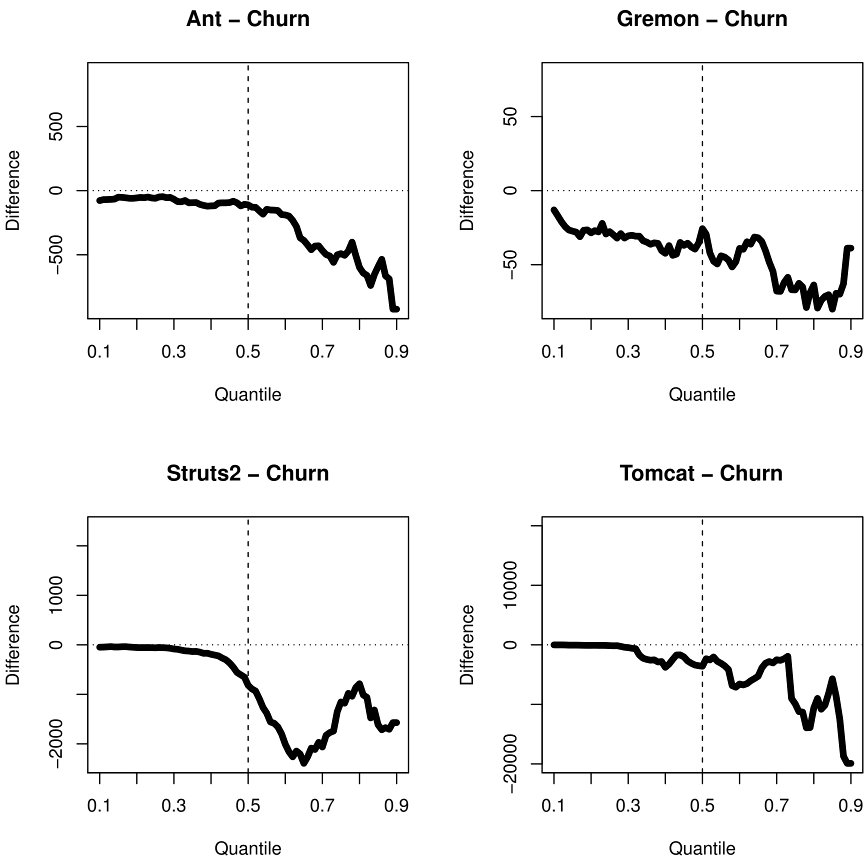


Figure 7: Quantile difference diagrams of cumulative code churns

Figure 7 illustrates the cumulative code churn comparisons using Quantile Difference Diagrams. The cumulative churn values are less, or at most equal on every quantile, for the commits related to maintainability increase, compared to those of maintainability decrease. The tendency of the line is negative: at higher values the line is located lower. At higher values – from about 0.7 – the amplitude is also higher than in case of lower values.

Answer to RQ4

We present the cumulative code churn related Quantile Difference Diagrams in Figure 7. Here the complete line is located below, or at most on the x-axis, meaning that the results presented in study [12] is valid for every quantile. The lines have a slight declining tendency, indicating the differences in the variance. We have not investigated this aspect yet, therefore this is a good candidate for a future work, revealed by the QDD.

4.5. Code ownership: impact on maintainability

In paper [11], we checked how code ownership influenced the maintainability change of the future commits. We calculated how many different developer had contributed to the file, and at certain commit we took their geometric mean. We performed similar division as presented is Section 4.4, and here also we performed the Wilcoxon test.

Table 5 contains the p-values of the statistic test results. These values are somewhat less significant than the cumulative code churn comparison.

System	p-value
Ant	0.033728
Gremon	0.059604
Struts 2	0.000014
Tomcat	0.213841

Table 5: Code ownership analysis results

We illustrate these results using Quantile Difference Diagrams in Figure 8. All the differences are non-positive, however, there are much more zeros, compared to the cumulative code churn related diagrams. Furthermore, the values on the y-axis (the differences) are lower in absolute, compared to the other diagram's (Figure 7) values. We find the greatest difference in case of Struts 2, which is not surprising according to the statistic test results.

Answer to RQ5

We present the ownership related Quantile Difference Diagrams in Figure 8. The complete lines are located below of the x-axis, or at most on it, which support the results presented in paper [11]. On the other hand, there are some spectacular differences between this and the previous diagram. Unlike in case of code churn analysis, here the line touches the x-axis several times. The difference represented by the line is not so big; it is always at most 1 (see the y-coordinate). The right hand side part of the line is more likely to be below the x-axis, compared to the left hand side, however, there is no a clear slope in the line.

5. Conclusions

In this paper we presented a case study of applying the Cumulative Characteristic Diagram and Quantile Difference Diagram on the already published results of finding various connections between version control history data and maintainability of the source code.

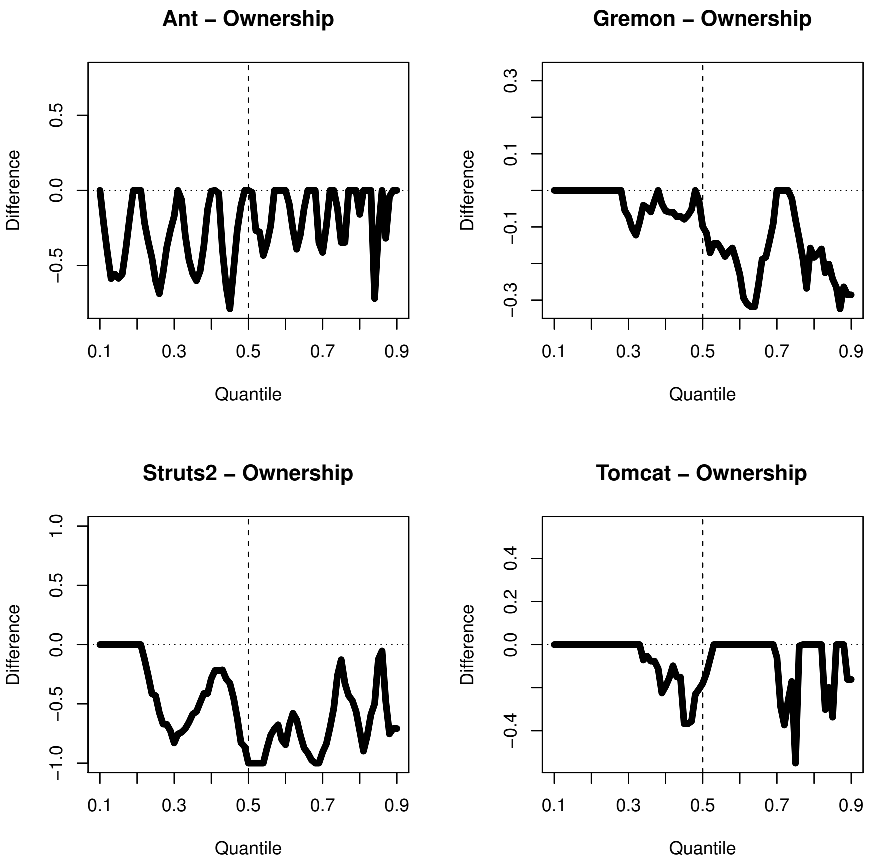


Figure 8: Quantile difference diagrams of code ownerships

First, we presented the motivating example of displaying the data using box plot diagrams. Then we presented the data with help of Cumulative Characteristic Diagrams, and performed a thorough analysis of them.

After that we considered operation Add, and illustrated the input data of the study related to the impact of version control operations on the value and the variance of the maintainability change, using both Cumulative Characteristic Diagrams and Quantile Difference Diagrams. Finally, we illustrated the data related to cumulative code churn and code ownership analysis, using Quantile Difference Diagrams.

As a final conclusion we can state that it was worthwhile to create the diagrams and apply them on the related studies. They spectacularly illustrate the statistic results, furthermore, they mostly reveal unknown connections, indicating the way for the future studies.

Acknowledgments

The author would like to thank Rudolf Ferenc and Péter Hegedűs for providing advices on this article.

References

[1] C. Faragó, "Visualization of univariate data for comparison," *Annales Mathematicae et Informaticae*, vol. 45, pp. 39–53, 2015.

[2] S. R. Chidamber and C. F. Kemerer, "A Metrics Suite for Object Oriented Design," *Transactions on Software Engineering*, vol. 20, no. 6, pp. 476–493, 1994.

[3] T. Gyimóthy, R. Ferenc, and I. Siket, "Empirical validation of object-oriented metrics on open source software for fault prediction," *Transactions on Software Engineering*, vol. 31, no. 10, pp. 897–910, 2005.

[4] T. Bakota, P. Hegedűs, P. Körtvélyesi, R. Ferenc, and T. Gyimóthy, "A probabilistic software quality model," in *Proceedings of the 27th International Conference on Software Maintenance (ICSM)*. IEEE Computer Society, 2011, pp. 243–252.

[5] ISO/IEC, *ISO/IEC 9126. Software Engineering – Product quality 6.5.* ISO/IEC, 2001.

[6] P. Hegedűs, "A probabilistic quality model for C# – an industrial case study," *Acta Cybernetica*, vol. 21, no. 1, pp. 135–147, 2013.

[7] P. Hegedűs, T. Bakota, G. Ladányi, C. Faragó, and R. Ferenc, "A drill-down approach for measuring maintainability at source code element level," *Electronic Communications of the EASST*, vol. 60, pp. 1–21, 2013.

[8] C. Faragó, P. Hegedűs, Á. Z. Végh, and R. Ferenc, "Connection between version control operations and quality change of the source code," *Acta Cybernetica*, vol. 21, no. 4, pp. 585–607, 2014.

[9] C. Faragó, P. Hegedűs, and R. Ferenc, "The impact of version control operations on the quality change of the source code," in *Proceedings of the 14th International Conference on Computational Science and Its Applications (ICCSA)*, vol. 8583 Lecture Notes in Computer Science (LNCS), no. PART 5. Springer International Publishing, 2014, pp. 353–369.

[10] C. Faragó, "Variance of source code quality change caused by version control operations," *Acta Cybernetica*, vol. 22, no. 1, pp. 35–56, 2015.

[11] C. Faragó, P. Hegedűs, and R. Ferenc, "Code ownership: Impact on maintainability," in *Proceedings of the 15th International Conference on Computational Science and Its Applications (ICCSA)*, vol. 9159 Lecture Notes in Computer Science (LNCS), no. PART 5. Springer International Publishing, 2015, pp. 3–19.

[12] ——, "Cumulative code churn: Impact on maintainability," in *Proceedings of the 15th International Working Conference on Source Code Analysis and Manipulation (SCAM)*. IEEE Computer Society, 2015, pp. 141–150.

[13] C. Faragó, P. Hegedűs, G. Ladányi, and R. Ferenc, "Impact of version history metrics on maintainability," in *Proceedings of the 8th International Conference on Advanced*

Software Engineering & Its Applications (ASEA). IEEE Computer Society, 2015, pp. 30–35.

[14] J. M. Chambers, W. S. Cleveland, B. Kleiner, and P. A. Tukey, "Graphical methods for data analysis," *Wadsworth, Belmont, CA*, 1983.

[15] P. Murrell, *R Graphics.* CRC Press, 2005.

[16] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015. [Online]. Available: http://www.R-project.org/

[17] R. McGill, J. W. Tukey, and W. A. Larsen, "Variations of box plots," *The American Statistician*, vol. 32, no. 1, pp. 12–16, 1978.

[18] Y. Benjamini, "Opening the box of a boxplot," *The American Statistician*, vol. 42, no. 4, pp. 257–262, 1988.

[19] M. Frigge, D. C. Hoaglin, and B. Iglewicz, "Some implementations of the boxplot," *The American Statistician*, vol. 43, no. 1, pp. 50–54, 1989.

[20] K. Potter, H. Hagen, A. Kerren, and P. Dannenmann, "Methods for presenting statistical information: The box plot," *Visualization of Large and Unstructured Data Sets, s*, vol. 4, pp. 97–106, 2006.

[21] D. Adler, *vioplot: Violin plot*, 2005, R package version 0.2.

[22] J. L. Hintze and R. D. Nelson, "Violin plots: a box plot-density trace synergism," *The American Statistician*, vol. 52, no. 2, pp. 181–184, 1998.

[23] P. Kampstra, "Beanplot: A boxplot alternative for visual comparison of distributions," *Journal of Statistical Software*, vol. 28, no. 1, pp. 1–9, 2008.

[24] K. M. Goldberg and B. Iglewicz, "Bivariate extensions of the boxplot," *Technometrics*, vol. 34, no. 3, pp. 307–320, 1992.

[25] P. J. Rousseeuw, I. Ruts, and J. W. Tukey, "The bagplot: a bivariate boxplot," *The American Statistician*, vol. 53, no. 4, pp. 382–387, 1999.

[26] H. P. Wolf and U. Bielefeld, *aplpack: Another Plot PACKage: stem.leaf, bagplot, faces, spin3R, plotsummary, plothulls, and some slider functions*, 2014, R package version 1.3.0.

[27] K. Hornik, A. Zeileis, and D. Meyer, "The strucplot framework: Visualizing multiway contingency tables with vcd," *Journal of Statistical Software*, vol. 17, no. 3, pp. 1–48, 2006.

[28] D. Sarkar, *Lattice: Multivariate Data Visualization with R.* Springer International Publishing, 2008.

[29] C. Faragó, *vudc: Visualization of Univariate Data for Comparison*, 2016, R package version 1.1.

Annales Mathematicae et Informaticae
46 (2016) pp. 77–96
http://ami.ektf.hu

Solutions of some particular pexiderized digital filtering functional equation

Charinthip Hengkrawit[a], Vichian Laohakosol[b*], Khanithar Naenudorn[c]

[a]Department of Mathematics and Statistics, Faculty of Science and Technology
Thammasat University, Bangkok 12120, Thailand
charinthip@mathstat.sci.tu.ac.th

[b]Department of Mathematics, Faculty of Science, Kasetsart University
Bangkok 10900, Thailand
fscivil@ku.ac.th

[c]Department of Mathematics, Faculty of Education, Dhonburi Rajabat University,
Bangkok 10600, Thailand
khanithar.n@gmail.com

Submitted February 11, 2016 — Accepted September 28, 2016

Abstract

Consider the pexiderized digital filtering functional equation

$$f_1(x+t,y+t)+f_2(x-t,y)+f_3(x,y-t)=f_4(x-t,y-t)+f_5(x+t,y)+f_6(x,y+t).$$

We determine three kinds of solutions, namely, biadditive, symmetric and skew-symmetric solution functions, subject to different sets of conditions on the functions involved.

Keywords: digital filtering functional equation, pexiderized form, biadditive function, symmetric function, skew-symmetric function.

MSC: 39B52, 39A14

*The second author is supported by the Center for Advanced Studies in Industrial Technology and the Faculty of Science, Kasetsart University, Bangkok, Thailand.

1. Introduction

Throughout let G be an abelian group which is 2-solvable, i.e., the equation $2u = v$ is solvable. A function $A\colon G \times G \to \mathbb{C}$ is said to be

- symmetric if $A(x, y) = A(y, x)$;
- skew-symmetric if $A(x, y) = -A(y, x)$;
- additive if $A(x + y) = A(x) + A(y)$:
- biadditive if it is additive in each of its variables.

Some well-known facts that we shall use implicitly are

- symmetric, skew-symmetric and biadditive properties are preserved under addition;
- if $A(x, y)$ is skew-symmetric, then $A(x, x) = 0$;
- if $A(x, y)$ is biadditive, then $A(x, 0) = 0 = A(0, y)$, $A(-x, y) = -A(x, y) = A(x, -y)$ and $A\left(\frac{x+y}{2}, \frac{x-y}{2}\right)$ is skew-symmetric.

In [4] the following functional equation related to digital filtering (see Proposition 1.2 below) is solved

$$f(x+t, y+t) + f(x-t, y) + f(x, y-t) = f(x-t, y-t) + f(x, y+t) + f(x+t, y),$$

where $f\colon G \times G \to \mathbb{C}$ and $x, y, t \in G$. Here we consider its pexiderized version, which is

$$\begin{aligned} &f_1(x+t, y+t) + f_2(x-t, y) + f_3(x, y-t) \\ &\qquad = f_4(x-t, y-t) + f_5(x+t, y) + f_6(x, y+t) \quad (x, y, t \in G), \end{aligned} \tag{PDF}$$

where $f_1, f_2, f_3, f_4, f_5, f_6\colon G \times G \to \mathbb{C}$. Since solving (PDF) generally seems quite difficult, we are content here to exhibit three kinds of solution functions of (PDF), namely, biadditive, symmetric and skew-symmetric solution functions. The case of biadditive functions is most satisfactory as complete shapes of solutions can be determined, while the remaining two cases are harder and we are forced to impose some more restrictions, which arise from certain symmetry of the functions involved. subject to three different sets of conditions on the functions involved. We will also appeal to the following results in [3], [4] and [1].

Proposition 1.1. *[3] If $f\colon G \to \mathbb{C}$ satisfies*

$$f(x+t, y+t) = f(x+t, y) + f(x, y+t) - f(x, y) \quad (x, y, t \in G),$$

then

$$f(x, y) = \phi(x) + \psi(y) + A(x, y)$$

for arbitrary mappings $\phi, \psi\colon G \to \mathbb{C}$ and arbitrary skew-symmetric biadditive map $A\colon G \times G \to \mathbb{C}$.

Proposition 1.2. *[4] The function $f\colon G\times G\to\mathbb{C}$ satisfies the functional equation*

$$f(x+t,y+t)+f(x-t,y)+f(x,y-t)=f(x-t,y-t)+f(x,y+t)+f(x+t,y)$$

for all x,y,t in G if and only if

$$f(x,y)=B(x,y)+\phi(x)+\psi(y)+\chi(x-y)$$

holds for all x,y in G, where $B\colon G\times G\to\mathbb{C}$ is biadditive and $\phi,\psi,\chi\colon G\to\mathbb{C}$ are arbitrary functions.

Proposition 1.3. *[1] If $f_1,f_2,f_3,f_4\colon G\times G\to\mathbb{C}$ satisfy the functional equation*

$$f_1(x+t,y+s)+f_2(x-t,y-s)=f_3(x+s,y-t)+f_4(x-s,y+t)\quad (x,y,s,t\in G),$$

then f_1,f_2,f_3 and f_4 are given by

$$f_1=w+h,\ f_2=w-h,\ f_3=w+k,\ f_4=w-k$$

where $w\colon G\times G\to\mathbb{C}$ is an arbitrary solution of the functional equation

$$w(x+t,y+s)+w(x-t,y-s)=w(x+s,y-t)+w(x-s,y+t)\quad (x,y,s,t\in G)$$

and $h,k\colon G\times G\to\mathbb{C}$ are arbitrary solutions of the system of difference functional equations

$$\Delta^3_{y,t}g(x,y)=0,\quad \Delta^3_{x,t}g(x,y)=0\quad (x,y,t\in G),$$

where the two partial difference operators are defined by

$$\Delta_{x,t}g(x,y)=g(x+t,y)-g(x,y),\ \Delta_{y,t}g(x,y)=g(x,y+t)-g(x,y).$$

2. Auxiliary results

It is convenient to introduce translation operators X^t and Y^t for $t\in G$, which are defined by

$$X^tf(x,y)=f(x+t,y),\quad Y^tf(x,y)=f(x,y+t).$$

In particular, $X^0=Y^0=1$ denote the identity operator.

Lemma 2.1. *Let $f_1,f_2,f_3,f_4,f_5,f_6\colon G\times G\to\mathbb{C}$, and for $i,j\in\{1,\dots,6\}$ with $i\neq j$, put*

$$S^{(i,j)}(x,y)=\frac{1}{2}\left\{f_i(x,y)+f_j(x,y)\right\},\ D^{(i,j)}(x,y)=\frac{1}{2}\left\{f_i(x,y)-f_j(x,y)\right\}.$$

Assume that f_1,f_2,f_3,f_4,f_5,f_6 satisfy (PDF).
A) Then

$$X^tY^tf_1+X^{-t}f_2+Y^{-t}f_3=X^{-t}Y^{-t}f_4+X^tf_5+Y^tf_6 \quad (2.1)$$

$$\left(X^tY^t - X^{-t}Y^{-t}\right) S^{(1,4)} + \left(X^{-t} - X^t\right) S^{(2,5)} + \left(Y^{-t} - Y^t\right) S^{(3,6)} = 0 \quad (2.2)$$

$$\left(X^tY^t + X^{-t}Y^{-t}\right) D^{(1,4)} + \left(X^{-t} + X^t\right) D^{(2,5)} + \left(Y^{-t} + Y^t\right) D^{(3,6)} = 0. \quad (2.3)$$

B) If, in addition, $f_1, f_2, f_3, f_4, f_5, f_6$ are symmetric or skew-symmetric functions, then

$$(X^{-t} - Y^{-t})D^{(2,3)} = \left(X^t - Y^t\right) D^{(5,6)} \quad (2.4)$$

$$X^tY^t(2f_1) + (X^{-t} + Y^{-t})S^{(2,3)} = X^{-t}Y^{-t}(2f_4) + (X^t + Y^t)S^{(5,6)} \quad (2.5)$$

$$X^tY^tS^{(1,4)} + X^{-t}S^{(2,6)} + Y^{-t}S^{(3,5)} = X^{-t}Y^{-t}S^{(1,4)} + X^tS^{(3,5)} + Y^tS^{(2,6)}. \quad (2.6)$$

Proof. A) Writing (PDF) in the operator form, we get

$$\begin{aligned} &X^tY^tf_1(x,y) + X^{-t}f_2(x,y) + Y^{-t}f_3(x,y) \\ &= X^{-t}Y^{-t}f_4(x,y) + X^tf_5(x,y) + Y^tf_6(x,y), \end{aligned}$$

which is (2.1). Replacing t by $-t$, we get

$$\begin{aligned} &X^{-t}Y^{-t}f_1(x,y) + X^tf_2(x,y) + Y^tf_3(x,y) \\ &= X^tY^tf_4(x,y) + X^{-t}f_5(x,y) + Y^{-t}f_6(x,y). \end{aligned}$$

The relations (2.2) and (2.3) follow from subtracting, respectively adding, the last two equations and rearranging terms.
B) Using the fact that $f_1, f_2, f_3, f_4, f_5, f_6$ are symmetric or skew-symmetric, (2.1) becomes

$$X^tY^tf_1 + Y^{-t}f_2 + X^{-t}f_3 = X^{-t}Y^{-t}f_4 + Y^tf_5 + X^tf_6. \quad (2.7)$$

Replacing t by $-t$, we get

$$X^tY^tf_4 + Y^{-t}f_5 + X^{-t}f_6 = X^{-t}Y^{-t}f_1 + Y^tf_2 + X^tf_3. \quad (2.8)$$

The relations (2.4) and (2.5) follow from subtracting, respectively, adding (2.7) and (2.1). The relation (2.6) comes from adding (2.8) and (2.1). □

3. Bi-additive solutions

In this section, biadditive solutions of (PDF) are completely determined.

Theorem 3.1. *If $f_1, f_2, f_3, f_4, f_5, f_6$ are biadditive functions satisfying* (PDF), *then*

$$\begin{aligned} f_1(x,y) &= B(x,y) - C(x,y) - D(x,y), \\ f_2(x,y) &= B(x,y) + C(x,y), \\ f_3(x,y) &= B(x,y) + D(x,y), \end{aligned}$$

$$f_4(x,y) = B(x,y) + C(x,y) + D(x,y),$$
$$f_5(x,y) = B(x,y) - C(x,y),$$
$$f_6(x,y) = B(x,y) - D(x,y),$$

where $B, C, D\colon G \times G \to \mathbb{C}$ *are arbitrary biadditive functions satisfying*

$$C(t,t) + D(t,t) = 0 \quad (t \in G).$$

Proof. Rewriting (2.2), we get

$$S^{(1,4)}(x+t, y+t) - S^{(1,4)}(x-t, y-t) + S^{(2,5)}(x-t, y) - S^{(2,5)}(x+t, y) + S^{(3,6)}(x, y-t) - S^{(3,6)}(x, y+t) = 0.$$

Since $S^{(1,4)}, S^{(2,5)}, S^{(3,6)}$ are biadditive, simplifying we get

$$S^{(1,4)}(t,y) - S^{(2,5)}(t,y) = -S^{(1,4)}(x,t) + S^{(3,6)}(x,t). \tag{3.1}$$

Replacing t by x, we have

$$S^{(1,4)}(x,y) - S^{(2,5)}(x,y) = -S^{(1,4)}(x,x) + S^{(3,6)}(x,x) =: \mathcal{S}_1(x).$$

Substituting $y = 0$, we get $\mathcal{S}_1(x) = 0$, and so

$$S^{(1,4)}(x,y) = S^{(2,5)}(x,y). \tag{3.2}$$

Similarly, replacing t by y in (3.1) and substituting $x = 0$, we get

$$S^{(1,4)}(x,y) = S^{(3,6)}(x,y). \tag{3.3}$$

From (3.2) and (3.3), we set

$$S^{(2,5)}(x,y) = S^{(1,4)}(x,y) = S^{(3,6)}(x,y) =: B(x,y). \tag{3.4}$$

On the other hand, rewriting (2.3), we get

$$D^{(1,4)}(x+t, y+t) + D^{(1,4)}(x-t, y-t) + D^{(2,5)}(x-t, y) + D^{(2,5)}(x+t, y) + D^{(3,6)}(x, y-t) + D^{(3,6)}(x, y+t) = 0.$$

Since $D^{(1,4)}, D^{(2,5)}, D^{(3,6)}$ are biadditive, simplifying we get

$$D^{(1,4)}(x,y) + D^{(1,4)}(t,t) + D^{(2,5)}(x,y) + D^{(3,6)}(x,y) = 0.$$

Setting $y = 0$, we have

$$D^{(1,4)}(t,t) = 0, \tag{3.5}$$

and so

$$D^{(1,4)}(x,y) + D^{(2,5)}(x,y) + D^{(3,6)}(x,y) = 0.$$

Putting

$$D^{(2,5)}(x,y) =: C(x,y) \tag{3.6}$$

$$D^{(3,6)}(x,y) =: D(x,y) \tag{3.7}$$

we arrive at

$$D^{(1,4)}(x,y) = -C(x,y) - D(x,y). \tag{3.8}$$

From (3.5), we have $C(t,t) + D(t,t) = 0$. The desired result follows from solving (3.4),(3.6),(3.7) and (3.8) for $f_1, \ldots, f_6$. □

4. Symmetric solutions

In this section, we consider solutions of (PDF) which are symmetric functions. This case is much more complicated than the previous one and so we start with several lemmas.

Lemma 4.1. *Let $B\colon G \times G \to \mathbb{C}$ be biadditive and let $\phi, \psi, \chi\colon G \to \mathbb{C}$ be arbitrary functions. If*

$$f(x,y) = B(x,y) + \phi(x) + \psi(y) + \chi(x-y) \tag{4.1}$$

is symmetric, then

$$f(x,y) = B(x,y) + \{\phi + \psi(x) + \psi(y)\} + \{\chi(-x) - \chi(x) + \chi(x-y)\},$$

where the biadditive function $B(x,y)$ is symmetric, ϕ is a complex constant, and $\chi(-x) - \chi(x)$ is an additive function of x.

Proof. Since $f(x,y)$ is symmetric, equating $f(x,y) = f(y,x)$, we get

$$B(x,y) + \phi(x) + \psi(y) + \chi(x-y) = B(y,x) + \phi(y) + \psi(x) + \chi(y-x). \tag{4.2}$$

Substituting $y = 0$, using $B(x,0) = 0 = B(0,x)$, putting $\phi = \phi(0) - \psi(0)$ and simplifying, we have

$$\phi(x) = \phi + \psi(x) + \chi(-x) - \chi(x).$$

Replacing this $\phi(x)$ in (4.2) and simplifying we get

$$B(x,y) + \chi(-x) - \chi(x) + \chi(x-y) = B(y,x) + \chi(-y) - \chi(y) + \chi(y-x). \tag{4.3}$$

Substituting $y = x - z$ and using biadditivity, we get

$$B(z,x) + \chi(-x) - \chi(x) + \chi(z) = B(x,z) + \chi(z-x) - \chi(x-z) + \chi(-z).$$

Replacing z by y, we get

$$B(y,x) + \chi(-x) - \chi(x) + \chi(y) = B(x,y) + \chi(y-x) - \chi(x-y) + \chi(-y). \tag{4.4}$$

Subtracting (4.4) from (4.3), we deduce that $B(x,y) = B(y,x)$, i.e., B is symmetric. Adding (4.3) and (4.4), and rearranging we deduce that

$$\{\chi(-x) - \chi(x)\} + \{\chi(x-y) - \chi(y-x)\} = \chi(-y) - \chi(y),$$

i.e., $\chi(-x) - \chi(x)$ is an additive function of x. Incorporating all the information obtained, the result follows. $\square$

Lemma 4.2. *Let the notation be as in Lemma 2.1. If $f_1, f_2, f_3, f_4, f_5, f_6$ are symmetric (or skew-symmetric) functions satisfying* (PDF), *then*

$$-\frac{1}{2}\{f_2(x,y) - f_3(x,y)\} =: -D^{(2,3)}(x,y) = w(x,y) + k(x,y) \tag{4.5}$$

$$\frac{1}{2}\{f_5(x,y) - f_6(x,y)\} =: D^{(5,6)}(x,y) = w(x,y) - k(x,y). \tag{4.6}$$

where the functions w and k are as described in Proposition 1.3.

Proof. Rewriting (2.4), we get

$$D^{(5,6)}(x+t,y) - D^{(2,3)}(x-t,y) = D^{(5,6)}(x,y+t) - D^{(2,3)}(x,y-t).$$

Using Proposition 1.3 with $s = 0$, $f_1 = D^{(5,6)} = f_4$, $f_2 = -D^{(2,3)} = f_3$, we have

$$w + h = f_1 = D^{(5,6)} = f_4 = w - k, \; w - h = f_2 = -D^{(2,3)} = f_3 = w + k,$$

i.e., $h = -k$, and the result follows. $\square$

Lemma 4.3. *Let $f_1, f_2, f_3, f_4, f_5, f_6 \colon G \times G \to \mathbb{C}$, and let*

$$K(x,y) := f_2 + f_5 - f_3 - f_6, \quad H(x,y) := f_2 - f_5 - f_3 + f_6.$$

If $f_1, f_2, f_3, f_4, f_5, f_6$ are symmetric functions satisfying (PDF), *then*

$$K(x,y) = \alpha_K(x+y) + \beta \tag{4.7}$$

$$H(x,y) = \alpha_H(x+y) + \frac{1}{2}\{\beta_H(x-y) + \beta_H(y-x)\}, \tag{4.8}$$

where $\alpha_K, \alpha_H, \beta_H \colon G \to \mathbb{C}$ are arbitrary functions, and β is a complex constant.

Proof. Using symmetry in (2.2), we get

$$\left(X^tY^t - X^{-t}Y^{-t}\right)S^{(1,4)} + \left(Y^{-t} - Y^t\right)S^{(2,5)} + \left(X^{-t} - X^t\right)S^{(3,6)} = 0. \tag{4.9}$$

Subtracting (4.9) from (2.2) and rearranging, we get

$$\left(X^{-t} - X^t\right)K(x,y) = \left(Y^{-t} - Y^t\right)K(x,y). \tag{4.10}$$

Operating both sides of (4.10) by $X^{-t} - X^t$ and using (4.10) again, we get

$$\left(X^{-2t} - 2 + X^{2t}\right)K(x,y) = \left(Y^{-t} - Y^t\right)\left(X^{-t} - X^t\right)K(x,y)$$

$$= \left(Y^{-2t} - 2 + Y^{2t}\right) K(x,y).$$

Simplifying and replacing $2t$ by t, we have

$$\left(X^{-t} + X^{t}\right) K(x,y) = \left(Y^{-t} + Y^{t}\right) K(x,y). \tag{4.11}$$

Defining the function $K_1 \colon G \times G \to \mathbb{C}$ by

$$K_1(x,y) := K\left(\frac{x+y}{2}, \frac{x-y}{2}\right), \text{ or equivalently, } \quad K(x,y) = K_1(x+y, x-y)$$

and rewriting (4.11) in terms of K_1, we get

$$\begin{aligned}&K_1(x+y-t, x-y-t) + K_1(x+y+t, x-y+t)\\ &= K_1(x+y-t, x-y+t) + K_1(x+y+t, x-y-t).\end{aligned}$$

Putting $u = x+y-t,\ v = x-y-t$, this last equation becomes

$$K_1(u,v) + K_1(u+2t, v+2t) = K_1(u, v+2t) + K_1(u+2t, v).$$

Replacing $2t$ by t and rearranging, we get

$$K_1(u+t, v+t) = K_1(u+t, v) + K_1(u, v+t) - K_1(u,v),$$

which is the McKiernan's functional equation mentioned in Proposition 1.1, whose solution is of the form

$$K_1(x,y) = \alpha_K(x) + \beta_K(y) + A_K(x,y),$$

with the functions α_K, β_K, A_K as stated above. Reverting back to K, we get

$$K(x,y) = \alpha_K(x+y) + \beta_K(x-y) + A_K(x+y, x-y).$$

Since A_K is skew-symmetric and biadditive, the shape of K reduces to

$$K(x,y) = \alpha_K(x+y) + \beta_K(x-y) - 2A_K(x,y). \tag{4.12}$$

Since K is symmetric, equating $K(x,y) = K(y,x)$, we get

$$\beta_K(x-y) - \beta_K(y-x) = 4A_K(x,y). \tag{4.13}$$

Substituting $y = 0$, we get $\beta_K(x) - \beta_K(-x) = 0$, i.e., β_K is an even function. From (4.10) and (4.11), we see at once that $X^t K(x,y) = Y^t K(x,y)$. Using this and (4.12), we get

$$\begin{aligned}&\alpha_K(x+y+t) + \beta_K(x-y+t) - 2A_K(x+t, y)\\ &= \alpha_K(x+y+t) + \beta_K(x-y-t) - 2A_K(x, y+t).\end{aligned}$$

Since A_K is biadditive and skew-symmetric, this last relation simplifies to

$$\beta_K(x-y+t)-\beta_K(x-y-t)=2A_K(t,x+y).$$

Putting $x=y$, and using the fact that β_K is an even function, we have $A_K(x,y)=0$, and so $\beta_K(x-y+t)-\beta_K(x-y-t)=0$. Taking $x-y-t=0$, we deduce that $\beta_K(x,y)=\beta$, a constant. The shape of K follows by collecting all the information found.

The determination of H proceeds analogously. Using symmetry in (2.3), we get

$$\left(X^tY^t+X^{-t}Y^{-t}\right)D^{(1,4)}+\left(Y^{-t}+Y^t\right)D^{(2,5)}+\left(X^{-t}+X^t\right)D^{(3,6)}=0. \tag{4.14}$$

Subtracting (4.14) from (2.3) and rearranging, we get

$$\left(X^{-t}+X^t\right)H(x,y)=\left(Y^{-t}+Y^t\right)H(x,y). \tag{4.15}$$

Defining the function $H_1\colon G\times G\to\mathbb{C}$ by

$$H_1(x,y):=H\left(\frac{x+y}{2},\frac{x-y}{2}\right),\ \text{ or equivalently,}\quad H(x,y)=H_1(x+y,x-y)$$

and rewriting (4.15) in terms of H_1, we get

$$\begin{aligned}&H_1(x+y-t,x-y-t)+H_1(x+y+t,x-y+t)\\&=H_1(x+y-t,x-y+t)+H_1(x+y+t,x-y-t).\end{aligned}$$

Putting $u=x+y-t,\ v=x-y-t$, then replacing $2t$ by t and rearranging, we get

$$H_1(u+t,v+t)=H_1(u+t,v)+H_1(u,v+t)-H_1(u,v),$$

which is the McKiernan's functional equation mentioned in Proposition 1.1, whose solution is of the form

$$H_1(x,y)=\alpha_H(x)+\beta_H(y)+A_H(x,y),$$

with the functions α_H,β_H,A_H as stated above. Reverting back to H, we get

$$H(x,y)=\alpha_H(x+y)+\beta_H(x-y)+A_H(x+y,x-y).$$

As in the previous case, using the fact that A_H is skew-symmetric and biadditive, the shape of H reduces to

$$H(x,y)=\alpha_H(x+y)+\beta_H(x-y)-2A_H(x,y).$$

Since H is symmetric, equating $H(x,y)=H(y,x)$, we get $\beta_H(x-y)-\beta_H(y-x)=4A_H(x,y)$. Incorporating all these details, the shape of H follows. $\square$

Remark 4.1. R1) Setting $t = 0$ in (4.14), and simplifying, we get

$$(f_1 - f_4) + (f_2 - f_5) + (f_3 - f_6) = 0.$$

R2) From Lemma 4.3 and Lemma 4.2, we have

$$\begin{aligned} &\alpha_K(x+y) + \beta \\ &= K = f_2 + f_5 - f_3 - f_6 = -4k(x,y) && (4.16) \\ &\alpha_H(x+y) + \frac{1}{2}\left\{\beta_H(x-y) + \beta_H(y-x)\right\} \\ &= H = f_2 - f_5 - f_3 + f_6 = -4w(x,y). && (4.17) \end{aligned}$$

Throughout the rest of this section, we shall deal only with the case where $f_1 = f_4$, which, by remark R1), gives rise to

$$f_2 - f_5 + f_3 - f_6 = 0.$$

Combining with (4.16) and (4.17), we get

$$f_2 - f_6 = -2k, \ \ f_3 - f_6 = 2w \tag{4.18}$$

Theorem 4.4. *Assume* $f_1, f_2, f_3, f_4, f_5, f_6 \colon G \times G \to \mathbb{C}$ *are symmetric functions satisfying* (PDF).
I. If

$$f_1 = f_4 = \frac{1}{2}\left(f_2 + f_3\right),$$

then there are biadditive, symmetric function $B(x,y)\colon G \times G \to \mathbb{C}$*, two constants* ϕ, β_1*, arbitrary functions* $\psi, \alpha_1, \alpha_2, \beta_2$ *and* $\chi\colon G \to \mathbb{C}$ *with* $\chi(x) - \chi(-x)$ *being an additive function in* x *such that*

$$\begin{aligned} f_1(x,y) &= f_4(x,y) = B(x,y) + \{\psi(x) + \psi(y) + \phi\} + \{\chi(-x) - \chi(x) + \chi(x-y)\} \\ f_2(x,y) &= f_1(x,y) + \{\alpha_1(x+y) + \beta_1\} + \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\} \\ f_3(x,y) &= f_1(x,y) - \{\alpha_1(x+y) + \beta_1\} - \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\} \\ f_5(x,y) &= f_1(x,y) + \{\alpha_1(x+y) + \beta_1\} - \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\} \\ f_6(x,y) &= f_1(x,y) - \{\alpha_1(x+y) + \beta_1\} + \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\}. \end{aligned}$$

II. If

$$f_1 = f_4 = \frac{1}{2}\left(f_2 + f_6\right) = \frac{1}{2}\left(f_3 + f_5\right),$$

then there are biadditive, symmetric function $B(x,y)\colon G \times G \to \mathbb{C}$*, two constants* ϕ, β_1*, arbitrary functions* ψ, α_1 *and* $\chi\colon G \to \mathbb{C}$ *with* $\chi(x) - \chi(-x)$ *being an additive function in* x *such that*

$$\begin{aligned} f_1(x,y) &= f_4(x,y) = B(x,y) + \{\psi(x) + \psi(y) + \phi\} + \{\chi(-x) - \chi(x) + \chi(x-y)\} \\ f_2(x,y) &= f_5(x,y) = f_1(x,y) + \{\alpha_1(x+y) + \beta_1\} \end{aligned}$$

$$f_3(x,y) = f_6(x,y) = f_1(x,y) - \{\alpha_1(x+y) + \beta_1\}.$$

III. If

$$f_1 = f_4 = \frac{1}{2}(f_2 + f_5) = \frac{1}{2}(f_3 + f_6),$$

then there are biadditive, symmetric function $B(x,y)\colon G \times G \to \mathbb{C}$, a constant ϕ, arbitrary functions ψ, α_2, β_2 and $\chi\colon G \to \mathbb{C}$ with $\chi(x) - \chi(-x)$ being an additive function in x such that

$$\begin{aligned}
f_1(x,y) &= f_4(x,y) = B(x,y) + \{\psi(x) + \psi(y) + \phi\} + \{\chi(-x) - \chi(x) + \chi(x-y)\}\\
f_2(x,y) &= f_6(x,y) = f_1(x,y) + \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\}\\
f_3(x,y) &= f_5(x,y) = f_1(x,y) - \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\}.
\end{aligned}$$

IV. If

$$f_1 = f_4 = f_6 \quad \text{and} \quad K(x,x) - H(x,x) = c$$

where c is a constant, then there are biadditive, symmetric function $B(x,y)\colon G \times G \to \mathbb{C}$, two constants ϕ, β_1, arbitrary functions $\psi, \alpha_1, \alpha_2, \beta_2$, $\chi\colon G \to \mathbb{C}$ with $\alpha_1(x) - \alpha_2(x)$ being a constant and $\chi(x) - \chi(-x)$ being an additive function in x such that

$$\begin{aligned}
f_1(x,y) &= f_4(x,y) = f_6(x,y) = B(x,y) + \{\psi(x) + \psi(y) + \phi\}\\
&\qquad\qquad\qquad\qquad\qquad\qquad + \{\chi(-x) - \chi(x) + \chi(x-y)\}\\
f_2(x,y) &= f_1(x,y) + \{\alpha_1(x+y) + \beta_1\}\\
f_3(x,y) &= f_1(x,y) - \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\}\\
f_5(x,y) &= f_1(x,y) + \{\alpha_1(x+y) + \beta_1\} - \{\alpha_2(x+y) + \beta_2(x-y) + \beta_2(y-x)\}.
\end{aligned}$$

Proof. I. Using $f_1 = f_4$, $f_2 + f_3 = f_5 + f_6$, substituting $g = 2f_1 = 2f_4 = f_2 + f_3 = f_5 + f_6$ in (2.5) and simplifying, we get

$$\begin{aligned}
&g(x+t, y+t) + g(x-t, y) + g(x, y-t)\\
&= g(x-t, y-t) + g(x, y+t) + g(x+t, y),
\end{aligned}$$

which is the Sahoo-Székelyhidi's functional equation mentioned in Proposition 1.2, and its solution is of the form

$$g(x,y) = B_1(x,y) + \phi_1(x) + \psi_1(y) + \chi_1(x-y),$$

where $B_1\colon G \times G \to \mathbb{C}$ is biadditive and $\phi_1, \psi_1, \chi_1\colon G \to \mathbb{C}$ are arbitrary functions. Since g is symmetric, applying Lemma 4.1, we deduce that

$$g(x,y) = B_1(x,y) + \{\phi_1 + \psi_1(x) + \psi_1(y)\} + \{\chi_1(-x) - \chi_1(x) + \chi_1(x-y)\}, \tag{4.19}$$

where ϕ_1 is a constant and $\chi_1(-x) - \chi_1(x)$ is an additive function of x, and by putting

$$B = \frac{B_1}{2},\ \phi = \frac{\phi_1}{2},\ \psi = \frac{\psi_1}{2},\ \chi = \frac{\chi_1}{2},$$

this gives the shapes of f_1 and f_4. Using (4.5), (4.16) and (4.17), we get

$$\begin{aligned}2f_3 &= (f_2+f_3)+(-f_2+f_3) = g+2(w+k)\\ &= g(x,y) - \frac{1}{4}\left\{\alpha_H(x+y)+\frac{1}{4}\left(\beta_H(x-y)+\beta_H(y-x)\right)\right\}\\ &\qquad -\frac{1}{4}\left\{\alpha_K(x+y)+\beta\right\}\end{aligned}$$

and the shape of f_3 follows by putting $\alpha_1 = \frac{1}{4}\alpha_K,\ \beta_1 = \frac{1}{4}\beta,\ \alpha_2 = \frac{1}{4}\alpha_H,\ \beta_2 = \frac{1}{8}\beta_H$. The shapes of other solution functions follow similarly noting from above and (4.6) that

$$f_2 = g - f_3,\ f_5+f_6 = g,\ f_5-f_6 = 2(w-k),$$

and then using (4.16) and (4.17).

II. The proof proceeds much the same as that of part I. Using $f_1 = f_4,\ f_2 + f_3 = f_5+f_6$, substituting $g = 2f_1 = 2f_4 = f_2+f_6 = f_3+f_5$ in (2.6) and simplifying, we see that g satisfies the the Sahoo-Székelyhidi's functional equation. Using symmetry, we deduce that it must be of the form (4.19). Thus,

$$\begin{aligned}&2f_1 = 2f_4 = f_2+f_6 = f_3+f_5 = g\\ &= B(x,y)+\{\phi+\psi(x)+\psi(y)\}+\{\chi(-x)-\chi(x)+\chi(x-y)\}.\end{aligned}$$

The shapes of the solution functions follow by using (4.18), (4.16) and (4.17).

III. Using $f_1 = f_4,\ f_2+f_3 = f_5+f_6$, substituting $g = 2f_1 = 2f_4 = f_2+f_5 = f_3+f_6$ in (2.6) and simplifying, we see that g satisfies the Sahoo-Székelyhidi's functional equation. Using symmetry, we deduce that its solution must be of the form (4.19). Thus,

$$\begin{aligned}&2f_1 = 2f_4 = f_2+f_5 = f_3+f_6 = g\\ &= B(x,y)+\{\phi+\psi(x)+\psi(y)\}+\{\chi(-x)-\chi(x)+\chi(x-y)\}.\end{aligned}$$

The shapes of the solution functions follow by using (4.18), (4.16) and (4.17).

IV. Solving for f_2, f_3, f_5 in terms of f_6 in (4.18) and (4.16), we get

$$f_2 = f_6 - 2k,\ f_3 = f_6+2w,\ f_5 = f_6+2(w-k).$$

Substituting these functions in (PDF), using (4.16) and (4.17), we get

$$\begin{aligned}&f_1(x+t,y+t)+f_6(x-t,y)+\frac{1}{2}\alpha_K(x+y-t)+f_6(x,y-t)-\frac{1}{2}\alpha_H(x+y-t)\\ &= f_4(x-t,y-t)+f_6(x+t,y)-\frac{1}{2}\alpha_H(x+y+t)\\ &\quad+\frac{1}{2}\alpha_K(x+y+t)+f_6(x,y+t).\end{aligned}\tag{4.20}$$

From Lemma 4.3, the condition $K(x,x)-H(x,x) = c$ leads to

$$\alpha_K(2x)+\beta-(\alpha_H(2x)+\beta_H(0)) = c,$$

i.e., the function $\alpha_K - \alpha_H$ is constant. Using this information and the hypotheses $f_1 = f_4 = f_6$, the equation (4.20) becomes the Sahoo-Székelyhidi's functional equation. Using symmetry, we deduce that its solution must be of the form (4.19). Thus,

$$f_1 = f_4 = f_6 = B(x,y) + \{\phi + \psi(x) + \psi(y)\} + \{\chi(-x) - \chi(x) + \chi(x-y)\}\,.$$

The shapes of the solution functions follow by using (4.18), (4.16) and (4.17). $\square$

5. Skew-symmetric solutions

In this section, we consider solutions of (PDF) which are skew-symmetric functions.

Lemma 5.1. *Let $B\colon G\times G \to \mathbb{C}$ be biadditive and let $\phi,\psi,\chi\colon G\to\mathbb{C}$ be arbitrary functions. If*

$$f(x,y) = B(x,y) + \phi(x) + \psi(y) + \chi(x-y) \tag{5.1}$$

is skew-symmetric, then

$$f(x,y) = B(x,y) - B(x,x) - \psi(x) + \psi(y) + \chi(x-y) - \Phi,$$

where $\Phi = \chi(0)$ is a constant and

$$\chi(x) + \chi(-x) = 2\Phi + B(x,x).$$

Proof. Since $f(x,y)$ is skew-symmetric, equating $f(x,y) = -f(y,x)$, we get

$$B(x,y) + \phi(x) + \psi(y) + \chi(x-y) = -B(y,x) - \phi(y) - \psi(x) - \chi(y-x). \tag{5.2}$$

Substituting $y = 0$, using $B(x,0) = 0 = B(0,x)$, putting $\Phi = -\phi(0) - \psi(0)$ and simplifying, we have

$$\phi(x) = \Phi - \psi(x) - \chi(-x) - \chi(x). \tag{5.3}$$

Replacing this $\phi(x)$ in (5.2) and simplifying we get

$$B(x,y) - \chi(-x) - \chi(x) + \chi(x-y) + 2\Phi = -B(y,x) + \chi(-y) + \chi(y) - \chi(y-x). \tag{5.4}$$

Taking $x = 0$, and simplifying, we get

$$\Phi = \chi(0).$$

Substituting $y = x - z$ and using biadditivity, we get

$$2B(x,x) - B(x,z) - \chi(-x) - \chi(x) + \chi(z) + 2\Phi = B(z,x) + \chi(z-x) + \chi(x-z) - \chi(-z).$$

Replacing z by y, we get

$$2B(x,x) - B(x,y) - \chi(-x) - \chi(x) + \chi(y) + 2\Phi$$

$$= B(y,x) + \chi(y-x) + \chi(x-y) - \chi(-y). \tag{5.5}$$

Combining (5.5) with (5.4) and simplifying, we deduce that

$$\chi(x) + \chi(-x) = 2\Phi + B(x,x). \tag{5.6}$$

From (5.6) and (5.3), we get

$$\phi(x) = -\Phi - \psi(x) - B(x,x). \tag{5.7}$$

Incorporating all the information obtained, the result follows. □

Lemma 5.2. *Let $f_1, f_2, f_3, f_4, f_5, f_6 \colon G \times G \to \mathbb{C}$, and let*

$$K(x,y) := f_2 + f_5 - f_3 - f_6, \quad H(x,y) := f_2 - f_5 - f_3 + f_6.$$

If $f_1, f_2, f_3, f_4, f_5, f_6$ are skew-symmetric functions satisfying (PDF)*, then*

$$K(x,y) = 0, \quad H(x,y) = -\beta_H(0) + \beta_H(x-y) - 2A_H(x,y), \tag{5.8}$$

where $\alpha_K, \beta_H \colon G \to \mathbb{C}$ are arbitrary functions, β a complex constant, A_H a biadditive, skew-symmetric function, and $\beta_H(t) + \beta_H(-t) = 2\beta_H(0)$.

Proof. Using skew-symmetry in (2.2), we get

$$\left(X^tY^t - X^{-t}Y^{-t}\right)S^{(1,4)} + \left(Y^{-t} - Y^t\right)S^{(2,5)} + \left(X^{-t} - X^t\right)S^{(3,6)} = 0. \tag{5.9}$$

Subtracting (5.9) from (2.2) and rearranging, we get

$$\left(X^{-t} - X^t\right)K(x,y) = \left(Y^{-t} - Y^t\right)K(x,y). \tag{5.10}$$

Operating both sides of (5.10) by $X^{-t} - X^t$ and using (5.10) again, we get

$$\begin{aligned}\left(X^{-2t} - 2 + X^{2t}\right)K(x,y) &= \left(Y^{-t} - Y^t\right)\left(X^{-t} - X^t\right)K(x,y)\\ &= \left(Y^{-2t} - 2 + Y^{2t}\right)K(x,y).\end{aligned}$$

Simplifying and replacing $2t$ by t, we have

$$\left(X^{-t} + X^t\right)K(x,y) = \left(Y^{-t} + Y^t\right)K(x,y). \tag{5.11}$$

Defining the function $K_1 \colon G \times G \to \mathbb{C}$ by

$$K_1(x,y) := K\left(\frac{x+y}{2}, \frac{x-y}{2}\right), \text{ or equivalently, } \quad K(x,y) = K_1(x+y, x-y)$$

and rewriting (5.11) in terms of K_1, we get

$$\begin{aligned}&K_1(x+y-t, x-y-t) + K_1(x+y+t, x-y+t)\\ &= K_1(x+y-t, x-y+t) + K_1(x+y+t, x-y-t).\end{aligned}$$

Putting $u = x + y - t,\ v = x - y - t$, this last equation becomes

$$K_1(u, v) + K_1(u + 2t, v + 2t) = K_1(u, v + 2t) + K_1(u + 2t, v).$$

Replacing $2t$ by t and rearranging, we get

$$K_1(u + t, v + t) = K_1(u + t, v) + K_1(u, v + t) - K_1(u, v),$$

which is the McKiernan's functional equation mentioned in Proposition 1.1, whose solution is of the form

$$K_1(x, y) = \alpha_K(x) + \beta_K(y) + A_K(x, y).$$

Reverting back to K, we get

$$K(x, y) = \alpha_K(x + y) + \beta_K(x - y) + A_K(x + y, x - y).$$

Since A_K is skew-symmetric and biadditive, the shape of K reduces to

$$K(x, y) = \alpha_K(x + y) + \beta_K(x - y) - 2A_K(x, y). \tag{5.12}$$

Since K is skew-symmetric, solving $K(x, y) = -K(y, x)$, we get

$$2\alpha_K(x + y) + \beta_K(x - y) + \beta_K(y - x) = 0. \tag{5.13}$$

Substituting $x = y$, we get $\alpha_K(x) = -\beta_K(0)$, a constant, and so (5.13) yields

$$\beta_K(t) + \beta_K(-t) = 2\beta_K(0). \tag{5.14}$$

From (5.10) and (5.11), we see at once that $X^t K(x, y) = Y^t K(x, y)$. Using this and (5.12), we get

$$\begin{aligned}&\alpha_K(x + y + t) + \beta_K(x - y + t) - 2A_K(x + t, y)\\ &= \alpha_K(x + y + t) + \beta_K(x - y - t) - 2A_K(x, y + t).\end{aligned}$$

Since A_K is biadditive and skew-symmetric, this last relation simplifies to

$$\beta_K(x - y + t) - \beta_K(x - y - t) = 2A_K(t, x + y).$$

Putting $x = y$, we get $\beta_K(t) - \beta_K(-t) = 2A_K(t, 2x)$, and adding to (5.14), we have

$$\beta_K(t) = \beta_K(0) + A_K(t, 2x).$$

Putting $x = 0$, we see that $\beta_K(t) = \beta_K(0) =: \beta$, a constant, yielding $A_K(x, y) = 0$, and so $K(x, y) = \alpha_K(x + y) + \beta$. Since K is skew-symmetric, equating $K(x, y) = -K(y, x)$, we deduce that $0 = \alpha_K(x + y) + \beta = K(x, y)$.

The determination of H proceeds analogously. Using skew-symmetry in (2.3), we get

$$\left(X^t Y^t + X^{-t} Y^{-t}\right) D^{(1,4)} + \left(Y^{-t} + Y^t\right) D^{(2,5)} + \left(X^{-t} + X^t\right) D^{(3,6)} = 0. \tag{5.15}$$

Subtracting (5.15) from (2.3) and rearranging, we get

$$\left(X^{-t}+X^{t}\right)H(x,y)=\left(Y^{-t}+Y^{t}\right)H(x,y). \tag{5.16}$$

Defining the function $H_1\colon G\times G\to\mathbb{C}$ by

$$H_1(x,y):=H\left(\frac{x+y}{2},\frac{x-y}{2}\right),\ \text{ or equivalently,}\quad H(x,y)=H_1(x+y,x-y)$$

and rewriting (5.16) in terms of H_1, we get

$$\begin{aligned}&H_1(x+y-t,x-y-t)+H_1(x+y+t,x-y+t)\\&=H_1(x+y-t,x-y+t)+H_1(x+y+t,x-y-t).\end{aligned}$$

Putting $u=x+y-t,\ v=x-y-t$, then replacing $2t$ by t and rearranging, we get

$$H_1(u+t,v+t)=H_1(u+t,v)+H_1(u,v+t)-H_1(u,v),$$

which is the McKiernan's functional equation mentioned in Proposition 1.1, whose solution is of the form

$$H_1(x,y)=\alpha_H(x)+\beta_H(y)+A_H(x,y),$$

with the functions α_H,β_H,A_H as stated above. Reverting back to H, we get

$$H(x,y)=\alpha_H(x+y)+\beta_H(x-y)+A_H(x+y,x-y).$$

As in the previous case, using the fact that A_H is skew-symmetric and biadditive, the shape of H reduces to

$$H(x,y)=\alpha_H(x+y)+\beta_H(x-y)-2A_H(x,y).$$

Since H is skew-symmetric, solving $H(x,y)=-H(y,x)$, we get

$$2\alpha_H(x+y)+\beta_H(x-y)+\beta_H(y-x)=0.$$

Substituting $x=y$, we get $\alpha_H(x)=-\beta_H(0)$, a constant and so

$$\beta_H(t)+\beta_H(-t)=2\beta_H(0). \tag{5.17}$$

Incorporating all these details, the shape of H follows. □

Remark 5.1. R1) Setting $t=0$ in (5.15), and simplifying, we get

$$(f_1-f_4)+(f_2-f_5)+(f_3-f_6)=0.$$

R2) From Lemma 5.2 and Lemma 4.2, we have

$$\begin{aligned}0=K&=f_2+f_5-f_3-f_6=-4k(x,y) &(5.18)\\ -\beta_H(0)+\beta_H(x-y)-2A_H(x,y)=H&=f_2-f_5-f_3+f_6=-4w(x,y). &(5.19)\end{aligned}$$

Throughout the rest of this section, we shall deal only with the case where $f_1 = f_4$, which, by remark R1), gives rise to

$$f_2 - f_5 + f_3 - f_6 = 0.$$

Combining with (5.18) and (5.19), we get

$$f_2 - f_6 = -2k, \ \ f_3 - f_6 = 2w \tag{5.20}$$

Theorem 5.3. *Assume* $f_1, f_2, f_3, f_4, f_5, f_6 \colon G \times G \to \mathbb{C}$ *are skew-symmetric functions satisfying* (PDF)*.*
I. If

$$f_1 = f_4 = \frac{1}{2}\left(f_2 + f_3\right),$$

then there are biadditive function $B(x,y) \colon G \times G \to \mathbb{C}$ *and biadditive, skew-symmetric function* $A(x,y) \colon G \times G \to \mathbb{C}$*, two constants* α_2, Φ*, arbitrary functions* $\psi, \chi, \beta_2 \colon G \to \mathbb{C}$ *with* $\beta_2(t) + \beta_2(-t) = 2\beta_2(0)$ *such that*

$$\begin{aligned} f_1(x,y) &= f_4(x,y) = B(x,y) - \psi(x) + \psi(y) + \chi(x-y) - \Phi - B(x,x) \\ f_2(x,y) &= f_6(x,y) = f_1(x,y) + \{\alpha_2 + \beta_2(x-y) - A(x,y)\} \\ f_3(x,y) &= f_5(x,y) = f_1(x,y) - \{\alpha_2 + \beta_2(x-y) - A(x,y)\}. \end{aligned}$$

II. If

$$f_1 = f_4 = \frac{1}{2}\left(f_2 + f_6\right) = \frac{1}{2}\left(f_3 + f_5\right),$$

then there are biadditive function $B(x,y) \colon G \times G \to \mathbb{C}$ *, a constant* Φ*, arbitrary functions* $\psi, \chi \colon G \to \mathbb{C}$ *such that*

$$\begin{aligned} f_1(x,y) &= f_2(x,y) = f_3(x,y) = f_4(x,y) = f_5(x,y) = f_6(x,y) \\ &= B(x,y) - \psi(x) + \psi(y) + \chi(x-y) - \Phi - B(x,x). \end{aligned}$$

III. If

$$f_1 = f_4 = \frac{1}{2}\left(f_2 + f_5\right) = \frac{1}{2}\left(f_3 + f_6\right),$$

then there are biadditive function $B(x,y) \colon G \times G \to \mathbb{C}$ *and biadditive, skew-symmetric function* $A(x,y) \colon G \times G \to \mathbb{C}$*, two constants* α_2, Φ*, arbitrary functions* $\psi, \chi, \beta_2 \colon G \to \mathbb{C}$ *with* $\beta_2(t) + \beta_2(-t) = 2\beta_2(0)$ *such that*

$$\begin{aligned} f_1(x,y) &= f_4(x,y) = B(x,y) - \psi(x) + \psi(y) + \chi(x-y) - \Phi - B(x,x) \\ f_2(x,y) &= f_6(x,y) = f_1(x,y) + \{\alpha_2 + \beta_2(x-y) - A(x,y)\} \\ f_3(x,y) &= f_5(x,y) = f_1(x,y) - \{\alpha_2 + \beta_2(x-y) - A(x,y)\}. \end{aligned}$$

IV. Assume $f_1 = f_6$ *and*

$$f_1(a,b) + f_6(c,d) + f_6(e,f) = f_1(a',b') + f_6(c',d') + f_6(e',f') \tag{5.21}$$

whenever $a+b+c+d+e+f = a'+b'+c'+d'+e'+f'$*. Then there are biadditive function* $B(x,y)\colon G\times G\to\mathbb{C}$*, a constant* Φ*, arbitrary functions* $\psi,\chi,\beta_2\colon G\to\mathbb{C}$ *with* $\beta_2(t)+\beta_2(-t)=2\beta_2(0)$ *such that*

$$\begin{aligned} f_1(x,y) &= f_2(x,y) = f_4(x,y) = f_6(x,y) \\ &= B(x,y) - \psi(x) + \psi(y) + \chi(x-y) - \Phi - B(x,x) \\ f_3(x,y) &= f_5(x,y) = f_1(x,y) + \beta_2(0) - \beta_2(x-y). \end{aligned}$$

Proof. I. Using $f_1 = f_4,\ f_2+f_3 = f_5+f_6$, substituting $g = 2f_1 = 2f_4 = f_2+f_3 = f_5+f_6$ in (2.5) and simplifying, we get

$$\begin{aligned} &g(x+t,y+t) + g(x-t,y) + g(x,y-t) \\ &= g(x-t,y-t) + g(x,y+t) + g(x+t,y), \end{aligned}$$

which is the Sahoo-Székelyhidi's functional equation mentioned in Proposition 1.2, and its solution is of the form

$$g(x,y) = B_1(x,y) + \phi_1(x) + \psi_1(y) + \chi_1(x-y),$$

where $B_1\colon G\times G\to\mathbb{C}$ is biadditive and $\phi_1,\psi_1,\chi_1\colon G\to\mathbb{C}$ are arbitrary functions. Since g is skew-symmetric, applying Lemma 5.1, we deduce that

$$g(x,y) = B_1(x,y) - \psi_1(x) + \psi_1(y) + \chi_1(x-y) - \Phi_1 - B_1(x,x), \tag{5.22}$$

where

$$\Phi_1 = \chi_1(0), \quad \chi_1(x) + \chi_1(-x) = 2\Phi_1 + B_1(x,x).$$

Putting

$$B = B_1/2,\ \psi = \psi_1/2,\ \chi = \chi_1/2,\ \Phi = \Phi_1/2,$$

this gives the shapes of f_1 and f_4. Using (4.5), (5.18) and (5.19), we get

$$\begin{aligned} 2f_3 &= (f_2+f_3) + (-f_2+f_3) = g + 2(w+k) \\ &= g(x,y) - \frac{1}{4}\left\{-\beta_H(0) + \beta_H(x-y)\right\} + \frac{1}{2}A_H(x,y) - 0 \end{aligned}$$

and the shape of f_3 follows by putting $\alpha_2 = -\frac{1}{4}\beta_H(0),\ \beta_2 = \frac{1}{4}\beta_H,\ A = \frac{1}{2}A_H$. The shapes of other solution functions follow similarly noting from above and (4.6) that

$$f_2 = g - f_3,\ f_5+f_6 = g,\ f_5 - f_6 = 2(w-k),$$

and then using (5.18) and (5.19).

II. The proof proceeds much the same as that of part I. Using $f_1 = f_4,\ f_2+f_3 = f_5+f_6$, substituting $g = 2f_1 = 2f_4 = f_2+f_6 = f_3+f_5$ in (2.6) and simplifying, we see that g satisfies the the Sahoo-Székelyhidi's functional equation. Using skew-symmetry, we deduce that it must be of the form (5.22). Thus,

$$2f_1 = 2f_4 = f_2+f_6 = f_3+f_5 = g$$

$$= B_1(x,y) - \psi_1(x) + \psi_1(y) + \chi_1(x-y) - \Phi_1 - B_1(x,x).$$

The shapes of the solution functions follow by using (5.20), (4.16) and (5.19).

III. Using $f_1 = f_4$, $f_2 + f_3 = f_5 + f_6$, substituting $g = 2f_1 = 2f_4 = f_2 + f_5 = f_3 + f_6$ in (2.6) and simplifying, we see that g satisfies the Sahoo-Székelyhidi's functional equation. Using skew-symmetry, we deduce that its solution must be of the form (5.22). Thus,

$$\begin{aligned}&2f_1 = 2f_4 = f_2 + f_5 = f_3 + f_6 = g\\ &\quad = B_1(x,y) - \psi_1(x) + \psi_1(y) + \chi_1(x-y) - \Phi_1 - B_1(x,x).\end{aligned}$$

The shapes of the solution functions follow by using (5.20), (5.18) and (5.19).

IV. Solving for f_2, f_3, f_5 in terms of f_6 in (5.20) and (5.18), we get

$$f_2 = f_6 - 2k,\ f_3 = f_6 + 2w,\ f_5 = f_6 + 2(w-k).$$

Substituting these functions in (PDF), using (5.18) and (5.19), we get

$$\begin{aligned}&f_1(x+t,y+t) + f_6(x-t,y) + f_6(x,y-t) - A_H(x,t)\\ &\quad = f_4(x-t,y-t) + f_6(x+t,y) + f_6(x,y+t) + A_H(t,y).\end{aligned}\tag{5.23}$$

Substituting $t=0$ in (5.23), we get

$$\begin{aligned}&f_1(x,y) + f_6(x,y) + f_6(x,y)\\ &\qquad = f_4(x,y) + f_6(x,y) + f_6(x,y),\end{aligned}$$

and so $f_1(x,y) = f_4(x,y)$. Substituting $x = 0$ in (5.23), we get

$$\begin{aligned}&f_1(t,y+t) + f_6(-t,y) + f_6(0,y-t)\\ &\qquad = f_4(-t,y-t) + f_6(t,y) + f_6(0,y+t) + A_H(t,y).\end{aligned}$$

Appealing to (5.21), we get $A_H(x,y) = 0$. Substituting $A_H(x,y) = 0$ in (5.23), we get

$$\begin{aligned}&f_1(x+t,y+t) + f_6(x-t,y) + f_6(x,y-t)\\ &\qquad = f_4(x-t,y-t) + f_6(x+t,y) + f_6(x,y+t).\end{aligned}$$

Using $f_1 = f_4 = f_6$, this last relation is the Sahoo-Székelyhidi's functional equation, and so its solution is the form

$$f_1 = f_4 = f_6 = B(x,y) - \psi(x) + \psi(y) + \chi(x-y) - \Phi - B(x,x).$$

Using (5.20), (5.18) and (5.19), we get

$$\begin{aligned}&f_2 = f_6 = B(x,y) - \psi(x) + \psi(y) + \chi(x-y) - \Phi - B(x,x)\\ &f_3 = f_6 + 2w = f_6 + \frac{1}{2}\beta_H(0) - \frac{1}{2}\beta_H(x-y).\end{aligned}$$

Putting $\beta_2 = \beta_H/2$ and observing from (5.17) that $\beta_2(t) + \beta_2(-t) = -2\alpha_2$, the shape of f_3 follows. The remaining functions are $f_5 = f_2 - f_6 + f_3 = f_3$. $\square$

References

[1] Haruki, S., A wavelike functional equation of Pexider type, *Aequationes Math.* 63 (2002), 201–209.

[2] Haruki, S., Nakagiri, S., A pexiderized wavelike partial difference functional equation, *Aequationes Math.* 74 (2007), 1–6.

[3] McKiernan, M. A., The general solution of some finite difference equations analogous to the wave equation, *Aequationes Math.* 8 (1972), 263–266.

[4] Sahoo, P. K., Székelyhidi, L., On a functional equation related to digital filtering, *Aequationes Math.* 62 (2001), 280–285.

Annales Mathematicae et Informaticae
46 (2016) pp. 97–114
http://ami.ektf.hu

The Miki-type identity for the Apostol-Bernoulli numbers

Orli Herscovici, Toufik Mansour

Department of Mathematics, University of Haifa, 3498838 Haifa, Israel,
orli.herscovici@gmail.com and tmansour@univ.haifa.ac.il

Submitted October 19, 2015 — Accepted September 7, 2016

Abstract

We study analogues of the Miki, Matiyasevich, and Euler identities for the Apostol-Bernoulli numbers and obtain the analogues of the Miki and Euler identities for the Apostol-Genocchi numbers.

Keywords: Apostol-Bernoulli numbers; Apostol-Genocchi numbers; Miki identity; Matiyasevich identity; Euler identity

MSC: 05A19; 11B68

1. Introduction

The Apostol-Bernoulli numbers are defined in [2] as

$$\frac{t}{\lambda e^t - 1} = \sum_{n=0}^{\infty} \mathcal{B}_n(\lambda)\frac{t^n}{n!}. \tag{1.1}$$

Note that at $\lambda = 1$ this generating function becomes

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!},$$

where B_n is the classical nth Bernoulli number. Moreover, $\mathcal{B}_0 = \mathcal{B}_0(\lambda) = 0$ while $B_0 = 1$ (see [9]). The Genocchi numbers are defined by the generating function

$$\frac{2t}{e^t + 1} = \sum_{n=0}^{\infty} G_n \frac{t^n}{n!},$$

which are closely related to the classical Bernoulli numbers and the special values of the Euler polynomials. It is known that $G_n = 2(1-2^n)B_n$ and $G_n = nE_{n-1}(0)$, where $E_n(0)$ is a value of the Euler polynomials evaluated at 0 (sometimes are called the Euler numbers) [4, 10, 11]. Likewise the Apostol-Bernoulli numbers, the Apostol-Genocchi numbers are defined by their generating function as

$$\frac{2t}{\lambda e^t + 1} = \sum_{n=0}^{\infty} \mathcal{G}_n(\lambda)\frac{t^n}{n!} \tag{1.2}$$

with $\mathcal{G}_0 = \mathcal{G}_0(\lambda) = 0$.

Over the years, different identities were obtained for the Bernoulli numbers (for instance, see [3, 4, 6, 7, 10, 12, 16, 17]). The Euler identity for the Bernoulli numbers is given by (see [6, 15])

$$\sum_{k=2}^{n-2} \binom{n}{k} B_k B_{n-k} = -(n+1)B_n, \qquad (n \geq 4). \tag{1.3}$$

Its analogue for convolution of Bernoulli and Euler numbers was obtained in [10] using the p-adic integrals. The similar convolution was obtained for the generalized Apostol-Bernoulli polynomials in [13]. In 1978, Miki [15] found a special identity involving two different types of convolution between Bernoulli numbers:

$$\sum_{k=2}^{n-2} \frac{B_k}{k}\frac{B_{n-k}}{n-k} - \sum_{k=2}^{n-2} \binom{n}{k} \frac{B_k}{k}\frac{B_{n-k}}{n-k} = 2H_n\frac{B_n}{n}, \qquad (n \geq 4), \tag{1.4}$$

where $H_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$ is the nth harmonic number. Different kinds of proofs of this identity were represented in [1, 5, 8]. Gessel [8] generalized the Miki identity for the Bernoulli polynomials. Another generalization of the Miki identity for the Bernoulli and Euler polynomials was obtained in [16]. In 1997, Matiyasevich [1, 14] found an identity involving two types of convolution between Bernoulli numbers

$$(n+2)\sum_{k=2}^{n-2} B_k B_{n-k} - 2\sum_{k=2}^{n-2} \binom{n+2}{k} B_k B_{n-k} = n(n+1)B_n. \tag{1.5}$$

The analogues of the Euler, Miki and Matiyasevich identities for the Genocchi numbers were obtained in [1]. In this paper, we represent the analogues of these identities for the Apostol-Bernoulli and the Apostol-Genocchi numbers.

2. The analogues for the Apostol-Bernoulli numbers

In our work, we use the generating functions method to obtain new analogues of the known identities for the Apostol-Bernoulli numbers (see [1, 9]). It is easy to show that

$$\frac{1}{\lambda e^a - 1} \cdot \frac{1}{\mu e^b - 1} = \frac{1}{\lambda\mu e^{a+b} - 1}\left(1 + \frac{1}{\lambda e^a - 1} + \frac{1}{\mu e^b - 1}\right). \tag{2.1}$$

Let us take $a = xt$ and $b = x(1-t)$ and multiply both sides of the identity (2.1) by $t(1-t)x^2$.

$$\begin{aligned}
&\frac{tx}{\lambda e^{tx}-1}\frac{(1-t)x}{\mu e^{(1-t)x}-1}\\
&\quad = \frac{t(1-t)x^2}{\lambda\mu e^{tx+(1-t)x}}\left(1+\frac{1}{\lambda e^{tx}-1}+\frac{1}{\mu e^{(1-t)x}-1}\right)\\
&\quad = \frac{x}{\lambda\mu e^{x}-1}\left(t(1-t)x+(1-t)\frac{tx}{\lambda e^{tx}-1}+t\frac{(1-t)x}{\mu e^{(1-t)x}-1}\right). \qquad (2.2)
\end{aligned}$$

By using (1.1) and the Cauchy product, we get on the LH side of (2.2)

$$\begin{aligned}
\frac{tx}{\lambda e^{tx}-1}\frac{(1-t)x}{\mu e^{(1-t)x}-1} &= \left(\sum_{n=0}^{\infty}\mathcal{B}_n(\lambda)\frac{t^n x^n}{n!}\right)\left(\sum_{n=0}^{\infty}\mathcal{B}_n(\mu)\frac{(1-t)^n x^n}{n!}\right)\\
&= \sum_{n=0}^{\infty}\left[\sum_{k=0}^{n}\binom{n}{k}\mathcal{B}_k(\lambda)t^k\mathcal{B}_{n-k}(\mu)(1-t)^{n-k}\right]\frac{x^n}{n!}, \qquad (2.3)
\end{aligned}$$

and on the RH side of (2.2) we obtain

$$\begin{aligned}
&\frac{x}{\lambda\mu e^{x}-1}\left(t(1-t)x+(1-t)\frac{tx}{\lambda e^{tx}-1}+t\frac{(1-t)x}{\mu e^{(1-t)x}-1}\right)\\
&= \sum_{n=0}^{\infty}\mathcal{B}_n(\lambda\mu)\frac{x^n}{n!}\cdot\\
&\qquad \cdot\left(t(1-t)x+(1-t)\sum_{n=0}^{\infty}\mathcal{B}_n(\lambda)\frac{t^n x^n}{n!}+t\sum_{n=0}^{\infty}\mathcal{B}_n(\mu)\frac{(1-t)^n x^n}{n!}\right)\\
&= t(1-t)\sum_{n=1}^{\infty}\mathcal{B}_{n-1}(\lambda\mu)n\frac{x^n}{n!}\\
&\qquad + (1-t)\sum_{n=0}^{\infty}\left[\sum_{k=0}^{n}\binom{n}{k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)t^{n-k}\right]\frac{x^n}{n!}\\
&\qquad\qquad + t\sum_{n=0}^{\infty}\left[\sum_{k=0}^{n}\binom{n}{k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu)(1-t)^{n-k}\right]\frac{x^n}{n!}. \qquad (2.4)
\end{aligned}$$

By comparing the coefficients of $\frac{x^n}{n!}$ on left (2.3) and right (2.4) hand sides, we get

$$\begin{aligned}
&\sum_{k=0}^{n}\binom{n}{k}t^k(1-t)^{n-k}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)\\
&\qquad = nt(1-t)\mathcal{B}_{n-1}(\lambda\mu)+(1-t)\sum_{k=0}^{n}\binom{n}{k}t^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)
\end{aligned}$$

$$+\,t\sum_{k=0}^{n}\binom{n}{k}(1-t)^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu). \tag{2.5}$$

It follows from (1.1) that $\mathcal{B}_n(1)=B_n$. It is well known that $B_0=1$, but from (1.1) we get $\mathcal{B}_0=0$. Therefore, we concentrate the members, containing the 0th index (the cases $k=0$ and $k=n$), out of the sums. The sum on the left hand side of (2.5) can be rewritten as

$$\begin{aligned}
&\sum_{k=0}^{n}\binom{n}{k}t^k(1-t)^{n-k}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)\\
&=\sum_{k=1}^{n-1}\binom{n}{k}t^k(1-t)^{n-k}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)\\
&\qquad\qquad+(1-t)^n\mathcal{B}_0(\lambda)\mathcal{B}_n(\mu)+t^n\mathcal{B}_n(\lambda)\mathcal{B}_0(\mu) \qquad (2.6)\\
&=\sum_{k=1}^{n-1}\binom{n}{k}t^k(1-t)^{n-k}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)+(1-t)^n\delta_{1,\lambda}\mathcal{B}_n(\mu)+t^n\mathcal{B}_n(\lambda)\delta_{1,\mu},
\end{aligned}$$

where $\delta_{p,q}$ is the Kronecker symbol. On the right hand side of (2.5) we have that the first sum can be rewritten as

$$\begin{aligned}
&(1-t)\sum_{k=0}^{n}\binom{n}{k}t^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)\\
&=(1-t)\sum_{k=1}^{n-1}\binom{n}{k}t^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)\\
&\qquad\qquad+(1-t)t^n\mathcal{B}_0(\lambda\mu)\mathcal{B}_n(\lambda)+(1-t)\mathcal{B}_n(\lambda\mu)\mathcal{B}_0(\lambda) \qquad (2.7)\\
&=(1-t)\sum_{k=1}^{n-1}\binom{n}{k}t^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)+(1-t)t^n\delta_{1,\lambda\mu}\mathcal{B}_n(\lambda)+(1-t)\mathcal{B}_n(\lambda\mu)\delta_{1,\lambda},
\end{aligned}$$

and the second sum can be rewritten as

$$\begin{aligned}
&t\sum_{k=0}^{n}\binom{n}{k}(1-t)^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu)\\
&=t\sum_{k=1}^{n-1}\binom{n}{k}(1-t)^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu)\\
&\qquad\qquad+t(1-t)^n\mathcal{B}_0(\lambda\mu)\mathcal{B}_n(\mu)+t\mathcal{B}_n(\lambda\mu)\mathcal{B}_0(\mu) \qquad (2.8)\\
&=t\sum_{k=1}^{n-1}\binom{n}{k}(1-t)^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu)+t(1-t)^n\delta_{1,\lambda\mu}\mathcal{B}_n(\mu)+t\mathcal{B}_n(\lambda\mu)\delta_{1,\mu}.
\end{aligned}$$

By substituting the detailed expressions (2.6)–(2.8) back into (2.5), we get

$$\sum_{k=1}^{n-1}\binom{n}{k}t^k(1-t)^{n-k}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)+(1-t)^n\delta_{1,\lambda}\mathcal{B}_n(\mu)+t^n\mathcal{B}_n(\lambda)\delta_{1,\mu}$$

$$\begin{aligned}
&= nt(1-t)\mathcal{B}_{n-1}(\lambda\mu) + (1-t)\mathcal{B}_n(\lambda\mu)\delta_{1,\lambda} \\
&\quad + (1-t)\sum_{k=1}^{n-1}\binom{n}{k}t^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda) + (1-t)t^n\delta_{1,\lambda\mu}\mathcal{B}_n(\lambda) \\
&\quad + t\sum_{k=1}^{n-1}\binom{n}{k}(1-t)^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu) + t(1-t)^n\delta_{1,\lambda\mu}\mathcal{B}_n(\mu) \\
&\quad + t\mathcal{B}_n(\lambda\mu)\delta_{1,\mu}.
\end{aligned} \tag{2.9}$$

By dividing both sides of (2.9) by $t(1-t)$, we obtain

$$\begin{aligned}
&\sum_{k=1}^{n-1}\binom{n}{k}t^{k-1}(1-t)^{n-k-1}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu) + \frac{(1-t)^{n-1}}{t}\delta_{1,\lambda}\mathcal{B}_n(\mu) + \frac{t^{n-1}}{1-t}\mathcal{B}_n(\lambda)\delta_{1,\mu} \\
&= n\mathcal{B}_{n-1}(\lambda\mu) + \sum_{k=1}^{n-1}\binom{n}{k}t^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda) + t^{n-1}\delta_{1,\lambda\mu}\mathcal{B}_n(\lambda) \\
&+ \frac{1}{t}\mathcal{B}_n(\lambda\mu)\delta_{1,\lambda} + \sum_{k=1}^{n-1}\binom{n}{k}(1-t)^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu) + (1-t)^{n-1}\delta_{1,\lambda\mu}\mathcal{B}_n(\mu) \\
&+ \frac{1}{1-t}\mathcal{B}_n(\lambda\mu)\delta_{1,\mu}.
\end{aligned} \tag{2.10}$$

We rewrite the (2.10) as

$$\begin{aligned}
&\sum_{k=1}^{n-1}\binom{n}{k}t^{k-1}(1-t)^{n-k-1}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu) \\
&\qquad = n\mathcal{B}_{n-1}(\lambda\mu) + \sum_{k=1}^{n-1}\binom{n}{k}t^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda) \\
&\qquad\qquad + \sum_{k=1}^{n-1}\binom{n}{k}(1-t)^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu) + A^\delta,
\end{aligned} \tag{2.11}$$

where

$$\begin{aligned}
A^\delta &= \frac{1}{t}(\mathcal{B}_n(\lambda\mu) - (1-t)^{n-1}\mathcal{B}_n(\mu))\delta_{1,\lambda} + (t^{n-1}\mathcal{B}_n(\lambda) \\
&\quad + (1-t)^{n-1}\mathcal{B}_n(\mu))\delta_{1,\lambda\mu} + \frac{1}{1-t}(\mathcal{B}_n(\lambda\mu) - t^{n-1}\mathcal{B}_n(\lambda))\delta_{1,\mu}.
\end{aligned} \tag{2.12}$$

By integrating (2.11) between 0 and 1 with respect to t and using the formulae

$$\int_0^1 t^p(1-t)^q dt = \frac{p!q!}{(p+q+1)!}, \qquad p, q \geq 0,$$

$$\int_0^1 \frac{1-t^{p+1}-(1-t)^{p+1}}{t(1-t)}dt = 2\int_0^1 \frac{1-t^p}{1-t}dt = 2H_p, \qquad p \geq 1,$$

we obtain

$$\begin{aligned}
&\int_0^1 \sum_{k=1}^{n-1} \binom{n}{k} t^{k-1}(1-t)^{n-k-1}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)dt \\
&= \int_0^1 n\mathcal{B}_{n-1}(\lambda\mu)dt + \int_0^1 \sum_{k=1}^{n-1} \binom{n}{k} t^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)dt \\
&+ \int_0^1 \sum_{k=1}^{n-1} \binom{n}{k} (1-t)^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu)dt + \int_0^1 A^\delta dt,
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
&\sum_{k=1}^{n-1} \binom{n}{k} \frac{(k-1)!(n-k-1)!}{(n-1)!}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu) \\
&= n\mathcal{B}_{n-1}(\lambda\mu) + \sum_{k=1}^{n-1} \binom{n}{k} \mathcal{B}_k(\lambda\mu)\frac{\mathcal{B}_{n-k}(\lambda)}{n-k} \\
&+ \sum_{k=1}^{n-1} \binom{n}{k} \mathcal{B}_k(\lambda\mu)\frac{\mathcal{B}_{n-k}(\mu)}{n-k} + \int_0^1 A^\delta dt. \qquad (2.13)
\end{aligned}$$

By dividing both sides of (2.13) by n and performing elementary transformations of the binomial coefficients of (2.13), we can state the following result.

Theorem 2.1. *For all $n \geq 2$,*

$$\begin{aligned}
&\sum_{k=1}^{n-1} \frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k} \\
&= \mathcal{B}_{n-1}(\lambda\mu) + \sum_{k=1}^{n-1} \binom{n-1}{k-1} \frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\lambda)+\mathcal{B}_{n-k}(\mu)}{n-k} + \frac{1}{n}\int_0^1 A^\delta dt, \qquad (2.14)
\end{aligned}$$

where A^δ is given by (2.12).

We have to consider different possible cases for λ and μ values.

Example 2.2. Let $\lambda = 1$, $\mu = 1$. It follows from (2.12) that

$$A^\delta = \frac{1-(1-t)^{n-1}}{t}B_n + (t^{n-1}+(1-t)^{n-1})B_n + \frac{1-t^{n-1}}{1-t}B_n. \qquad (2.15)$$

Therefore, the integrating of (2.15) between 0 and 1 with respect to t gives

$$\int_0^1 A^\delta dt = \int_0^1 \frac{1-t^n-(1-t)^n}{t(1-t)} B_n dt + \int_0^1 t^{n-1} B_n dt + \int_0^1 (1-t)^{n-1} B_n dt = 2H_n B_n. \tag{2.16}$$

By substituting (2.16) back into (2.14) and replacing all $\mathcal{B}$ by B consistently with the case condition, we get

$$\sum_{k=1}^{n-1} \frac{B_k}{k}\frac{B_{n-k}}{n-k} - 2\sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{B_k}{k}\frac{B_{n-k}}{n-k} = B_{n-1} + 2H_n\frac{B_n}{n}.$$

Note that for even $n \geq 4$, all summands, containing odd-indexed Bernoulli numbers, equal zero. Thus, the sums must be limited from $k=2$ up to $n-2$ over even indexes only. Moreover, the term B_{n-1} on the RH side disappears from the same reason. Now we have

$$\sum_{k=2}^{n-2} \frac{B_k}{k}\frac{B_{n-k}}{n-k} - 2\sum_{k=2}^{n-2}\binom{n-1}{k-1}\frac{B_k}{k}\frac{B_{n-k}}{n-k} = 2H_n\frac{B_n}{n}.$$

In order to obtain the Miki identity (1.4), let us consider the sum

$$2\sum_{k=2}^{n-2}\binom{n-1}{k-1}\frac{B_k}{k}\frac{B_{n-k}}{n-k} = \frac{2}{n}\sum_{k=2}^{n-2}\frac{1}{n-k}\binom{n}{k}B_kB_{n-k}.$$

Finally, using $2\sum\limits_{k=2}^{n-2}\frac{1}{n-k}\binom{n}{k}B_kB_{n-k} = n\sum\limits_{k=2}^{n-2}\binom{n}{k}\frac{B_k}{k}\frac{B_{n-k}}{n-k}$ (see [1]), we obtain the known Miki identity (1.4) (see [1, 8, 15]).

Corollary 2.3. *Let $\mu \neq 1$. For all $n \geq 2$, the following identities are valid*

$$\sum_{k=1}^{n-1} \frac{B_k}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\mu)}{k}\frac{B_{n-k}+\mathcal{B}_{n-k}(\mu)}{n-k} = \mathcal{B}_{n-1}(\mu) + H_{n-1}\frac{\mathcal{B}_n(\mu)}{n}, \tag{2.17}$$

$$\sum_{k=1}^{n-1} \frac{\mathcal{B}_k(\frac{1}{\mu})}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k} - \frac{1}{n}\sum_{k=0}^{n-1}\binom{n}{k}B_k\frac{\mathcal{B}_{n-k}(\frac{1}{\mu})+\mathcal{B}_{n-k}(\mu)}{n-k} = B_{n-1}. \tag{2.18}$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, then

$$\sum_{k=1}^{n-1} \frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\lambda)+\mathcal{B}_{n-k}(\mu)}{n-k} = \mathcal{B}_{n-1}(\lambda\mu). \tag{2.19}$$

Proof. In the case $\lambda = 1$, $\mu \neq 1$, we have from (2.12) that $A^\delta = \frac{1-(1-t)^{n-1}}{t}\mathcal{B}_n(\mu)$. The integrating between 0 and 1 with respect to t gives

$$\int_0^1 A^\delta dt = \int_0^1 \frac{1-(1-t)^{n-1}}{t}\mathcal{B}_n(\mu)dt = H_{n-1}\mathcal{B}_n(\mu). \tag{2.20}$$

By substituting (2.20) into (2.14), we obtain

$$\sum_{k=1}^{n-1}\frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k}$$
$$= \mathcal{B}_{n-1}(\lambda\mu) + \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\lambda)+\mathcal{B}_{n-k}(\mu)}{n-k} + \frac{1}{n}H_{n-1}\mathcal{B}_n(\mu).$$

By taking into account that $\lambda = 1$ and $\mathcal{B}_p(1) = B_p$, we get the identity (2.17).

In order to prove (2.18), we suppose that $\lambda = \frac{1}{\mu} \neq 1$. Then, from (2.12), we obtain that $A^\delta = t^{n-1}\mathcal{B}_n(\frac{1}{\mu}) + (1-t)^{n-1}\mathcal{B}_n(\frac{1}{\mu})$. By integrating of A^δ between 0 and 1 with respect to t, we get

$$\int_0^1 A^\delta dt = \int_0^1 t^{n-1}\mathcal{B}_n(\frac{1}{\mu})dt + \int_0^1 (1-t)^{n-1}\mathcal{B}_n(\mu)dt = \frac{\mathcal{B}_n(\frac{1}{\mu})+\mathcal{B}_n(\mu)}{n}. \tag{2.21}$$

By substituting (2.21) into (2.14), we obtain

$$\sum_{k=1}^{n-1}\frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k} = \mathcal{B}_{n-1}(\lambda\mu) + \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\lambda)}{n-k}$$
$$+ \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\mu)}{n-k} + \frac{\mathcal{B}_n(\lambda)+\mathcal{B}_n(\frac{1}{\lambda})}{n^2}.$$

By substituting $\lambda = \frac{1}{\mu}$ into the last equation and using the facts that $\mathcal{B}_p(1) = B_p$ and $B_0 = 1$, we obtain (2.18).

Equation (2.19) follows from the fact that $A^\delta = 0$ for $\lambda, \mu, \lambda\mu \neq 1$. □

By integrating both sides of (2.9) from 0 to 1 with respect to t and multiplying by $(n+1)(n+2)$, we obtain the following result, which is an analogue of the Matiyasevich identity (1.5).

Theorem 2.4. *For all $n \geq 2$,*

$$(n+2)\sum_{k=1}^{n-1}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu) - \sum_{k=1}^{n-1}\binom{n+2}{k}\mathcal{B}_k(\lambda\mu)\left(\mathcal{B}_{n-k}(\lambda)+\mathcal{B}_{n-k}(\mu)\right)$$
$$= \frac{n(n+1)(n+2)}{6}\mathcal{B}_{n-1}(\lambda\mu) \tag{2.22}$$
$$+ \frac{(n-1)(n+2)}{2}\left(\mathcal{B}_n(\mu)\delta_{1,\lambda} + \mathcal{B}_n(\lambda)\delta_{1,\mu}\right) + (\mathcal{B}_n(\lambda)+\mathcal{B}_n(\mu))\delta_{1,\lambda\mu}.$$

Example 2.5. Let $\lambda = 1$, $\mu = 1$. Then, by using the fact that $\mathcal{B}_p(1) = B_p$, we obtain

$$(n+2)\sum_{k=1}^{n-1} B_k B_{n-k} - 2\sum_{k=1}^{n-1}\binom{n+2}{k} B_k B_{n-k} = n(n+1)B_n + \frac{n(n+1)(n+2)}{6}B_{n-1}. \tag{2.23}$$

Finally, by assuming that n is even and $n \geq 4$, we get that all terms, containing odd indexed Bernoulli numbers, equal zero. Under this condition the $(n-1)$st Bernoulli number on the RH side disappears, and the summation limits are from 2 till $n-2$. Thus, we obtain (1.5) (see also [1]).

Corollary 2.6. *Let $\mu \neq 1$. Then, for all $n \geq 2$, the following identities are valid:*

$$(n+2)\sum_{k=1}^{n-1} B_k \mathcal{B}_{n-k}(\mu) - \sum_{k=1}^{n-1}\binom{n+2}{k}\mathcal{B}_k(\mu)\left(B_{n-k} + \mathcal{B}_{n-k}(\mu)\right) = \frac{n(n+1)(n+2)}{6}\mathcal{B}_{n-1}(\mu) + \frac{(n-1)(n+2)}{2}\mathcal{B}_n(\mu), \tag{2.24}$$

$$(n+2)\sum_{k=1}^{n-1} \mathcal{B}_k(\frac{1}{\mu})\mathcal{B}_{n-k}(\mu) - \sum_{k=1}^{n-1}\binom{n+2}{k} B_k \left(\mathcal{B}_{n-k}(\frac{1}{\mu}) + \mathcal{B}_{n-k}(\mu)\right) = \frac{n(n+1)(n+2)}{6}B_{n-1} + \mathcal{B}_n(\frac{1}{\mu}) + \mathcal{B}_n(\mu). \tag{2.25}$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, then

$$(n+2)\sum_{k=1}^{n-1} \mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu) - \sum_{k=1}^{n-1}\binom{n+2}{k}\mathcal{B}_k(\lambda\mu)\left(\mathcal{B}_{n-k}(\lambda) + \mathcal{B}_{n-k}(\mu)\right) = \frac{n(n+1)(n+2)}{6}\mathcal{B}_{n-1}(\mu). \tag{2.26}$$

Proof. By substituting $\lambda = 1$ into (2.22) and using the facts that $\mathcal{B}_p(1) = B_p$ and $\delta_{1,\mu} = \delta_{1,\lambda\mu} = 0$, we obtain (2.24). By substituting $\lambda = \frac{1}{\mu}$ into (2.22) and using the fact that $\delta_{1,\lambda} = \delta_{1,\mu} = 0$, we obtain (2.25). Equation (2.26) follows from (2.22) by using the fact that $\delta_{1,\lambda} = \delta_{1,\mu} = \delta_{1,\lambda\mu} = 0$. □

By dividing (2.9) by t and substituting $t = 0$, we obtain the following analogue of the Euler identity (1.3).

Theorem 2.7 (The Euler identity analogue)**.** *For all $n \geq 2$,*

$$\sum_{k=1}^{n-1}\binom{n}{k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu) = n\mathcal{B}_1(\lambda)\mathcal{B}_{n-1}(\mu) - n\mathcal{B}_{n-1}(\lambda\mu) - n\mathcal{B}_{n-1}(\lambda\mu)\mathcal{B}_1(\lambda) - (n-1)\mathcal{B}_n(\mu)\delta_{1,\lambda} - \mathcal{B}_n(\lambda)\delta_{1,\mu} - \mathcal{B}_n(\mu)\delta_{1,\lambda\mu}. \tag{2.27}$$

Proof. By dividing (2.9) by t, we obtain

$$\begin{aligned}
&\sum_{k=1}^{n-1}\binom{n}{k}t^{k-1}(1-t)^{n-k}\mathcal{B}_k(\lambda)\mathcal{B}_{n-k}(\mu)+\frac{(1-t)^n}{t}\delta_{1,\lambda}\mathcal{B}_n(\mu)+t^{n-1}\mathcal{B}_n(\lambda)\delta_{1,\mu}\\
&\qquad=n(1-t)\mathcal{B}_{n-1}(\lambda\mu)+(1-t)\sum_{k=1}^{n-1}\binom{n}{k}t^{n-k-1}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)\\
&\qquad+(1-t)t^{n-1}\delta_{1,\lambda\mu}\mathcal{B}_n(\lambda)+\frac{(1-t)}{t}\mathcal{B}_n(\lambda\mu)\delta_{1,\lambda} \qquad (2.28)\\
&\qquad+\sum_{k=1}^{n-1}\binom{n}{k}(1-t)^{n-k}\mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu)\\
&\qquad+(1-t)^n\delta_{1,\lambda\mu}\mathcal{B}_n(\mu)+\mathcal{B}_n(\lambda\mu)\delta_{1,\mu}.
\end{aligned}$$

Consider now the difference $\frac{(1-t)^n}{t}\delta_{1,\lambda}\mathcal{B}_n(\mu)-\frac{1-t}{t}\mathcal{B}_n(\lambda\mu)\delta_{1,\lambda}$. It is obviously that

$$\begin{aligned}
&\frac{(1-t)^n}{t}\delta_{1,\lambda}\mathcal{B}_n(\mu)-\frac{1-t}{t}\mathcal{B}_n(\lambda\mu)\delta_{1,\lambda}\\
&\qquad=\frac{(1-t)^n}{t}\delta_{1,\lambda}\mathcal{B}_n(\mu)-\frac{1-t}{t}\mathcal{B}_n(\mu)\delta_{1,\lambda}\\
&\qquad=\delta_{1,\lambda}\mathcal{B}_n(\mu)\frac{\sum\limits_{j=0}^{n}\binom{n}{j}(-t)^j-1+t}{t} \qquad (2.29)\\
&\qquad=\delta_{1,\lambda}\mathcal{B}_n(\mu)\left(-\sum_{j=2}^{n}\binom{n}{j}(-t)^{j-1}-(n-1)\right).
\end{aligned}$$

By substituting $t=0$ into (2.28) and using (2.29), we obtain (2.27). □

Example 2.8. Let $\lambda=1$, $\mu=1$. Then, by using the fact that $B_0=1$, we get

$$\sum_{k=0}^{n-1}\binom{n}{k}B_kB_{n-k}=-nB_n-nB_{n-1}.$$

Note that for $n\geq 4$, the odd Bernoulli numbers equal to zero and, thus, only one of the members on the right hand side will stay. Therefore, by assuming that $n\geq 4$ and n is even, we obtain the Euler identity (1.3) (see also [1, 6]).

Corollary 2.9. *For all $n\geq 2$ and $\mu\neq 1$, the following identities are valid:*

$$\sum_{k=1}^{n-1}\binom{n}{k}\mathcal{B}_k(\mu)\mathcal{B}_{n-k}(\mu)=-(n-1)\mathcal{B}_n(\mu)-n\mathcal{B}_{n-1}(\mu), \qquad (2.30)$$

$$\sum_{k=0}^{n-1}\binom{n}{k}B_k\mathcal{B}_{n-k}(\frac{1}{\mu})=n\mathcal{B}_1(\mu)\mathcal{B}_{n-1}(\frac{1}{\mu})+nB_{n-1}\mathcal{B}_1(\frac{1}{\mu}). \qquad (2.31)$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, *then*

$$\sum_{k=1}^{n-1} \binom{n}{k} \mathcal{B}_k(\lambda\mu)\mathcal{B}_{n-k}(\mu) = n\mathcal{B}_1(\lambda)\mathcal{B}_{n-1}(\mu) - n(1+\mathcal{B}_1(\lambda))\mathcal{B}_{n-1}(\lambda\mu).$$

Identity (2.30) is obtained by substituting $\lambda = 1$ into (2.27), and Identity (2.31) is obtained in case $\lambda\mu = 1$. Note that here we use the fact that $\mathcal{B}_1(\mu) = \frac{1}{\mu-1}$ and, therefore, $\mathcal{B}_1(1/\mu) = -(\mathcal{B}_1(\mu)+1)$.

3. Identities for the Apostol-Genocchi numbers

Following the same technique we used in the previous section, we will obtain the analogues of the Miki and Euler identities for the Apostol-Genocchi numbers. It is easy to show that

$$\frac{1}{\lambda e^a + 1} \cdot \frac{1}{\mu e^b + 1} = \frac{1}{\lambda\mu e^{a+b} - 1}\left(1 - \frac{1}{\lambda e^a + 1} - \frac{1}{\mu e^b + 1}\right). \tag{3.1}$$

Let us take $a = xt$ and $b = (1-t)x$ and multiply both sides of the (3.1) by $4t(1-t)x^2$. We get

$$\begin{aligned}&\frac{2tx}{\lambda e^{tx}+1} \cdot \frac{2(1-t)x}{\mu e^{(1-t)x}+1}\\ &\qquad = 2 \cdot \frac{x}{\lambda\mu e^x - 1}\left(2t(1-t)x - (1-t)\frac{2tx}{\lambda e^{tx}+1} - t\frac{2(1-t)x}{\mu e^{(1-t)x}+1}\right),\end{aligned}$$

By using (1.1) and (1.2), we get

$$\begin{aligned}&\left(\sum_{n=0}^{\infty} \mathcal{G}_n(\lambda)\frac{t^n x^n}{n!}\right)\left(\sum_{n=0}^{\infty} \mathcal{G}_n(\mu)\frac{(1-t)^n x^n}{n!}\right)\\ &= 2\sum_{n=0}^{\infty} \mathcal{B}_n(\lambda\mu)\frac{x^n}{n!} \cdot\\ &\qquad \cdot \left(2t(1-t)x - (1-t)\sum_{n=0}^{\infty} \mathcal{G}_n(\lambda)\frac{t^n x^n}{n!} - t\sum_{n=0}^{\infty} \mathcal{G}_n(\mu)\frac{(1-t)^n x^n}{n!}\right).\end{aligned}$$

Therefore, by applying the Cauchy product and extracting the coefficients of $\frac{x^n}{n!}$, we obtain

$$\begin{aligned}&\sum_{k=0}^{n} \binom{n}{k} \mathcal{G}_k(\lambda)\mathcal{G}_{n-k}(\mu)t^k(1-t)^{n-k}\\ &\qquad = 4t(1-t)n\mathcal{B}_{n-1}(\lambda\mu) - 2(1-t)\sum_{k=0}^{n} \binom{n}{k} \mathcal{B}_k(\lambda\mu)\mathcal{G}_{n-k}(\lambda)t^{n-k}\end{aligned}$$

$$- 2t\sum_{k=0}^{n}\binom{n}{k}\mathcal{B}_k(\lambda\mu)\mathcal{G}_{n-k}(\mu)(1-t)^{n-k}. \tag{3.2}$$

Now we divide (3.2) by $t(1-t)$ and then integrate with respect to t from 0 to 1. By using the facts that $\mathcal{B}_0 = 0$, $B_0 = 1$, and $\mathcal{G}_0 = G_0 = 0$, we obtain the following statement, that is an analogue of the Miki identity (1.4) for the Apostol-Genocchi numbers.

Theorem 3.1. *For all $n \geq 2$,*

$$\sum_{k=1}^{n-1}\frac{\mathcal{G}_k(\lambda)}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} + 2\sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{G}_{n-k}(\lambda)+\mathcal{G}_{n-k}(\mu)}{n-k}$$
$$= 4\mathcal{B}_{n-1}(\lambda\mu) - \frac{2}{n^2}\left(\mathcal{G}_n(\lambda)+\mathcal{G}_n(\mu)\right)\delta_{1,\lambda\mu}.$$

Example 3.2. Let $\lambda = \mu = 1$. Then

$$\sum_{k=1}^{n-1}\frac{G_k}{k}\frac{G_{n-k}}{n-k} + 4\sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{B_k}{k}\frac{G_{n-k}}{n-k} = 4B_{n-1} - \frac{4G_n}{n^2}.$$

Let us suppose now that $n \geq 4$ and n is even. Then, the facts that both odd indexed Bernoulli and Genocchi numbers equal zero imply

$$\sum_{k=2}^{n-2}\frac{G_k}{k}\frac{G_{n-k}}{n-k} + 4\sum_{k=2}^{n-2}\binom{n-1}{k-1}\frac{B_k}{k}\frac{G_{n-k}}{n-k} = -\frac{4G_n}{n^2}.$$

Multiplying both sides of this equation by n and using $\frac{n}{k(n-k)} = \frac{1}{k} + \frac{1}{n-k}$ and $\frac{n}{k}\binom{n-1}{k-1} = \binom{n}{k}$ yield

$$2\sum_{k=2}^{n-2}\frac{G_kG_{n-k}}{n-k} + 4\sum_{k=2}^{n-2}\binom{n}{k}\frac{B_kG_{n-k}}{n-k} = -\frac{4G_n}{n}.$$

By dividing both sides by 2 and replacing the indexes k by $n-k$ and vice versa, we obtain the following analogue of the Miki identity (1.4) for the Genocchi numbers

$$\sum_{k=2}^{n-2}\frac{G_kG_{n-k}}{k} + 2\sum_{k=2}^{n-2}\binom{n}{k}\frac{G_kB_{n-k}}{k} = -\frac{2G_n}{n}.$$

Note that this coincides with [1, Proposition 4.1] for the numbers B'_n, which are defined as $G_n = 2B'_n$.

Corollary 3.3. *Let $\mu \neq 1$. For $n \geq 2$,*

$$\sum_{k=1}^{n-1}\frac{G_k}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} + 2\sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\mu)}{k}\frac{G_{n-k}+\mathcal{G}_{n-k}(\mu)}{n-k} = 4\mathcal{B}_{n-1}(\mu),$$

$$\sum_{k=1}^{n-1} \frac{\mathcal{G}_k(\frac{1}{\mu})}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} + 2\sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{B_k}{k}\frac{\mathcal{G}_{n-k}(\frac{1}{\mu})+\mathcal{G}_{n-k}(\mu)}{n-k}$$
$$= 4B_{n-1} - \frac{2}{n^2}\left(\mathcal{G}_n(\frac{1}{\mu}) + \mathcal{G}_n(\mu)\right).$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, *then*

$$\sum_{k=1}^{n-1} \frac{\mathcal{G}_k(\lambda)}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} + 2\sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{B}_k(\lambda\mu)}{k}\frac{\mathcal{G}_{n-k}(\lambda)+\mathcal{G}_{n-k}(\mu)}{n-k} = 4\mathcal{B}_{n-1}(\lambda\mu).$$

In order to obtain the analogues of the Euler identity, we divide (3.2) by $t(1-t)$ and subsitute $t = 0$.

Theorem 3.4. *For all* $n \geq 2$,

$$\sum_{k=1}^{n-1}\binom{n}{k}\mathcal{B}_k(\lambda\mu)\mathcal{G}_{n-k}(\mu) = n\mathcal{B}_{n-1}(\lambda\mu)(2-\mathcal{G}_1(\lambda)) - \frac{n\mathcal{G}_1(\lambda)\mathcal{G}_{n-1}(\mu)}{2} - \mathcal{G}_n(\mu)\delta_{1,\lambda\mu}.$$

Example 3.5. Let $\lambda = \mu = 1$. Then, since $G_1 = 1$, we obtain

$$\sum_{k=1}^{n-1}\binom{n}{k}B_k G_{n-k} = nB_{n-1} - \frac{n}{2}G_{n-1} - G_n.$$

By using the fact that all odd indexed Bernoulli and Genocchi numbers starting from $n = 3$ disappear, we obtain for all even $n \geq 4$, $\sum_{k=2}^{n-2}\binom{n}{k}B_k G_{n-k} = -G_n$, where the summation is over even indexed numbers (see also [1]).

Here are some identities of the Euler type for the Apostol-Genocchi numbers following from Theorem 3.4.

Corollary 3.6. *Let* $\lambda \neq 1$. *For* $n \geq 2$,

$$\sum_{k=1}^{n-1}\binom{n}{k}\mathcal{B}_k(\lambda)\mathcal{G}_{n-k}(\lambda) = n\mathcal{B}_{n-1}(\lambda) - \frac{n\mathcal{G}_{n-1}(\lambda)}{2},$$
$$\sum_{k=1}^{n-1}\binom{n}{k}\mathcal{B}_k(\lambda)G_{n-k} = n\mathcal{B}_{n-1}(\lambda)(2-\mathcal{G}_1(\lambda)) - \frac{n\mathcal{G}_1(\lambda)G_{n-1}}{2},$$
$$\sum_{k=0}^{n-2}\binom{n}{k}B_k\mathcal{G}_{n-k}(\frac{1}{\lambda}) = -\frac{n\mathcal{G}_1(\lambda)\mathcal{G}_{n-1}(\frac{1}{\lambda})}{2}.$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, *then*

$$\sum_{k=1}^{n-1}\mathcal{B}_k(\lambda\mu)\mathcal{G}_{n-k}(\mu) = n\mathcal{B}_{n-1}(\lambda\mu)(2-\mathcal{G}_1(\lambda)) - \frac{n\mathcal{G}_1(\lambda)\mathcal{G}_{n-1}(\mu)}{2}.$$

Here we used the facts that $B_0 = 1$ and $2 - \mathcal{G}_1(\lambda) = \mathcal{G}_1(\frac{1}{\lambda})$. Another series of the identities of the Miki and the Euler types for the Apostol-Genocchi numbers can be obtained in the same manner, when the following, easily proved, equation

$$\frac{1}{\lambda e^a - 1} \cdot \frac{1}{\mu e^b + 1} = \frac{1}{\lambda\mu e^{a+b} + 1}\left(1 + \frac{1}{\lambda e^a - 1} - \frac{1}{\mu e^b + 1}\right)$$

is taken as a basis for the generating function approach. The following result may be proved in the same way as Theorem 3.1. Let us take $a = xt$ and $b = (1-t)x$ and multiply both sides of the two last identities by $4t(1-t)x^2$. We get

$$\begin{aligned} &2 \cdot \frac{tx}{\lambda e^{tx} - 1} \cdot \frac{2(1-t)x}{\mu e^{(1-t)x} + 1} \\ &\qquad = \frac{2x}{\lambda\mu e^x + 1}\left(2t(1-t)x + 2(1-t)\frac{tx}{\lambda e^{tx} - 1} - t\frac{2(1-t)x}{\mu e^{(1-t)x} + 1}\right). \end{aligned} \tag{3.3}$$

Again, we use (1.1) and (1.2) and apply the Cauchy product in order to extract the coefficients of $\frac{x^n}{n!}$ on both sides of (3.3). Thus, we obtain

$$\begin{aligned} &2\sum_{k=0}^{n}\binom{n}{k}\mathcal{B}_k(\lambda)\mathcal{G}_{n-k}(\mu)t^k(1-t)^{n-k} \\ &\qquad = 2t(1-t)n\mathcal{G}_{n-1}(\lambda\mu) + 2(1-t)\sum_{k=0}^{n}\binom{n}{k}\mathcal{G}_k(\lambda\mu)\mathcal{B}_{n-k}(\lambda)t^{n-k} \\ &\qquad\quad - t\sum_{k=0}^{n}\binom{n}{k}\mathcal{G}_k(\lambda\mu)\mathcal{G}_{n-k}(\mu)(1-t)^{n-k}. \end{aligned} \tag{3.4}$$

Now we divide both equations by $t(1-t)$ and then integrate with respect to t from 0 to 1. By using the facts that $\mathcal{B}_0 = 0$, $B_0 = 1$, and $\mathcal{G}_0 = G_0 = 0$, we obtain the following statement, that is another analogue of the Miki identity for the Apostol-Genocchi numbers.

Theorem 3.7. *For all $n \geq 2$,*

$$\begin{aligned} &\sum_{k=1}^{n-1}\frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{G}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\lambda) - \frac{1}{2}\mathcal{G}_{n-k}(\mu)}{n-k} \\ &\qquad = \mathcal{G}_{n-1}(\lambda\mu) + \frac{\mathcal{G}_n(\mu)}{n}H_{n-1}\delta_{1,\lambda}. \end{aligned} \tag{3.5}$$

Example 3.8. Let $\lambda = \mu = 1$. Then, for all $n \geq 2$,

$$\sum_{k=1}^{n-1}\frac{B_k}{k}\frac{G_{n-k}}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{G_k}{k}\frac{B_{n-k} - \frac{1}{2}G_{n-k}}{n-k} = G_{n-1} + \frac{G_n}{n}H_{n-1}.$$

It is known that the Genocchi and Bernoulli numbers are related as

$$G_n = 2(1-2^n)B_n$$

(see [1]). By substituting this identity into the difference $B_{n-k} - \frac{1}{2}G_{n-k}$ under the second summation, we obtain

$$\sum_{k=1}^{n-1} \frac{B_k}{k}\frac{G_{n-k}}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{G_k}{k}\frac{B_{n-k}-(1-2^{n-k})B_{n-k}}{n-k} = G_{n-1} + \frac{G_n}{n}H_{n-1}.$$

Note that for $n \geq 3$, the odd-indexed Bernoulli and Genocchi numbers disappear, therefore, let us assume now that n is even and $n \geq 4$. Thus, we have

$$\sum_{k=2}^{n-2} \frac{B_k}{k}\frac{G_{n-k}}{n-k} - \sum_{k=2}^{n-2}\binom{n-1}{k-1}\frac{G_k}{k}\frac{2^{n-k}B_{n-k}}{n-k} = \frac{G_n}{n}H_{n-1}.$$

Using the binomial identity $\binom{n-1}{k-1} = \binom{n-1}{n-k}$ leads to

$$\sum_{k=2}^{n-2} \frac{B_k}{k}\frac{G_{n-k}}{n-k} - \sum_{k=2}^{n-2}\binom{n-1}{n-k}\frac{G_k}{k}\frac{2^{n-k}B_{n-k}}{n-k} = \frac{G_n}{n}H_{n-1}.$$

We replace k by $n-k$ under the second summation. Finally, using the notation $G_n = 2B'_n$, proposed in [1], and dividing both sides by 2 lead to the statement (4.2) of [1, Proposition 4.1]

$$\sum_{k=2}^{n-2} \frac{B_k}{k}\frac{B'_{n-k}}{n-k} - \sum_{k=2}^{n-2}\binom{n-1}{k}\frac{2^kB_k}{k}\frac{B'_{n-k}}{n-k} = \frac{B'_n}{n}H_{n-1}.$$

Corollary 3.9. *Let $\mu \neq 1$. For all $n \geq 2$,*

$$\sum_{k=1}^{n-1} \frac{B_k}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{G}_k(\mu)}{k}\frac{B_{n-k}-\frac{1}{2}\mathcal{G}_{n-k}(\mu)}{n-k}$$
$$= \mathcal{G}_{n-1}(\mu) + \frac{\mathcal{G}_n(\mu)}{n}H_{n-1}.$$

Due to the asymmetry of λ and μ in the (3.5), we get the following corollary of the Theorem 3.7.

Corollary 3.10. *Let $\lambda \neq 1$. For all $n \geq 2$,*

$$\sum_{k=1}^{n-1} \frac{\mathcal{B}_k(\lambda)}{k}\frac{G_{n-k}}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{G}_k(\lambda)}{k}\frac{\mathcal{B}_{n-k}(\lambda)-\frac{1}{2}G_{n-k}}{n-k} = \mathcal{G}_{n-1}(\lambda),$$
$$\sum_{k=1}^{n-1} \frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{G}_{n-k}(\frac{1}{\lambda})}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{G_k}{k}\frac{\mathcal{B}_{n-k}(\lambda)-\frac{1}{2}\mathcal{G}_{n-k}(\frac{1}{\lambda})}{n-k} = G_{n-1}. \tag{3.6}$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, then

$$\sum_{k=1}^{n-1} \frac{\mathcal{B}_k(\lambda)}{k}\frac{\mathcal{G}_{n-k}(\mu)}{n-k} - \sum_{k=1}^{n-1}\binom{n-1}{k-1}\frac{\mathcal{G}_k(\lambda\mu)}{k}\frac{\mathcal{B}_{n-k}(\lambda)-\frac{1}{2}\mathcal{G}_{n-k}(\mu)}{n-k} = \mathcal{G}(\lambda\mu).$$

By dividing (3.2) and (3.4) by t and then substituting $t = 0$, we obtain the following analogue of the Euler identity.

Theorem 3.11. *For all $n \geq 2$,*

$$\sum_{k=1}^{n-1} \binom{n}{k} \mathcal{G}_k(\lambda\mu)\mathcal{G}_{n-k}(\mu) = 2n\mathcal{G}_{n-1}(\lambda\mu) + 2(n-1)\mathcal{G}_n(\lambda\mu)\delta_{1,\lambda} + 2n\mathcal{B}_1(\lambda)\left(\mathcal{G}_{n-1}(\lambda\mu) - \mathcal{G}_{n-1}(\mu)\right). \tag{3.7}$$

Example 3.12. Let $\lambda = \mu = 1$. Then

$$\sum_{k=1}^{n-1} \binom{n}{k} G_k G_{n-k} = 2nG_{n-1} + 2(n-1)G_n.$$

By using the fact that all odd indexed Bernoulli and Genocchi numbers starting from $n = 3$ disappear, we obtain a more familiar form for all even $n \geq 4$, $\sum_{k=2}^{n-2} \binom{n}{k} G_k G_{n-k} = 2(n-1)G_n$, where the summation is over even indexed numbers (see also [1]).

Corollary 3.13. *Let $\lambda \neq 1$ and $n \geq 2$. Then the following identities are valid*

$$\sum_{k=1}^{n-1} \binom{n}{k} \mathcal{G}_k(\lambda)\mathcal{G}_{n-k}(\lambda) = 2n\mathcal{G}_{n-1}(\lambda) + 2(n-1)\mathcal{G}_n(\lambda), \tag{3.8}$$

$$\sum_{k=1}^{n-1} \binom{n}{k} \mathcal{G}_k(\lambda)G_{n-k} = 2n\mathcal{G}_{n-1}(\lambda) + 2n\mathcal{B}_{n-1}(\lambda)(\mathcal{G}_{n-1}(\lambda) - G_{n-1}), \tag{3.9}$$

$$\sum_{k=1}^{n-1} \binom{n}{k} G_k\mathcal{G}_{n-k}(\frac{1}{\lambda}) = 2nG_{n-1} + 2n\mathcal{B}_1(\frac{1}{\lambda}) \left(\mathcal{G}_{n-1}(\frac{1}{\lambda}) - G_{n-1}\right). \tag{3.10}$$

Moreover, if $\lambda, \mu, \lambda\mu \neq 1$, then

$$\sum_{k=1}^{n-1} \binom{n}{k} \mathcal{G}_k(\lambda\mu)\mathcal{G}_{n-k}(\mu) = 2n\mathcal{G}_{n-1}(\lambda\mu) + 2n\mathcal{B}_{n-1}(\lambda)(\mathcal{G}_{n-1}(\lambda\mu) - \mathcal{G}_{n-1}(\mu)). \tag{3.11}$$

Proof. Replacing λ and μ in (3.7), and substituting $\mu = 1$ lead to

$$\sum_{k=1}^{n-1} \binom{n}{k} \mathcal{G}_k(\lambda)\mathcal{G}_{n-k}(\lambda) = 2n\mathcal{G}_{n-1}(\lambda) + 2(n-1)\mathcal{G}_n(\lambda) + 2n\left(-\frac{1}{2}\right)\left(\mathcal{G}_{n-1}(\lambda) - \mathcal{G}_{n-1}(\lambda)\right).$$

The last summand equals zero, and we obtain the identiy (3.8). By substituting $\mu = 1$ into (3.7) we obtain (3.9). Substituting $\mu = \frac{1}{\lambda}$ into (2.14) and using the fact that $1 + \mathcal{B}_1(\lambda) = -\mathcal{B}_1(\frac{1}{\lambda})$ lead to (3.10). The second summand on the RH of the (3.7) disappears since $\lambda \neq 1$, and we obtain (3.11). □

Remark 3.14. As it was mentioned above, the classical Bernoulli and Genocchi numbers are connected via the following relationship $G_n = 2(1-2^n)B_n$. It is easy to see that also the Apostol-Bernoulli and Apostol-Genocchi numbers satisfy $\mathcal{G}_n(\lambda) = -2\mathcal{B}_n(-\lambda)$. Moreover, the Apostol-Bernoulli numbers satisfy $\mathcal{B}_{2n}(\lambda) = \mathcal{B}_{2n}(\frac{1}{\lambda})$ and $\mathcal{B}_{2n+1}(\lambda) = -\mathcal{B}_{2n+1}(\frac{1}{\lambda})$ for $\lambda \neq 1$. In the same manner, the Apostol-Genocchi numbers satisfy $\mathcal{G}_{2n}(\lambda) = \mathcal{G}_{2n}(\frac{1}{\lambda})$ and $\mathcal{G}_{2n+1}(\lambda) = -\mathcal{G}_{2n+1}(\frac{1}{\lambda})$ for $n > 0$. These relationships allow to obtain new identities from those considered in the current paper.

Acknowledgement. The research of the first author was supported by the Ministry of Science and Technology, Israel.

References

[1] Agoh, T., On the Miki and Matiyasevich identities for Bernoulli numbers, *Integers* **14** (2014) A17.

[2] Apostol, T.M., On the Lerch zeta function, *Pacific J. Math.* **1**(2) (1951) 161–167.

[3] Bayad, A., Kim, T., Identities for the Bernoulli, the Euler and the Genocchi numbers and polynomials, *Adv. Stud. Cont. Math.* **20**(2) (2010) 247–253.

[4] Chang, C.-H., Ha, C.-W., On recurrence relations for Bernoulli and Euler numbers, *Bull. Austral. Math. Soc.* **64** (2001) 469–474.

[5] Crabb, M.C., The Miki-Gessel Bernoulli number identity, *Glasgow Math. J.* **47** (2005) 327–328.

[6] Dilcher, K., Sums of products of Bernoulli numbers, *J. Number Theory* **60** (1996) 23–41.

[7] Dunne, G.V., Schubert, C., Bernoulli number identities from quantum field theory and topological string theory, *arXiv:math/0406610v2*, 2014.

[8] Gessel, I., On Miki's identity for Bernoulli numbers, *J. Number Theory* **110** (2005) 75–82.

[9] He, Y., Wang, C., Some formulae of products of the Apostol-Bernoulli and Apostol-Euler polynomials, *Discr. Dyn. Nature Soc.*, **2012**, 2012.

[10] Hu, S., Kim, D., Kim, M.-S., New identities involving Bernoulli, Euler and Genocchi numbers, *Adv. Diff. Eq.* **2013** (2013) 74.

[11] Jolany, H., Sharifi, H., Alikelaye, E., Some results for the Apostol-Genocchi polynomials of higher-order, *Bull. Malays. Math. Sci. Soc. (2)* **36**(2) (2013) 465–479.

[12] Kim, T., Rim, S.H., Simsek, Y., Kim, D., On the analogs of Bernoulli and Euler numbers, related identities and zeta and L-functions, *J. Korean Math. Soc.* **45**(2) (2008) 435–453.

[13] Luo, Q.-M., Srivastava, H.M., Some generalizations of the Apostol-Bernoulli and Apostol-Euler polynomials, *J. Math. Anal. Appl.* **308**(1) (2005) 290–302.

[14] Matiyasevich, Y., Identities with Bernoulli numbers, http://logic.pdmi.ras.ru/~yumat/personaljournal/ identitybernoulli/bernulli.htm, 1997.

[15] Miki, H., A relation between Bernoulli numbers, *J. Number Theory* **10** (2978) 297–302.

[16] Pan, H., Sun, Z.-W., New identities involving Bernoulli and Euler polynomials, *arXiv:math/0407363v2*, 2004.

[17] Simsek, Y., Kim, T., Kim, D., A new Kim's type Bernoulli and Euler numbers and related identities and zeta and L-functions, *arXiv:math/0607653v1*, 2006.

Annales Mathematicae et Informaticae
46 *(2016) pp. 115–120*
http://ami.ektf.hu

A note on the derived length of the group of units of group algebras of characteristic two*

Tibor Juhász

Institute of Mathematics and Informatics
Eszterházy Károly University
Eger, Hungary
juhasz.tibor@uni-eszterhazy.hu

Submitted November 25, 2016 — Accepted December 14, 2016

In memoriam Mihály Rados (1941–2016)

Abstract

Denote by FG the group algebra of a group G over a field F, by $U(FG)$ its group of units, and by $\mathrm{dl}(U(FG))$ the derived length of $U(FG)$. We know very little about $\mathrm{dl}(U(FG))$, especially when F has characteristic 2. In this short note, it is shown that, if F is of characteristic 2, G' is cyclic of order 2^n and the nilpotency class of G is less than $n+1$, then $\mathrm{dl}(U(FG))$ is equal to n or $n+1$. In addition, if $n > 1$ and $G' = \mathrm{Syl}_2(G)$, then $\mathrm{dl}(U(FG)) = n$.

Keywords: Group ring, group of units, derived length

MSC: 16S34, 16U60, 20C07, 20F14

1. Introduction

Let FG be the group algebra of a group G over a field F of prime characteristic p, and let $U(FG)$ be the group of units of FG. It is determined in [4] when $U(FG)$ is solvable, however, we know very little about the derived length of $U(FG)$.

*This research was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP 4.2.4. A/2-11-1-2012-0001 'National Excellence Program'

Assume first that p is an odd prime. For this case, the group algebras FG with metabelian group of units are classified in [16], under restriction G is finite, and this result is extended to torsion G in [6]. In [7, 8] the finite groups G are described, such that $U(FG)$ has derived length 3. According to [1], if G is a finite p-group with cyclic commutator subgroup, then $\mathrm{dl}(U(FG)) = \lceil \log_2(|G'|+1) \rceil$, where $\lceil \cdot \rceil$ is the upper integer part function. The aim of [2] and [10] is to extend this result, and determine the value of $\mathrm{dl}(U(FG))$ for arbitrary groups G with G' is a cyclic p-group, where p is still an odd prime. As it turned out, if G is nilpotent and torsion, then the derived length of $U(FG)$ remains $\lceil \log_2(|G'|+1) \rceil$, but for non-nilpotent or non-torsion G it can be different. However, the description is not complete yet, for the open cases we refer the reader to [10].

For $p = 2$ and finite group G, necessary and sufficient conditions for $U(FG)$ to be metabelian is given in [9], and independently, in [14]. This result is extended in [6] as follows: if F is a field of characteristic 2, and G is a nilpotent torsion group, then $U(FG)$ is metabelian exactly when G' is a central elementary abelian group of order dividing 4. In [13], it is established that if G is a group of maximal class of order 2^n, then $\mathrm{dl}(U(FG))$ is less or equal to $n-1$. To the best of the author's knowledge, for $p = 2$ there is no other result concerning the derived length of $U(FG)$. The aim of this paper to draw the attention to this uncovered area by sharing the author's experience and an introductory result.

The group of units of a group algebra can be investigated via the Lie structure of the group algebra. For example, we can obtain an upper bound on the derived length of $U(FG)$, by the help of the strong Lie derived length of FG. Let $\delta^{(0)}(FG) = FG$, and for $i \geq 1$, denote by $\delta^{(i)}(FG)$ the associative ideal generated by all the Lie commutators $[x, y] = xy - yx$ with $x, y \in \delta^{(i-1)}(FG)$. FG is said to be strongly Lie solvable, if there exists i, for which $\delta^{(i)}(FG) = 0$, and the first such i is called the strong Lie derived length of FG, which will be denoted by $\mathrm{dl}^L(FG)$. For $x, y \in U(FG)$ we have that the group commutator $(x, y) = x^{-1}y^{-1}xy$ is equal to $1 + x^{-1}y^{-1}[x, y]$, which implies that $\delta_i(U(FG)) \subseteq 1 + \delta^{(i)}(FG)$ for all i, where $\delta_i(U(FG))$ denotes the ith term of the derived series of $U(FG)$. Therefore, if FG is strongly Lie solvable, then $\mathrm{dl}(U(FG)) \leq \mathrm{dl}^L(FG)$.

According to [15, Theorem 5.1], FG is strongly Lie solvable if and only if either G is abelian, or G' is a finite p-group and F is a field of characteristic p. By [11, Proposition 1], if FG is strongly Lie solvable such that G is nilpotent and $\gamma_3(G) \subseteq (G')^p$, then $\mathrm{dl}^L(FG) = \lceil \log_2(t(G')+1) \rceil$, where by $t(G')$ we mean the nilpotency index of the augmentation ideal of the subalgebra FG'.

Assume now that G is a group with cyclic commutator subgroup of order 2^n and F is a field of characteristic 2. Then G is nilpotent with nilpotency class $\mathrm{cl}(G) \leq n+1$, so we can apply the above formulas to get

$$\mathrm{dl}(U(FG)) \leq \mathrm{dl}^L(FG) = \lceil \log_2(2^n+1) \rceil = n+1.$$

Hence, if $n = 1$, then $\mathrm{dl}(U(FG)) = 2$. For the case when $n > 1$ and $\mathrm{cl}(G) \leq n$, we are able to give a lower bound on $\mathrm{dl}(U(FG))$ as well.

Theorem 1.1. *Let F be a field of characteristic 2, and let G be a group with cyclic commutator subgroup of order 2^n, where $n > 1$. Then $\mathrm{dl}(U(FG)) \geq n$, whenever G has nilpotency class at most n.*

According to [12, Theorem 1], under conditions of Theorem 1.1, $U(FG)$ is nilpotent and, by [5, Theorem 4.3], if $G' = \mathrm{Syl}_2(G)$, then $\mathrm{cl}(U(FG)) = 2^n - 1$. Using the well-known relation $\delta_i(U(FG)) \subseteq \gamma_{2^i}(U(FG))$ between terms of the derived series and the lower central series of groups, we have the following assertion.

Corollary 1.2. *Let F be a field of characteristic 2, and let G be a group with cyclic commutator subgroup of order 2^n, where $n > 1$. If $G' = \mathrm{Syl}_2(G)$ and $\mathrm{cl}(G) \leq n$, then $\mathrm{dl}(U(FG)) = n$.*

For instance, if

$$G = \langle a, b, c \mid c^{2^n} = 1, b^{-1}ab = ac, ac = ca, bc = cb \rangle,$$

with $n > 1$, and $\mathrm{char}(F) = 2$, then $\mathrm{dl}(U(FG)) = n$. This example also witnesses that for non-torsion G, $U(FG)$ can be metabelian, even if G' is cyclic of order 4.

The GAP system for computational discrete algebra [17] and its package, the LAGUNA [3] open the door to compute the derived length of $U(FG)$ for G of not too large size. Computing $\mathrm{dl}(U(FG))$ for some group G of order not greater than 512 and F of 2 elements, it seems that $\mathrm{dl}(U(FG))$ will always be at least n, even if $\mathrm{cl}(G) = n + 1$. However, it would also be interesting to know when $\mathrm{dl}(U(FG))$ is n or when it is $n + 1$.

2. Proof of Theorem 1.1

We will use the following notations. For a normal subgroup H of G we denote by $\mathfrak{I}(H)$ the ideal in FG generated by all elements of the form $h - 1$ with $h \in H$. For the subsets $X, Y \subseteq FG$ by $[X, Y]$ we mean the additive subgroup of FG generated by all Lie commutators $[x, y]$ with $x \in X$ and $y \in Y$.

Write $G' = \langle x \mid x^{2^n} = 1 \rangle$, and assume that $n > 1$. Then for any $m > 1$, $y \in \gamma_m(G)$ and $g \in G$ we have $g^{-1}yg = y^k$, where k is odd, thus $(y, g) = y^{k-1} \in \gamma_m(G)^2$. Hence, $\gamma_{m+1}(G) \subseteq \gamma_m(G)^2$ for all $m > 1$, so G is nilpotent of class at most $n + 1$. Evidently, if $\gamma_3(G) \subseteq (G')^4$, then $\mathrm{cl}(G)$ cannot exceed n. We show first the converse, that is, if $\mathrm{cl}(G) \leq n$, then

$$\gamma_3(G) \subseteq (G')^4. \tag{2.1}$$

This is clear, if $n = 2$. For $n \geq 3$, it is well known that the automorphism group of G' is the direct product of the cyclic group $\langle \alpha \rangle$ of order 2 and the cyclic group $\langle \beta \rangle$ of order 2^{n-2}, where the action of these automorphisms on G' is given by $\alpha(x) = x^{-1}$, $\beta(x) = x^5$. Consequently, for every $g \in G$ there exists $i \geq 0$, such that either $g^{-1}xg = x^{5^i}$ or $g^{-1}xg = x^{-5^i}$. Assume that there is a $g \in G$ such that $g^{-1}xg = x^{-5^i}$ for some i, and let $y \in \gamma_m(G)$ with $m > 1$. Then $(y, g) = y^{-1-5^i} \in$

$\gamma_{m+1}(G)$, and as $-1-5^i \equiv 2 \pmod 4$, we have that $\gamma_{m+1}(G) = (\gamma_m(G))^2$. This means that $\mathrm{cl}(G) = n+1$, which is a contradiction. Therefore, for any $g \in G$ there exists i such that $g^{-1}xg = x^{5^i}$ and $(x,g) = x^{-1+5^i} = x^{4k}$ for some integer k, which forces 2.1.

Let F be a field of characteristic 2. The next step is to show by induction that

$$[\omega(FG')^m, FG] \subseteq \Im(G')^{m+3} \tag{2.2}$$

for all $m \geq 1$. Let $y \in G'$ and $g \in G$. Then, using that $\gamma_3(G) \subseteq (G')^4$, we have

$$[y+1, g] = [y,g] = gy((y,g)+1) \in \Im(\gamma_3(G)) \subseteq \Im(G')^4.$$

Since the Lie commutators of the form $[y+1,g]$ span the subspace $[\omega(FG'), FG]$, (2.2) holds for $m=1$. Assume now (2.2) for some $m \geq 1$. Then,

$$\begin{aligned}[\omega(FG')^{m+1}, FG] &\subseteq \omega(FG')^m[\omega(FG'), FG] + [\omega(FG')^m, FG]\omega(FG') \\ &\subseteq \Im(G')^{m+4},\end{aligned}$$

as desired. Furthermore, by using (2.2), for all $k,l \geq 1$ we have

$$\begin{aligned}&[\Im(G')^k, \Im(G')^l] \\ &\quad = [FG\omega(FG')^k, FG\omega(FG')^l] \\ &\quad \subseteq FG[\omega(FG')^k, FG\omega(FG')^l] + [FG, FG\omega(FG')^l]\omega(FG')^k \\ &\quad \subseteq FG[\omega(FG')^k, FG]\omega(FG')^l + FG[FG, \omega(FG')^l]\omega(FG')^k \\ &\qquad + [FG, FG]\omega(FG')^{k+l} \\ &\quad \subseteq \Im(G')^{k+l+1}.\end{aligned} \tag{2.3}$$

At this stage, it may be worth mentioning that without the assumption $\mathrm{cl}(G) \leq n$ we can only claim that $\gamma_3(G) \subseteq (G')^2$ and $[\omega(FG')^m, FG] \subseteq \omega(FG')^{m+1}$ instead of (2.1) and (2.2). Although those would be enough for (2.3), but not for what follows.

Denote by S the set of those $a \in G$, for which there exists $b \in G$, such that $\langle (a,b) \rangle = G'$. We are going to show that for all $k \geq 1$ and $a \in S$, there exists $w_k \in \Im(G')^{3\cdot 2^{k-1}}$, such that

$$1 + a(x+1)^{3\cdot 2^{k-1}-1} + w_k \in \delta_k(U(FG)). \tag{2.4}$$

This implies that $\delta_k(U(FG))$ contains non-identity element, while $3\cdot 2^{k-1} - 1 < 2^n$, and then

$$\mathrm{dl}(U(FG)) \geq \left\lceil \log_2\left(\frac{2}{3}(2^n+1)\right)\right\rceil = n,$$

and the proof of Theorem 1.1 will be done.

Let $a \in S$. Then there exists $b \in G$ such that $(a,b) = x^i$, where i is odd. By (2.2), $[x+1, b] \in \Im(G')^4$, and

$$u := (1 + a(x+1), b) = 1 + (1+a(x+1))^{-1}b^{-1}[a(x+1), b]$$

$$\equiv 1 + (1 + a(x+1))^{-1} b^{-1} [a, b](x+1)$$
$$\equiv 1 + (1 + a(x+1))^{-1} a(x^i + 1)(x+1) \pmod{\mathfrak{I}(G')^3}.$$

Since $1 + a(x+1)$ belongs to the normal subgroup $1 + \mathfrak{I}(G')$, so does its inverse, and

$$u \equiv 1 + a(x^i + 1)(x+1) \pmod{\mathfrak{I}(G')^3}.$$

Using that $x^i + 1 \equiv i(x+1) = x+1 \pmod{\omega(FG')^2}$, we obtain that

$$u \equiv 1 + a(x+1)^2 \pmod{\mathfrak{I}(G')^3},$$

and (2.4) is confirmed for $k = 1$. Assume, by induction, the truth of (2.4) for some $k \geq 1$, and let $a \in S$. Then there exists $b \in G$ such that $\langle (a, b) \rangle = G'$, and of course, b also belongs to S. Moreover, $b^{-1}a \in S$, because $(b^{-1}a, b) = (a, b)$. By the inductive hypothesis, there exist $w_k, w'_k \in \mathfrak{I}(G')^{3 \cdot 2^{k-1}}$, such that

$$u := 1 + b^{-1}a(x+1)^{3 \cdot 2^{k-1} - 1} + w_k \in \delta_k(U(FG))$$

and

$$v := 1 + b(x+1)^{3 \cdot 2^{k-1} - 1} + w'_k \in \delta_k(U(FG)).$$

According to (2.3),

$$[u, v] \equiv [b^{-1}a(x+1)^{3 \cdot 2^{k-1} - 1}, b(x+1)^{3 \cdot 2^{k-1} - 1}] \pmod{\mathfrak{I}(G')^{3 \cdot 2^k}}.$$

Applying (2.2), we have that $[(x+1)^{3 \cdot 2^{k-1} - 1}, b]$ and $[b^{-1}a, (x+1)^{3 \cdot 2^{k-1} - 1}]$ belong to $\mathfrak{I}(G')^{3 \cdot 2^{k-1} + 2}$, and

$$\begin{aligned} [u, v] &\equiv b^{-1}a[(x+1)^{3 \cdot 2^{k-1} - 1}, b](x+1)^{3 \cdot 2^{k-1} - 1} \\ &\quad + b[b^{-1}a, (x+1)^{3 \cdot 2^{k-1} - 1}](x+1)^{3 \cdot 2^{k-1} - 1} + [b^{-1}a, b](x+1)^{3 \cdot 2^k - 2} \\ &\equiv a(x^i + 1)(x+1)^{3 \cdot 2^k - 2} \equiv a(x+1)^{3 \cdot 2^k - 1} \pmod{\mathfrak{I}(G')^{3 \cdot 2^k}}, \end{aligned}$$

where i is not divisible by 2. Since $u^{-1}, v^{-1} \in 1 + \mathfrak{I}(G')$, so

$$(u, v) = 1 + u^{-1}v^{-1}[u, v] \equiv 1 + a(x+1)^{3 \cdot 2^k - 1} \pmod{\mathfrak{I}(G')^{3 \cdot 2^k}}$$

and the induction is done.

Acknowledgements. The author would like to express his sincere thanks to Alexander Konovalov for the help in creating the GAP script, furthermore to my colleague, Tibor Tajti for providing a capable technical background to run it.

References

[1] Bagiński, C., A note on the derived length of the unit group of a modular group algebra, *Comm. Algebra* 30 (2002), 4905–4913.

[2] Balogh, Z., Li, Y., On the derived length of the group of units of a group algebra, *J. Algebra Appl.* 6 (2007), No. 6, 991–999.

[3] Bovdi, V., Konovalov, A., Rossmanith, R., Schneider, C., LAGUNA — Lie AlGebras and UNits of group Algebras, Version 3.7.0; 11 November 2014 (`http://www.cs.st-andrews.ac.uk/~alexk/laguna/`).

[4] Bovdi, A., Group algebras with a solvable group of units, *Comm. Algebra* 36 (2008), No. 2, 315–324.

[5] Bovdi, A., Kurdics, J., Lie properties of group algebra and the nilpotency class of the group of units, *J. Algebra* 212 (1999), No. 1, 28–64.

[6] Catino, F., Spinelli, E., On the derived length of the unit group of a group algebra, *J. Group Theory* 13 (2010), No. 4, 577–588.

[7] Chandra, H., Sahai, M., On group algebras with unit groups of derived length three in characteristic three, *Publ. Math.* 82 (2013), No. 3–4, 697–708.

[8] Chandra, H., Sahai, M., Group algebras with unit groups of derived length three, *J. Algebra Appl.* 9 (2010), No. 2, 305–314.

[9] Coleman, D.B., Sandling, R., Mod 2 group algebras with metabelian unit groups. *J. Pure Appl. Algebra* 131 (1998), 25–36.

[10] Juhász, T., The derived length of the unit group of a group algebra – The case $G' = \mathrm{Syl}_p(G)$, *J. Algebra Appl.*, In press, DOI: `http://dx.doi.org/10.1142/S0219498817501420`.

[11] Juhász, T., On the derived length of Lie solvable group algebras, *Publ. Math.* 68 (2006), No. 1–2, 243–256.

[12] Khripta, I.I., The nilpotency of the multiplicative group of a group ring, *Mat. Zametki* 11 (1972), 191–200.

[13] Konovalov, A., Wreath products in the unit group of modular group algebras of 2-groups of maximal class, PhD Thesis, Mathematical Institute, Ukrainian National Academy of Sciences, Kiev, 1995.

[14] Kurdics, J., On group algebras with metabelian unit groups, *Period. Math. Hung.* 32 (1996), No. 1–2, 57–64.

[15] Sehgal, S.K., Topics in group rings, Marcel Dekker, New York, 1978.

[16] Shalev, A., Meta-abelian unit groups of group algebras are usually abelian, *J. Pure Appl. Algebra* 72 (1991), 295–302.

[17] The GAP Group, GAP – Groups, Algorithms, and Programming, Version 4.8.6; 2016 (`http://www.gap-system.org`).

Annales Mathematicae et Informaticae
46 *(2016) pp. 121–142*
http://ami.ektf.hu

Maintainability of classes in terms of bug prediction

Gergely Ladányi

University of Szeged
Department of Software Engineering
Hungary
lgergely@inf.u-szeged.hu

Submitted October 27, 2015 — Accepted April 18, 2016

Abstract

Measuring software product maintainability is a central issue in software engineering which led to a number of different practical quality models. Besides system level assessments it is also desirable that these models provide technical quality information at source code element level (e.g. classes, methods) to aid the improvement of the software. Although many existing models give an ordered list of source code elements that should be improved, it is unclear how these elements are affected by other important quality indicators of the system, e.g. bug density.

In this paper we empirically investigate the bug prediction capabilities of the class level maintainability measures of our ColumbusQM probabilistic quality model using open-access PROMSIE bug dataset. We show that in terms of correctness and completeness, ColumbusQM competes with statistical and machine learning prediction models especially trained on the bug data using product metrics as predictors. This is a great achievement in the light of that our model needs no training and its purpose is different (e.g. to estimate testability, or development costs) than those of the bug prediction models.

Keywords: ISO/IEC 25010, ColumbusQM, Software maintainability, Bug, Bug prediction, Class level maintainability, PROMISE

1. Introduction

Maintainability is probably one of the most attractive, observed and evaluated quality characteristics of software products. The importance of maintainability lies in its direct connection with many factors that influence the overall costs of a software system, for example the effort needed for new developments [1], mean time between failures of the system [2], bug fixing time [3], or operational costs [4]. After the appearance of the ISO/IEC 9126 standard [5] for software product quality, the development of new practical models which measure the maintainability of systems in a direct way has exploded [6, 7, 8, 9, 10].

Although these models provide a system level overview about the maintainability of a software which is a valuable information in itself for e.g. making decisions, backing up intuition, or assessing risks, just a portion of them provide low-level, actionable information for developers (i.e. list of source code elements and quality attributes that should be improved). Current approaches usually just enumerate the most complex methods, most coupled classes or other source code elements that carry some source code metric value. There is a lack of empirical evidences that these elements are indeed the most critical from maintainability point of view and changing them will improve some of the quality factors related to maintainability.

We used our earlier results to calculate maintainability on system and lower levels. The ColumbusQM model introduced in our previous work [10] is able to calculate maintainability on systems level. Later we extended the ColumbusQM with the drill-down approach [11] to calculate maintainability on lower levels as well (classes, methods, etc.). The drill-down approach calculates a so-called *relative maintainability index (RMI)* for each source code element which measures the extent to which they affect the overall system maintainability. The RMI is a small number that is either positive when it improves the overall rating or negative when it decreases the system level maintainability. We also developed a web based graphical user interface called QualityGate [12] to continuously monitor the maintainability of a software using version control systems.

The contribution of this study is the comparison of the RMI based ordering of classes with widely used statistical and machine learning prediction models, e.g. decision trees, neural networks, or regression. The performance of the RMI based ordering of classes proves to be competitive compared to these prediction techniques.

The paper is organized as follows. Section 2 presents the work related to ours. The data collection and analysis methodology is introduced in Section 3, while the analysis results are described in Section 4. Finally, we list the threats to validity in Section 5, and conclude the paper in Section 6.

2. Related work

In this section will give an overview about the related papers dealing with software quality measurement and fault prediction.

Both of software fault detection [13, 14] and software quality models [15, 16] date back to the 70's and evolving since then [17, 18]. Although the software quality measurement has become popular recently by the release of quality standards like ISO/IEC 9126 [5] and its successor the ISO/IEC 25010 [19]. Using the definition of the characteristics and subcharacteristics reseachers developed several software quality models which are able to measure the software quality of a system, but only a few works on lower class or method level. Even fewer works investigate empirically the relation of the maintainability and other factors, like bug density at finer levels.

Heitlager et al. [6] presented a bottom-up approach to measure software maintainability. They split the basic metric values into five categories from poor to excellent using threshold values [20]. Then they aggregated these qualifications for for higher properties, such as maintainability. Bijlsma et al. [21] examined the correlation of the SIG model rating with four maintainability related factors: Time, Throughput, Productivity, Efficiency. They found that their model has a strong predictive power for the maintenance burden that is associated to the system. Our maintainability model is in many aspects similar to the SIG model; however, we use a probabilistic approach for aggregation opposed to the threshold based approach, and also generate a list of source code elements with the highest risk that should be improved first. Moreover, we investigate the relation of maintainability and bugs at lower level rather than system level where immediate actions can be taken.

The technical debt based models like SQALE [8] or SQUALE [7] introduce low-level rules to connect the ISO/IEC 9126 characteristics with metrics. These rules refer to different properties of the source code (e.g. the comment ratio should be above 25%) and violating them has a reparation cost. These models provide a list of critical elements simply by ordering them based on their total reparation costs. Although it assures the biggest system level maintainability increase there is no guarantee that one corrects the most critical elements (e.g. elements with the most bugs, or elements that are used by many other components).

Chulani et. al. [22] introduced the Orthogonal Defect Classification Constructive Quality Model (ODC COQUALMO) as an extension of the Constructive Cost Model (COCOMO) [23, 24]. The model was calibrated with empirical defect distributions and it contains two sub-models. Defect introduction sub-model predicts the number of defects that will appear in each defect category. The defect removal model produces an estimate of the number of defects that will be removed from these categories. The idea behind this and our study is similar, the main difference is that we did not add a bug predictor sub-model to our quality model but we investigated the connection between the defects and the final aggregated value of the model.

Chawla [25] proposed the SQMMA (Software Quality Model for Maintainability Analysis) approach based on the ISO/IEC 25010 standard which provides comprehensive formulas to calculate the Maintainablity and its subcharacteristic. They normalized the average metric values respect to the first selected release of Tomcat and they aggregated towards higer nodes using weighted sum. They also compared the number of buggy files with the system-level quality measurements through four versions of the Tomcat. They observed that the pattern of Maintainability consistently matches (in reverse) with the number of buggy files in the system. We also compared the maintainability with the bugs in a system, but we worked on the level of classes instead of the system.

Papers related to software defects very often use databases with information about bugs and different metrics about the source code elements. Zimmermann et al. [26] has published a bug database for Eclipse and used the data for predicting defects by logistic regression and complexity metrics. Moser et al. [27] annotated this dataset with change metrics and compared its bug prediction ability with code metrics. We also used code metrics and a bug database, but we examined the connection between the number of bugs and the RMI value of classes rather than source code metrics directly. Moreover, the bug dataset used by Moser has become part of the PROMISE [28] database, but we could not consider these data as it provides bug information only for packages and Java files, while we need bug data on the level of classes.

Jureczko et al. [29] describes an analysis that was conducted on newly collected repository with 92 versions of 38 proprietary, open-source and academic projects. The dataset is part of the PROMISE dataset and part of the our study as well since it provides bug information on class level. To study the problem of cross project defect prediction they performed clustering on software projects in order to identify groups of software projects with similar characteristic from the defect prediction point of view. The conducted analysis reveals that there exist clusters from the defect prediction point of view, and two of those clusters were successfully identified. Later Madeyski et al. [30] used the dataset to empirically investigate how process metrics can significantly improve defect prediction.

There are also other works relying on the PROMISE dataset. Menzies et al. aim to comparatively evaluate local versus global lessons learned [31] for effort estimation and defect prediction. They applied automated clustering tools to effort and defect datasets from the PROMISE repository and rule learners generated lessons learned from all the data. The work of Wang and Yao [32] deals with improving the bug prediction models by handling imbalanced training data and uses PROMISE dataset to validate the approach.

Vasilescu et. al. [33] aggregated the SLOC class level metric to package level using various aggregation techniques (Theil, Gini, Kolm, Atkinson, indices, sum, mean, median). They found that the choice of the aggregation technique does influence the correlation of the aggregated values and the number of defects. Contrary to them we did not aggregate class level metrics to package level metrics, but we aggregated them to a maintainability index using the ColumbusQM quality model weighed by experts and compared the results with the number of bugs in the classes.

3. Methodology

We started the empirical examination of the bugs in the aspects of maintainability by collecting bug datasets. In the literature the most widely used bug dataset is the PROMISE dataset [34]. It contains bug information on class level for various open-source and proprietary systems. Since for the maintainability calculation the source code is necessary we only used the open-source systems from the dataset. In order to decrease the possible bias in the machine learning prediction models we filter out the very small systems (i.e. systems with fewer than 6 classes) and those having very high ratio of buggy classes (i.e. over 75% of the classes contain bugs). At the end of the process we collected source code and bug information for each 30 versions of the 16 open-source systems. For each Java class found in these 30 versions we calculate the class level RMI value (relative maintainability index) according to our drill-down approach [11].

3.1. The applied quality model

First we calculated the absolute maintainability values for the different versions of the systems. We used the ColumbusQM, our probabilistic software quality model [10] that is able to measure the quality characteristics defined by the ISO/IEC 25010 standard. The computation of the high-level quality characteristics is based on a directed acyclic graph (see Figure 1) whose nodes correspond to quality properties that can either be internal (low-level) or external (high-level). Internal quality properties characterize the software product from an internal (developer) view and are usually estimated by using source code metrics. External quality properties characterize the software product from an external (end user) view and are usually aggregated somehow by using internal and other external quality properties. The nodes representing internal quality properties are called *sensor nodes* as they measure internal quality directly (white nodes in Figure 1). The other nodes are called *aggregate nodes* as they acquire their measures through aggregation of the lower-level nodes. In addition to the aggregate nodes defined by the standard (black nodes) we introduced new ones (light gray nodes) and kept those of contained only in the old standard (dark gray nodes).

The description of the different quality attributes can be found in Table 1.

Dependencies between an internal and an external, or two external properties are represent by the edges of the graph. The aim is to evaluate all the external quality properties by performing an aggregation along the edges of the graph, called Attribute Dependency Graph (ADG). We calculate a so called *goodness value* (from the [0,1] interval) to each node in the ADG that expresses how good or bad (1 is the best) is the system regarding that quality attribute. The probabilistic statistical aggregation algorithm uses a benchmark as the basis of the qualification, which is a source code metric repository database with 100 open-source and industrial software systems.

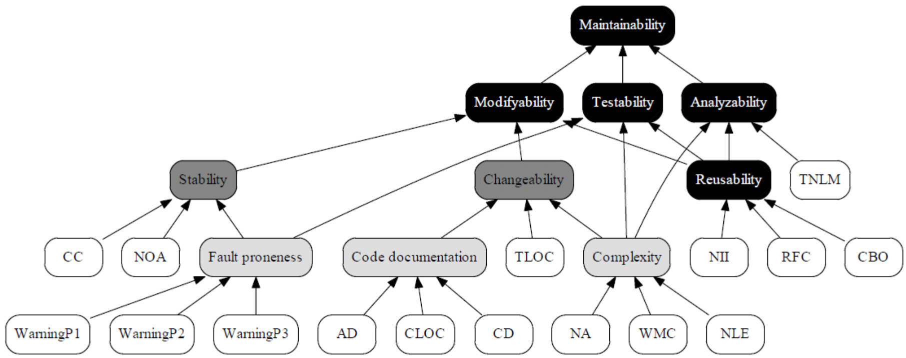


Figure 1: ColumbusQM – Java ADG [35]

3.2. The drill-down approach

The above approach is used to obtain a system-level measure for source code maintainability. Our aim is to drill down to lower levels in the source code and to get a similar measure for the building blocks of the code base (e.g. classes or methods). For this, we defined the *relative maintainability index*[1] *(RMI)* for the source code elements [11], which measures the extent to which they affect the system level goodness values. The basic idea is to calculate the system level goodness values, leaving out the source code elements one-by-one. After a particular source code element is left out, the system level goodness values will change slightly for each node in the ADG. The difference between the original goodness value computed for the whole system and the goodness value computed without the particular source code element is called the *relative maintainability index* of the source code element itself. The RMI is a small number that is either positive when it improves the overall rating or negative when it decreases the system level maintainability. The absolute value of the index measures the extent of the influence to the overall system level maintainability. In addition, a relative index can be computed for each node of the ADG, meaning that source code elements can affect various quality aspects in different ways and to different extents.

More details and the validation of the approach can be found in our previous paper [11].

3.3. Comparison of ColumbusQM and prediction models

One way to look at the maintainability scores is that they rank classes according to their level of maintainability. Nonetheless, if we assume that classes with the worst maintainability scores contain most of the bugs we can easily turn RMI into

[1] We use the terms relative maintainability index, relative maintainability score, and RMI interchangeably throughout the paper. Moreover, we may also refer to them by omitting the word "relative" for simplicity reasons.

Table 1: The low-level quality properties of our model [35]

Sensor nodes	
CC	Clone coverage. The percentage of copied and pasted source code parts, computed for the classes of the system.
NOA	Number of Ancestors. Number of classes, interfaces, enums and annotations from which the class is directly or indirectly inherited.
WarningP1	The number of critical rule violations in the class.
WarningP2	The number of major rule violations in the class.
WarningP3	The number of minor rule violations in the class.
AD	Api Documentation. Ratio of the number of documented public methods in the class.
CLOC	Comment Lines of Code. Number of comment and documentation code lines of the class.
CD	Comment Density. The ratio of comment lines compared to the sum of its comment and logical lines of code.
TLOC	Total Lines of Code. Number of code lines of the class, including empty and comment lines.
NA	Number of attributes in the class.
WMC	Weighted Methods per Class. Complexity of the class expressed as the number of linearly independent control flow paths in it. It is calculated as the sum of the McCabe's Cyclomatic Complexity (McCC) values of its local methods and init blocks.
NLE	Nesting Level Else-If. Complexity of the class expressed as the depth of the maximum embeddedness of its conditional and iteration block scopes, where in the if-else-if construct only the first if instruction is considered.
NII	Number of Incoming Invocations. Number of other methods and attribute initializations, which directly call the local methods of the class.
RFC	Response set For Class. Number of local (i.e. not inherited) methods in the class plus the number of directly invoked other methods by its methods or attribute initializations.
TNLM	Total Number of Local Methods. Number of local (i.e. not inherited) methods in the class, including the local methods of its nested, anonymous, and local classes.
CBO	Coupling Between Object classes. Number of directly used other classes (e.g. by inheritance, function call, type reference, attribute reference).

a simple classification method. For that we should define "classes with the worst maintainability scores" more precisely. To be able to compare the RMI based classification to other prediction models, we simply use the natural RMI threshold of 0, i.e. our simple model classifies all the classes as buggy which have negative maintainability scores, and non-buggy all the rest.

Now we are ready to compare the ColumbusQM based ordering (i.e. classification with the above extension) to other well-known statistical and machine learning prediction models. We examine two types of models, three classification methods: J48 decision tree algorithm, neural network model, and a logistic regression based algorithm. Additionally, we apply three regression techniques that differ from the above classifiers in that they assign a real number (the predicted number of bugs in the class) to each class instead of predicting only whether it is buggy or not. We consider the RepTree decision tree based regression algorithm, linear regression, and neural network based regression. In case of regression algorithms we say that the algorithm predicts a class as buggy if it predicts more than 0.5 bugs for it, because above this number the class will more likely be buggy than not buggy.

These models need a training phase to build a model for prediction. We chose to put all the classes from the different versions of the systems together and train these algorithms on this huge dataset using 10-fold cross validation. We allow them to use all the available source code metrics which the Columbus static analyzer tool called *SourceMeter* provides – not only those used by ColumbusQM – as predictors. After the training we run the prediction on each of the 30 separate versions of the 13 systems. For building the prediction models we use the Weka tool [36].

We used Spearman's rank correlation coefficient to measure the strength of the similarity between the orderings of machine learning models and RMI. We also evaluate the performance of the prediction models and our maintainability model in terms of the classical measures of *precision* and *recall*. In addition, we also calculate the *completeness* value [37], which measures the number of bugs (faults) in classes classified fault-prone, divided by the total number of faults in the system. This number differs from the usual recall value as it measures the percentage of faults – and not only the faulty classes – that has been found by the prediction model.

Each model predicts classes as either fault-prone or not fault-prone, so the classification is binary (in case of regression models we make the prediction binary as described above). The definition of the performance measures used in this work are as follows:

$$\text{Precision} := \frac{\#\text{ classes correctly classified as buggy}}{\#\text{ total classes classified as buggy}}$$

$$\text{Recall} := \frac{\#\text{ classes correctly classified as buggy}}{\#\text{ total buggy classes in the system}}$$

$$\text{Completeness} := \frac{\#\text{ bugs in classes classified buggy}}{\#\text{ total bugs in the system}}$$

To be able to directly compare the results of different models we also calculate the *F-measure*, which is the harmonic mean of precision and recall:

$$F - \text{measure} := 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

As an additional aggregate measure we define the $\dot{F}$*-measure* to be the harmonic mean of precision and completeness:

$$\dot{F} - \text{measure} := 2 \cdot \frac{\text{precision} \cdot \text{completeness}}{\text{precision} + \text{completeness}}$$

4. Results

The empirical analysis is performed on 30 releases of 13 different open-source systems that take up to 2M lines of source code. The bug data for these systems is available in the PROMISE [28] online bug repository which we used for collecting the bug numbers at class level. For each version of the subject systems we calculated the system level quality according to Section 3.1 and all the relative maintainability scores for classes as described in Section 3.2 using ColumbusQM [10].

Table 2: Descriptive statistics of the analyzed systems [35]

System	Nr. of classes	Nr. of bugs	Buggy classes
ant-1.3	115	33	20
ant-1.4	163	45	38
ant-1.5	266	35	32
ant-1.6	319	183	91
ant-1.7	681	337	165
camel-1.0	295	11	10
camel-1.2	506	484	191
camel-1.4	724	312	134
camel-1.6	795	440	170
ivy-1.4	209	17	15
ivy-2.0	294	53	37
jedit-3.2	255	380	89
jedit-4.0	288	226	75
jedit-4.1	295	215	78
jedit-4.2	344	106	48
jedit-4.3	439	12	11
log4j-1.0	118	60	33
log4j-1.1	100	84	35
lucene-2.0	180	261	87
pbeans-2.0	37	16	8
poi-2.0	289	39	37
synapse-1.0	139	20	15
synapse-1.1	197	96	57
synapse-1.2	228	143	84
tomcat-6.0	732	114	77
velocity-1.6	189	161	66
xalan-2.4	634	154	108
xalan-2.6	816	605	395
xerces-1.2	291	61	43
xerces-1.3	302	186	65
Average	**341.33**	**162.97**	**77.13**

Table 2 shows some basic descriptive statistics about the analyzed systems. The second column contains the total number of classes in the systems (that we could successfully map) while the third column shows the total number of bugs. The fourth column presents the number of classes containing at least one bug.

4.1. Prediction Model Results

Recall, that to compare our model to other prediction models we put all the classes together and let the machine learning algorithms to build prediction models based on all the available product metrics using 10-fold cross validation. We built three binary classifiers based on J48 decision tree, logistic regression, and neural network; and three regression based prediction models using RepTree decision tree, linear regression, and neural network.

Classification algorithms. First, we present the results of the comparison of classifiers and maintainability score. For each of the classifiers we ran the classification on the classes of the 30 versions of the subject systems. As described in Section 3, we also classified the same classes based on our maintainability scores. Finally, we calculated the precision, recall, and completeness measures (for the definitions of these measures see Section 3) for the different predictions for each subject system.

Table 3 and 4 lists all the precision, recall, and completeness values for the four different methods and for all the 30 versions of the 13 subject systems. Although the results are varying for the different systems, in general we can say that the precisions of the three learning based models are higher than that of the RMI based model. The average precision values are 0.68, 0.59, 0.68, and 0.35 for J48, logistic regression (LR), neural network (NN), and RMI, respectively. Nonetheless, in terms of recall and completeness especially, RMI looks superior to the other methods. The average completeness values are 0.38, 0.24, 0.31, and 0.81 for J48, logistic regression, neural network, and RMI, respectively. Figure 2 shows these values on a bar-chart together with the F-measure (harmonic mean of the precision and recall) and $\dot{F}$-measure (harmonic mean of the precision and completeness values). It is easy to see that according to both the F-measure and $\dot{F}$-measure, RMI and J48 methods are performing the best. But while J48 achieves this with very high precision and an average recall and completeness, RMI has far the highest recall and completeness values combined with moderate precision. Another important observation is that for every method the average recall values are smaller than the completeness values which suggests that the bug distribution among the classes of the projects is fairly uniform.

RMI performs the worst (in terms of precision) in cases of *camel v1.0* and *jedit v4.3*. Exactly these two systems are those where the number of bugs per class are the smallest (11/295 and 12/439 respectively). This biases not only our method but the other algorithms, too. They achieve the lowest precision on these two systems.

There is a block of precision values close to 1 for systems from *log4j v1.0* to *pbeans v2.0* for the three learning based models. However, their completeness measure is very low (and recall is even lower). On the contrary, RMI has a very

Table 3: Comparison of the precision, recall, and completeness of different models

System	J48 Prec.	J48 Rec.	J48 Comp.	LR Prec.	LR Rec.	LR Comp.
ant-1.3	0.82	0.45	0.39	0.67	0.10	0.12
ant-1.4	0.45	0.13	0.13	0.00	0.00	0.00
ant-1.5	0.52	0.47	0.49	0.22	0.06	0.09
ant-1.6	0.76	0.41	0.57	0.74	0.15	0.25
ant-1.7	0.68	0.38	0.53	0.80	0.21	0.38
camel-1.0	0.20	0.10	0.09	0.50	0.10	0.09
camel-1.2	1.00	0.14	0.27	0.89	0.04	0.13
camel-1.4	0.75	0.20	0.25	0.70	0.05	0.15
camel-1.6	0.68	0.11	0.29	0.58	0.06	0.15
ivy-1.4	0.30	0.20	0.29	0.50	0.20	0.29
ivy-2.0	0.65	0.35	0.43	0.58	0.30	0.38
jedit-3.2	0.80	0.27	0.57	0.75	0.17	0.45
jedit-4.0	0.74	0.33	0.60	0.76	0.25	0.54
jedit-4.1	0.89	0.44	0.64	0.91	0.27	0.48
jedit-4.2	0.63	0.56	0.71	0.56	0.38	0.57
jedit-4.3	0.13	0.55	0.58	0.09	0.36	0.33
log4j-1.0	1.00	0.12	0.13	1.00	0.06	0.07
log4j-1.1	1.00	0.09	0.12	0.50	0.03	0.07
lucene-2.0	1.00	0.15	0.25	1.00	0.07	0.18
pbeans-2	0.75	0.38	0.56	1.00	0.13	0.19
poi-2.0	0.50	0.19	0.21	0.47	0.19	0.21
synapse-1.0	0.80	0.27	0.35	0.00	0.00	0.00
synapse-1.1	0.80	0.14	0.19	0.00	0.00	0.00
synapse-1.2	0.82	0.17	0.24	1.00	0.05	0.07
tomcat-1	0.47	0.40	0.46	0.48	0.36	0.48
velocity-1.6	0.85	0.26	0.39	0.67	0.12	0.20
xalan-2.4	0.56	0.50	0.55	0.50	0.31	0.38
xalan-2.6	0.89	0.53	0.58	0.94	0.33	0.40
xerces-1.2	0.38	0.19	0.25	0.35	0.16	0.23
xerces-1.3	0.58	0.23	0.42	0.58	0.22	0.40
Average	**0.68**	**0.29**	**0.38**	**0.59**	**0.16**	**0.24**

high recall and completeness in these cases (over 0.79 and 0.83, respectively) and still having acceptably high precision values (above 0.5).

For *synapse v1.0* and *v1.1* logistic regression and for *ant v1.3* and *v1.4* neural network achieves 0 precision, recall and completeness. Again, RMI based prediction achieves a very high completeness and recall with a moderate, but still acceptable level of precision in these cases, too.

We note again, that while our maintainability model uses only 16 metrics we let the learning algorithms to use all the 59 available product metrics which we calculated. If we restricted the set of predictors for the classifier algorithms to only those that ColumbusQM uses, we got somewhat different results. Figure 3 shows the average values of the resulting precision, recall, completeness, F-measure, and $\dot{F}$-measure. In this case, the precision of the learning algorithms dropped a bit while recall and completeness levels remained. The F-measure of the RMI is higher than that of any other method while in terms of $\dot{F}$-measure J48 and RMI perform the best.

We empirically investigated that RMI competes with the classification based algorithms. With Spearman's rank correlation we measured the strength of the similarity between the oredering of the RMI and the machine learning models.

Table 4: Comparison of the precision, recall, and completeness of different models

System	NN Prec.	NN Rec.	NN Comp.	RMI Prec.	RMI Rec.	RMI Comp.
ant-1.3	0.00	0.00	0.00	0.27	0.90	0.91
ant-1.4	0.00	0.00	0.00	0.27	0.61	0.64
ant-1.5	0.50	0.28	0.31	0.18	0.78	0.80
ant-1.6	0.77	0.30	0.44	0.46	0.85	0.91
ant-1.7	0.72	0.33	0.51	0.38	0.82	0.88
camel-1.0	0.20	0.10	0.09	0.05	0.90	0.91
camel-1.2	0.97	0.19	0.26	0.41	0.60	0.77
camel-1.4	0.68	0.11	0.24	0.26	0.80	0.87
camel-1.6	0.75	0.11	0.24	0.26	0.73	0.86
ivy-1.4	0.40	0.13	0.24	0.15	1.00	1.00
ivy-2.0	0.86	0.16	0.21	0.25	0.86	0.91
jedit-3.2	0.87	0.15	0.44	0.49	0.69	0.86
jedit-4.0	0.86	0.25	0.55	0.40	0.77	0.91
jedit-4.1	0.91	0.27	0.48	0.39	0.78	0.85
jedit-4.2	0.71	0.42	0.60	0.23	0.88	0.93
jedit-4.3	0.20	0.73	0.75	0.03	0.64	0.67
log4j-1.0	1.00	0.03	0.15	0.58	0.79	0.83
log4j-1.1	1.00	0.03	0.11	0.68	0.80	0.86
lucene-2.0	1.00	0.08	0.24	0.69	0.68	0.81
pbeans-2	1.00	0.25	0.50	0.55	0.75	0.81
poi-2.0	0.45	0.14	0.15	0.19	0.65	0.64
synapse-1.0	0.50	0.07	0.05	0.15	0.87	0.85
synapse-1.1	1.00	0.09	0.11	0.35	0.75	0.79
synapse-1.2	1.00	0.05	0.09	0.49	0.77	0.80
tomcat-1	0.56	0.36	0.47	0.22	0.84	0.89
velocity-1.6	0.91	0.15	0.30	0.56	0.74	0.80
xalan-2.4	0.49	0.34	0.36	0.34	0.69	0.74
xalan-2.6	0.89	0.47	0.54	0.59	0.52	0.61
xerces-1.2	0.45	0.23	0.31	0.24	0.47	0.52
xerces-1.3	0.76	0.29	0.47	0.30	0.43	0.59
Average	**0.68**	**0.20**	**0.31**	**0.35**	**0.74**	**0.81**

The measurements for the classication algorithms are in Table 5. If the machine learning models used only the quality model metrics the correlation in average is strong (0.669) for the Logistic Regression and moderately strong for the J48 (0.504) and the Neural Network (0.463). As it was expected if the machine learning models were able to use other metrics the correlation became lower because of the difference between the used metrics.

Regression based algorithms. Next, we analyze the comparison results of the regression based algorithms and RMI. Recall, that regression algorithms provide a continuous function for predicting the number of bugs instead of only classifying classes as buggy or non-buggy. According to the method described in Section 3, we consider a class as buggy in this case if the appropriate regression model predicts at least 0.5 bugs for it. The detailed results of the precision, recall, and completeness values are shown in Table 6 and 7.

In this case the overall picture is somewhat different. The regression based methods work more similarly to RMI meaning that they achieve higher recall and completeness in return for lower precision. This can also be observed in Figure 4. All the bars on the chart are similarly distributed as in case of RMI. Neural network

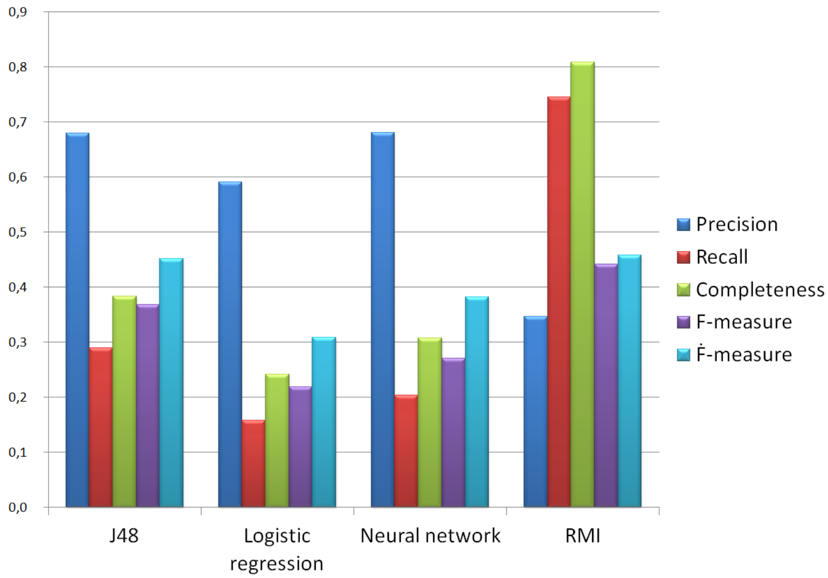


Figure 2: Average performance of methods using all metrics

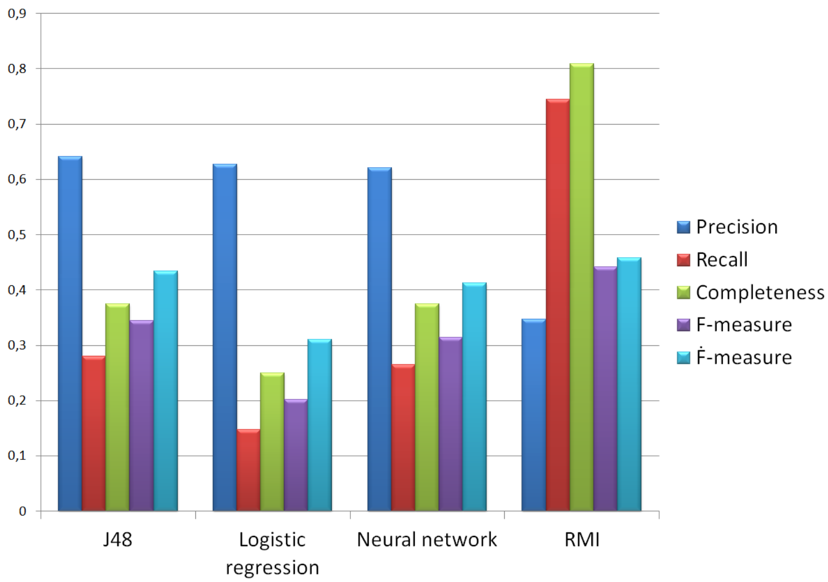


Figure 3: Average performance of methods using only the 16 metrics used by ColumbusQM

Table 5: Spearman's rank correlation between the ranking of RMI and the classification algorithms

System	Only QM			All		
	J48	LR	NN	J48	LR	NN
ant-1.3.csv	0.503	0.687	0.587	0.252	0.361	0.241
ant-1.4.csv	0.409	0.544	0.428	0.192	0.215	0.084
ant-1.5.csv	0.553	0.699	0.566	0.348	0.421	0.166
ant-1.6.csv	0.541	0.71	0.577	0.334	0.376	0.118
ant-1.7.csv	0.622	0.82	0.641	0.359	0.515	0.253
camel-1.0.csv	0.56	0.75	0.369	0.177	0.298	0.12
camel-1.2.csv	0.521	0.753	0.411	0.258	0.425	0.107
camel-1.4.csv	0.588	0.821	0.42	0.316	0.482	0.147
camel-1.6.csv	0.607	0.864	0.446	0.176	0.552	0.222
ivy-1.4.csv	0.517	0.666	0.474	0.326	0.435	0.276
ivy-2.0.csv	0.507	0.624	0.479	0.349	0.425	0.311
jedit-3.2.csv	0.194	0.239	0.035	0.131	0.077	0.102
jedit-4.0.csv	0.343	0.437	0.207	0.277	0.27	0.267
jedit-4.1.csv	0.413	0.624	0.325	0.31	0.367	0.238
jedit-4.2.csv	0.449	0.551	0.294	0.243	0.25	0.182
jedit-4.3.csv	0.386	0.567	0.392	0.366	0.302	0.135
log4j-1.0.csv	0.588	0.682	0.468	0.422	0.534	0.426
log4j-1.1.csv	0.82	0.854	0.751	0.567	0.726	0.528
lucene-2.0.csv	0.438	0.513	0.41	0.292	0.329	0.272
pbeans-2.csv	0.534	0.526	0.364	0.356	0.33	0.286
poi-2.0.csv	0.578	0.804	0.729	0.492	0.653	0.342
synapse-1.0.csv	0.454	0.725	0.401	0.379	0.307	0.338
synapse-1.1.csv	0.585	0.777	0.489	0.478	0.487	0.467
synapse-1.2.csv	0.642	0.799	0.52	0.566	0.496	0.483
tomcat-1.csv	0.44	0.611	0.502	0.324	0.368	0.16
velocity-1.6.csv	0.542	0.687	0.548	0.432	0.527	0.398
xalan-2.4.csv	0.497	0.747	0.585	0.394	0.591	0.377
xalan-2.6.csv	0.103	0.521	0.335	0.075	0.399	0.256
xerces-1.2.csv	0.58	0.687	0.521	0.423	0.564	0.32
xerces-1.3.csv	0.618	0.769	0.609	0.491	0.588	0.365
Average	**0.504**	**0.669**	**0.463**	**0.337**	**0.422**	**0.266**

achieves the highest precision but far the lowest recall and completeness. The main feature of RMI remained unchanged, namely it has far the highest recall and completeness in average. In terms of the F-measures, RepTree, linear regression and RMI perform almost identically. Figure 5 shows the performance measures where the regression models used only the 16 metrics used by ColumbusQM. Contrary to the classifier algorithms, this caused no remarkable change in this case.

We also measuered the Spearman's rank correlation coefficient between the regression based machine learning algorithms and the RMI. The measurements can be found in Table 8. In this case the linear regression algorithm produce the most similar ranking in average comparing to the RMI (0.683). If the machine learning models were able to use all of the metrics the similarity between the rankins became lower in this case as well. Moreover it also interesting that in average the regression based algorithms could achieve better correlation comparing to the classification algorithms. This is probably because the classification algorithms many times predicted with 0.0 or 1.0 probability which makes it harder to make a more balanced ordering of the classes.

It is clear from the presented data that if one seeks for an algorithm that predicts buggy classes with few false positive hits RMI based method is not the optimal

Table 6: Comparison of the precision, recall, and completeness of different regression models

System	RT Prec.	RT Rec.	RT Comp.	LR Prec.	LR Rec.	LR Comp.
ant-1.3	0.34	0.70	0.70	0.48	0.65	0.61
ant-1.4	0.29	0.42	0.49	0.28	0.34	0.38
ant-1.5	0.27	0.66	0.69	0.25	0.59	0.60
ant-1.6	0.66	0.73	0.84	0.61	0.68	0.79
ant-1.7	0.54	0.63	0.76	0.54	0.62	0.76
camel-1.0	0.12	0.30	0.27	0.11	0.50	0.45
camel-1.2	0.63	0.17	0.33	0.69	0.37	0.57
camel-1.4	0.44	0.27	0.45	0.40	0.51	0.68
camel-1.6	0.39	0.22	0.42	0.30	0.37	0.62
ivy-1.4	0.21	0.60	0.65	0.24	0.67	0.71
ivy-2.0	0.36	0.70	0.77	0.30	0.76	0.79
jedit-3.2	0.64	0.57	0.83	0.64	0.66	0.87
jedit-4.0	0.57	0.68	0.86	0.54	0.69	0.88
jedit-4.1	0.55	0.69	0.81	0.56	0.77	0.87
jedit-4.2	0.31	0.81	0.91	0.31	0.90	0.95
jedit-4.3	0.04	0.64	0.67	0.04	0.64	0.67
log4j-1.0	0.77	0.30	0.45	0.88	0.42	0.55
log4j-1.1	0.89	0.49	0.67	0.86	0.34	0.55
lucene-2.0	0.83	0.34	0.55	0.79	0.47	0.63
pbeans-2	0.71	0.63	0.69	0.60	0.38	0.56
poi-2.0	0.21	0.54	0.54	0.20	0.46	0.46
synapse-1.0	0.35	0.53	0.60	0.25	0.27	0.35
synapse-1.1	0.56	0.42	0.54	0.63	0.33	0.47
synapse-1.2	0.67	0.44	0.50	0.74	0.38	0.46
tomcat-1	0.25	0.78	0.84	0.26	0.69	0.78
velocity-1.6	0.58	0.44	0.58	0.56	0.55	0.70
xalan-2.4	0.41	0.73	0.77	0.38	0.70	0.75
xalan-2.6	0.74	0.58	0.66	0.80	0.54	0.63
xerces-1.2	0.22	0.47	0.52	0.25	0.44	0.51
xerces-1.3	0.33	0.49	0.61	0.27	0.34	0.53
Average	**0.46**	**0.53**	**0.63**	**0.46**	**0.53**	**0.64**

solution. But the purpose of RMI is clearly not that. It strives for highlighting the most problematic classes from maintainability point of view and for giving an ordering among classes in which they should be improved. In this respect, it is an additional extra that it outperforms the pure bug prediction algorithms without any learning phase. Moreover, RMI based prediction is superior in terms of completeness, which is the primary target when improving the code – to catch more bugs with less resources. Adding this to the fact that typically less than half of the classes gets a negative RMI, we can say that it is a practically useful method. To summarize, RMI has lower but still acceptable level of precision and very high recall and completeness compared to the different learning algorithms resulting in competitive performance in terms of the F-measures.

5. Threats to validity

There are some threats to the validity of our study results. First of all, the correctness of the bug data contained in the PROMISE dataset is taken for granted. However, there might be errors in the collected bug data that could compromise

Table 7: Comparison of the precision, recall, and completeness of different regression models

System	NN Prec.	NN Rec.	NN Comp.	RMI Prec.	RMI Rec.	RMI Comp.
ant-1.3	0.30	0.15	0.15	0.27	0.90	0.91
ant-1.4	0.41	0.18	0.22	0.27	0.61	0.64
ant-1.5	0.34	0.44	0.46	0.18	0.78	0.80
ant-1.6	0.63	0.36	0.54	0.46	0.85	0.91
ant-1.7	0.67	0.44	0.60	0.38	0.82	0.88
camel-1.0	0.08	0.10	0.09	0.05	0.90	0.91
camel-1.2	0.84	0.11	0.25	0.41	0.60	0.77
camel-1.4	0.59	0.16	0.33	0.26	0.80	0.87
camel-1.6	0.55	0.16	0.35	0.26	0.73	0.86
ivy-1.4	0.17	0.13	0.24	0.15	1.00	1.00
ivy-2.0	0.39	0.32	0.45	0.25	0.86	0.91
jedit-3.2	0.81	0.33	0.66	0.49	0.69	0.86
jedit-4.0	0.69	0.41	0.68	0.40	0.77	0.91
jedit-4.1	0.72	0.46	0.65	0.39	0.78	0.85
jedit-4.2	0.49	0.63	0.78	0.23	0.88	0.93
jedit-4.3	0.06	0.45	0.50	0.03	0.64	0.67
log4j-1.0	1.00	0.12	0.25	0.58	0.79	0.83
log4j-1.1	0.75	0.09	0.21	0.68	0.80	0.86
lucene-2.0	0.93	0.16	0.39	0.69	0.68	0.81
pbeans-2	1.00	0.25	0.50	0.55	0.75	0.81
poi-2.0	0.38	0.16	0.18	0.19	0.65	0.64
synapse-1.0	0.00	0.00	0.00	0.15	0.87	0.85
synapse-1.1	1.00	0.07	0.10	0.35	0.75	0.79
synapse-1.2	1.00	0.10	0.11	0.49	0.77	0.80
tomcat-1	0.38	0.47	0.59	0.22	0.84	0.89
velocity-1.6	0.59	0.15	0.32	0.56	0.74	0.80
xalan-2.4	0.45	0.43	0.47	0.34	0.69	0.74
xalan-2.6	0.87	0.28	0.40	0.59	0.52	0.61
xerces-1.2	0.36	0.28	0.34	0.24	0.47	0.52
xerces-1.3	0.51	0.28	0.48	0.30	0.43	0.59
Average	**0.46**	**0.53**	**0.63**	**0.46**	**0.53**	**0.64**

our study. But the probability of this is negligible, and there are many other works in the literature that relies on PROMISE dataset similarly to ours.

Another problem is that it is hard to generalize the results as we studied only 30 versions of 13 Java open-source systems from the PROMISE bug repository. There are other open-access bug datasets that could also be examined to support the generality of the current findings for other programming languages as well.

Only about 25% of the classes in the PROMISE repository contain bugs, therefore the training data we used for the machine learning models is somewhat imbalanced. Compensating this effect itself is a subject of many research efforts. However, we think that for a first comparison of the performance of our maintainability score based model to other prediction algorithms this level of imbalance is acceptable (i.e. does not bias the learning algorithms significantly).

We found inconsistencies between the bug databases and the downloaded source code of the systems. Some of the classes were missing either from the bug data or from the source code of some projects. In such cases we simply left out these classes from the further analysis. Even though the proportion of these classes was very small, it is certainly a threat to validity, but we think its effect is negligible.

Finally, as we used machine learning bug prediction models the chosen algo-

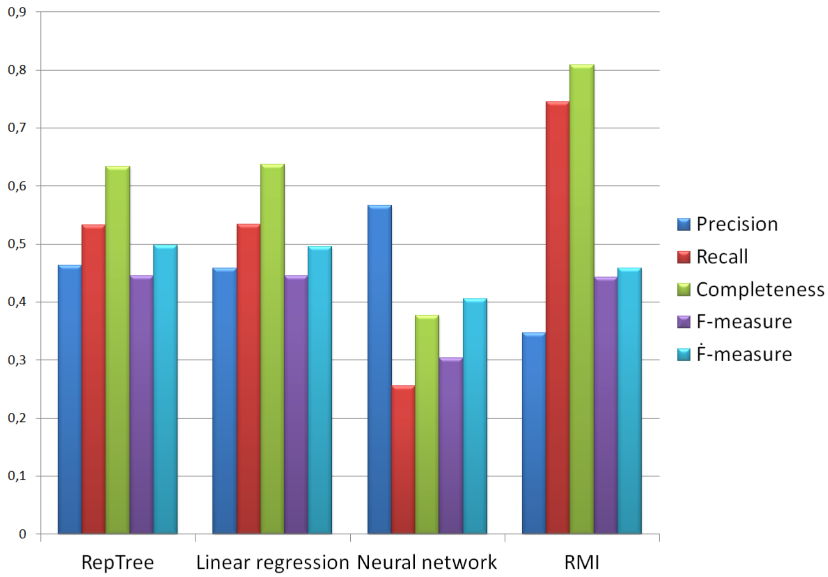


Figure 4: Average performance of regression methods using all metrics

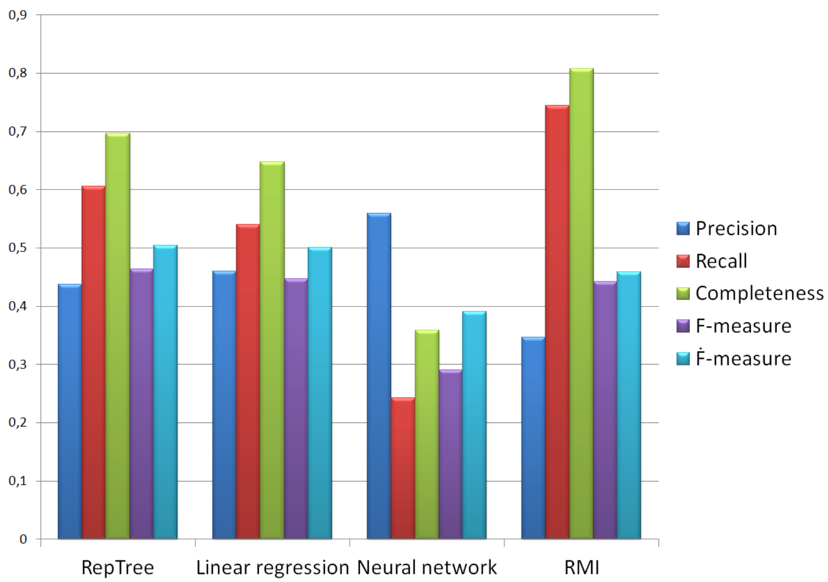


Figure 5: Average performance of regression using only the 16 metrics used by ColumbusQM

Table 8: Correlation between the ranking of RMI and the regression algorithms

System	Only QM			All		
	RT	LR	NN	RT	LR	NN
ant-1.3.csv	0.596	0.708	0.639	0.57	0.533	0.485
ant-1.4.csv	0.553	0.585	0.527	0.475	0.363	0.289
ant-1.5.csv	0.597	0.678	0.629	0.576	0.519	0.425
ant-1.6.csv	0.638	0.713	0.638	0.604	0.52	0.429
ant-1.7.csv	0.713	0.793	0.717	0.682	0.67	0.558
camel-1.0.csv	0.545	0.773	0.564	0.399	0.522	0.224
camel-1.2.csv	0.634	0.776	0.571	0.434	0.629	0.359
camel-1.4.csv	0.698	0.834	0.614	0.551	0.713	0.447
camel-1.6.csv	0.759	0.879	0.643	0.416	0.765	0.54
ivy-1.4.csv	0.524	0.651	0.529	0.471	0.563	0.371
ivy-2.0.csv	0.548	0.61	0.547	0.47	0.498	0.435
jedit-3.2.csv	0.197	0.246	0.214	0.172	0.139	0.136
jedit-4.0.csv	0.335	0.413	0.401	0.356	0.309	0.269
jedit-4.1.csv	0.489	0.585	0.493	0.523	0.452	0.343
jedit-4.2.csv	0.45	0.547	0.476	0.496	0.386	0.337
jedit-4.3.csv	0.489	0.592	0.53	0.526	0.422	0.39
log4j-1.0.csv	0.588	0.642	0.448	0.511	0.562	0.405
log4j-1.1.csv	0.784	0.869	0.652	0.747	0.712	0.493
lucene-2.0.csv	0.485	0.558	0.471	0.421	0.408	0.432
pbeans-2.csv	0.405	0.61	0.436	0.445	0.429	0.203
poi-2.0.csv	0.618	0.826	0.725	0.767	0.752	0.564
synapse-1.0.csv	0.588	0.749	0.618	0.581	0.47	0.261
synapse-1.1.csv	0.627	0.797	0.698	0.674	0.655	0.424
synapse-1.2.csv	0.704	0.811	0.731	0.721	0.677	0.396
tomcat-1.csv	0.537	0.624	0.501	0.509	0.434	0.37
velocity-1.6.csv	0.576	0.756	0.664	0.526	0.63	0.452
xalan-2.4.csv	0.639	0.715	0.613	0.722	0.697	0.672
xalan-2.6.csv	0.382	0.526	0.52	0.446	0.524	0.468
xerces-1.2.csv	0.599	0.773	0.285	0.692	0.603	0.585
xerces-1.3.csv	0.62	0.859	0.333	0.741	0.573	0.578
Average	**0.564**	**0.683**	**0.548**	**0.541**	**0.538**	**0.411**

rithms and tuning of their parameters are also important. During the comparison process we tried to choose the most well-known regression and classification algorithms and we used their default parameters set by Weka. Probably there are better bug prediction models than those we used, but our goal was not to find the best one but to generally compare our relative maintainability index to the bug prediction ability of the well-known machine learning algorithms.

6. Conclusions

In this paper we examined the connection of the maintainability scores of Java classes calculated by our ColumbusQM quality model and their fault-proneness (i.e. the number of bugs the classes contain). As there is only a small number of quality models providing output at source element level, currently there is a lack of research dealing with this topic.

However, the primary target of quality models (including ours) in general is not bug prediction, but it is important to investigate the usefulness of these models in practice. Nonetheless, to get a picture about how useful this feature of the model is, we compared its bug prediction capability to other well-known statistical and machine learning algorithms. The results show that if the two model uses the same metrics the Spearman's rank correlation between the predicted values of the models is strong or moderately strong. If the machine learning models were able to use other metrics as well the correlation became weaker. The empirical investigation showed that there are different balances between the precision and recall of the different methods, but in overall (i.e. according to the F-measure of the prediction performance) our model is clearly competitive with the other approaches. While typical classifier algorithms tend to have higher precision but lower recall, our quality model based prediction has far the highest recall with an acceptable level of precision. What is even more, completeness, which expresses the number of detected bugs compared to the total number of all bugs, is also the best among all algorithms.

We stress that the main result here is that a general maintainability model like ColumbusQM is able to draw the attention to the classes containing this large amount of bugs independent of the analyzed system (i.e. without any training on data). This property ensures that the source code elements one starts to improve contain the largest amount of bugs while high precision would only mean that the classes one consider contain bugs for sure (but not necessarily in large amount).

Acknowledgements. I would like to thank Péter Hegedűs, Rudolf Ferenc, István Siket and Tibor Gyimóthy for their help in this research work and their useful advices.

References

[1] Krishnamoorthy Srinivasan and Douglas Fisher. Machine Learning Approaches to Estimating Software Development Effort. *IEEE Transactions on Software Engineering*, 21(2):126–137, 1995.

[2] Yennun Huang, Chandra Kintala, Nick Kolettis, and N Dudley Fulton. Software Rejuvenation: Analysis, Module and Applications. In *25th International Symposium on Fault-Tolerant Computing, 1995. FTCS-25.*, pages 381–390. IEEE, 1995.

[3] Cathrin Weiss, Rahul Premraj, Thomas Zimmermann, and Andreas Zeller. How Long will it Take to Fix This Bug? In *Proceedings of the 4th International Workshop on Mining Software Repositories*, May 2007.

[4] John D. Musa. Operational Profiles in Software-Reliability Engineering. *IEEE Softw.*, 10(2):14–32, March 1993.

[5] ISO/IEC. *ISO/IEC 9126. Software Engineering – Product quality 6.5.* ISO/IEC, 2001.

[6] I. Heitlager, T. Kuipers, and J. Visser. A Practical Model for Measuring Maintainability. *Proceedings of the 6th International Conference on Quality of Information and Communications Technology*, pages 30–39, 2007.

[7] Karine Mordal-Manet, Francoise Balmas, Simon Denier, Stephane Ducasse, Harald Wertz, Jannik Laval, Fabrice Bellingard, and Philippe Vaillergues. The SQUALE

Model – A Practice-based Industrial Quality Model. In *Proceedings of the 25rd International Conference on Software Maintenance (ICSM 2009)*, pages 531–534. IEEE Computer Society, 2009.

[8] J. L. Letouzey and T. Coq. The SQALE Analysis Model: An Analysis Model Compliant with the Representation Condition for Assessing the Quality of Software Source Code. In *2010 2nd International Conference on Advances in System Testing and Validation Lifecycle (VALID)*, pages 43–48. IEEE, August 2010.

[9] Stefan Wagner, Klaus Lochmann, Lars Heinemann, Michael Kläs, Adam Trendowicz, Reinhold Plösch, Andreas Seidl, Andreas Goeb, and Jonathan Streit. The Quamoco Product Quality Modelling and Assessment Approach. In *Proceedings of the 2012 International Conference on Software Engineering*, ICSE 2012, pages 1133–1142, Piscataway, NJ, USA, 2012. IEEE Press.

[10] T. Bakota, P. Hegedűs, P. Körtvélyesi, R. Ferenc, and T. Gyimóthy. A Probabilistic Software Quality Model. In *Proceedings of the 27th IEEE International Conference on Software Maintenance (ICSM 2011)*, pages 368–377, Williamsburg, VA, USA, 2011. IEEE Computer Society.

[11] Péter Hegedűs, Tibor Bakota, Gergely Ladányi, Csaba Faragó, and Rudolf Ferenc. A Drill-Down Approach for Measuring Maintainability at Source Code Element Level. *Electronic Communications of the EASST*, 60, 2013.

[12] T. Bakota, P. Hegedűs, I. Siket, G. Ladányi, and R. Ferenc. Qualitygate sourceaudit: A tool for assessing the technical quality of software. In *Software Maintenance, Reengineering and Reverse Engineering (CSMR-WCRE), 2014 Software Evolution Week - IEEE Conference on*, pages 440–445, Feb 2014.

[13] B. Randell. System structure for software fault tolerance. *Software Engineering, IEEE Transactions on*, SE-1(2):220–232, June 1975.

[14] James J. Horning, Hugh C. Lauer, P. M. Melliar-Smith, and Brian Randell. A program structure for error detection and recovery. In *Operating Systems, Proceedings of an International Symposium*, pages 171–187, London, UK, UK, 1974. Springer-Verlag.

[15] Hans Kaspar Myron Lipow Gordon J. Macleod Michael J. Merrit Barry W. Boehm, John R. Brown. *Characteristics of Software Quality(TRW series of software technology)*. Elsevier Science Ltd;, North-Holland, 1978.

[16] J.A. McCall, P.K. Richards, and G.F. Walters. *Factors in Software Quality. Volume I: Concepts and Definitions of Software Quality*. AD A049. General Electric, 1977.

[17] Romi Satria Wahono. A systematic literature review of software defect prediction: Research trends, datasets, methods and frameworks. *Journal of Software Engineering*, 1(1):1–16, 2015.

[18] Ronald Jabangwe, Jürgen Börstler, Darja Šmite, and Claes Wohlin. Empirical evidence on the link between object-oriented measures and external quality attributes: a systematic literature review. *Empirical Software Engineering*, 20(3):640–693, 2014.

[19] ISO/IEC. *ISO/IEC 25010:2011. Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models*. ISO/IEC, 2011.

[20] Tiago L. Alves, Christiaan Ypma, and Joost Visser. Deriving Metric Thresholds from Benchmark Data. In *Proceedings of the 26th IEEE International Conference on Software Maintenance (ICSM2010)*, 2010.

[21] Dennis Bijlsma, Miguel Alexandre Ferreira, Bart Luijten, and Joost Visser. Faster Issue Resolution with Higher Technical Quality of Software. *Software Quality Control*, 20(2):265–285, June 2012.

[22] Raymond Madachy and Barry Boehm. Assessing Quality Processes with ODC COQUALMO. In *Making Globally Distributed Software Development a Success Story*, pages 198–209. Springer, 2008.

[23] Barry Boehm, Bradford Clark, Ellis Horowitz, Chris Westland, Ray Madachy, and Richard Selby. Cost Models for Future Software Life Cycle Processes: COCOMO 2.0. *Annals of Software Engineering*, 1(1):57–94, 1995.

[24] Stefan Wagner, Andreas Goeb, Lars Heinemann, and Michael Kl˙ Operationalised product quality models and assessment: The quamoco approach. *Information and Software Technology*, 62:101 – 123, 2015.

[25] Mandeep K. Chawla and Indu Chhabra. Sqmma: Software quality model for maintainability analysis. In *Proceedings of the 8th Annual ACM India Conference*, Compute '15, pages 9–17, New York, NY, USA, 2015. ACM.

[26] Thomas Zimmermann, Rahul Premraj, and Andreas Zeller. Predicting defects for eclipse. In *Proceedings of the 3rd International Workshop on Predictor Models in Software Engineering*, PROMISE '07, Washington, DC, USA, 2007. IEEE Computer Society.

[27] Raimund Moser, Witold Pedrycz, and Giancarlo Succi. A Comparative Analysis of the Efficiency of Change Metrics and Static Code Attributes for Defect Prediction. In *Proceedings of the 30th International Conference on Software Engineering (ICSE '08)*, pages 181–190, New York, NY, USA, 2008. ACM.

[28] Tim Menzies, Bora Caglayan, Zhimin He, Ekrem Kocaguneli, Joe Krall, Fayola Peters, and Burak Turhan. The promise repository of empirical software engineering data, June 2012.

[29] Marian Jureczko and Lech Madeyski. Towards identifying software project clusters with regard to defect prediction. In *Proceedings of the 6th International Conference on Predictive Models in Software Engineering*, PROMISE '10, pages 9:1–9:10, New York, NY, USA, 2010. ACM.

[30] Lech Madeyski and Marian Jureczko. Which process metrics can significantly improve defect prediction models? an empirical study. *Software Quality Journal*, 23(3):393–422, September 2015.

[31] Tim Menzies, Andrew Butcher, David Cok, Andrian Marcus, Lucas Layman, Forrest Shull, Burak Turhan, and Thomas Zimmermann. Local versus Global Lessons for Defect Prediction and Effort Estimation. *IEEE Transactions on Software Engineering*, 39(6):822–834, 2013.

[32] Shuo Wang and Xin Yao. Using Class Imbalance Learning for Software Defect Prediction. *IEEE Transactions on Reliability*, 62(2):434–443, 2013.

[33] Bogdan Vasilescu, Alexander Serebrenik, and Mark van den Brand. By No Means: a Study on Aggregating Software Metrics. In *Proceedings of the 2nd International Workshop on Emerging Trends in Software Metrics*, WETSoM '11, pages 23–26. ACM, 2011.

[34] T. Hall, S. Beecham, D. Bowes, D. Gray, and S. Counsell. A systematic literature review on fault prediction performance in software engineering. *Software Engineering, IEEE Transactions on*, 38(6):1276–1304, Nov 2012.

[35] Péter Hegedűs. Advances in Software Product Quality Measurement and its Applications in Software Evolution. In *Proceedings of the 31st International Conference on Software Maintenance and Evolution – ICSME'15*. IEEE, 2015, accepted, to appear.

[36] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 2009.

[37] Lionel C Briand, Walcelio L. Melo, and Jurgen Wust. Assessing the Applicability of Fault-proneness Models Across Object-oriented Software Projects. *IEEE Transactions on Software Engineering*, 28(7):706–720, 2002.

Annales Mathematicae et Informaticae
46 *(2016) pp. 143–163*
http://ami.ektf.hu

Methods to analysis of queueing models with state-dependent jump priorities

Agassi Melikov[a], Anar Rustamov[b]
Turan Jafarzade[c], János Sztrik[d]

[a]Institute of Control Science
agassi.melikov@rambler.ru

[b]Qafqaz University
anar.rustemov@gmail.com

[c]National Aviation Academy
turan_jafarzade@hotmail.com

[d]University of Debrecen
sztrik.janos@inf.unideb.hu

Submitted April 20, 2016 — Accepted September 7, 2016

Abstract

In this paper, exact and approximate approaches for studying queuing models with state-dependent jump priorities are developed. Both models with finite separate buffers and finite common buffer for heterogeneous calls are investigated. It is shown that both models might be described by two-dimensional Markov Chains (2-D MC). Exact approach based on solution of appropriate system of balance equations (SBE) for state probabilities faced with big computational challenges for large scale models. To overcome the indicated difficulties an approximate approach based on the state space merging algorithms is developed. This approach allows to construct simple algorithms to calculate the Quality of Service (QoS) metrics of the examined models. The results of numerical experiments are demonstrated.

Keywords: queuing models, jump priority, Markov chains, space merging, numerical analysis

MSC: 60K25, 68M20, 90B22

1. Introduction

Priorities are effective tools to solve the problems of quality of service (QoS) provisioning of heteregeneous calls in queuing systems. By nature the priorities can be broadly divided into two classes: *static* and *dynamic*. Static priorities (relative or preemptive) are defined in advance and they do not change during the whole system operation time [1]. In literature relative static priorities in queuing systems with buffers sometimes are called HOL-priorities (Head-Of-Line), .i.e. in static priorities call for service is chosen from the head of line according to the highest priority. Dynamic priorities in turn are divided into two classes: *dynamical vs time* and *dynamical vs state*. In *dynamical versus time* priorities the priority of the calls can be changed according their waiting times (or sojourn time) [2]. In *dynamical versus state* priorities (they sometimes are called *state-dependent priorities*) calls can change priority according the state of the system where the state is described by vector whose components indicate number of heterogeneous calls in the queue (or in the system) [3].

The drawback of static priorities is that when they are used in real systems the delay of low priority calls is too large espesially for the system with heavy loads of high priority calls. Dynamic priorities allow to avoid the starvation of low priority calls. Detailed review of priority schemas might be found in [4].

As a rule, classical priorities (static or dynamic) are used to determine type of call from the buffer which must be send to channel for servicing. However, some scientific and practical interest represents the priorities which are introduced to change (either increase or decrease) the priorities of calls in buffer. These changes are realized instantaneously so such kind of priorities are called jump priorities (JP). They might be either static or dynamic too. Let us breifly review existing results related to such kind of priorities.

The pioneer work on the analyzing dynamical vs time HOL-priorities with priority jumps (HOL-JP) is [5]. In this paper dynamical vs time HOL-JP was proposed where calls with low priority can jump to another buffer with high priority after waiting some (deterministic) period of time in native buffer; this process goes until a call of any type gets access to a channel or reaches a queue with highest priority. Formulas for calculation of the mean waiting time of the heterogeneous calls were developed in [5].

Dynamical vs state HOL-JP in discrete-time queuing models were proposed in [6-10]. In these models authors included two kinds of calls - high priority calls (H-calls) and low priority calls (L-calls). A scheme of head-of-line merge-by-probability (HOL-MBP) according to which at the end of each time slot all L-calls go to the end of the queue of H-calls with the fixed probability $\beta, 0 < \beta < 1$, was proposed in [6]. A modification of the HOL-MBP scheme was considered in [7]. It was named head-of-line jump-or-serve (HOL-JOS) and, in contrast to the scheme of [6]], in it only one L-call goes from the queue head into the H-queue. Unlike the HOL-JOS scheme, in HOL-JIA$_1$ (Head-Of-Line Jump If-Arrival) scheme [8] transition of the L-call into the H-queue depends not only on the state of the H-queue at the beginning of the slot, but also on the number of arrivals of L-calls during this slot. The only distinction of the HOL-JIA$_1$ scheme from the HOL-JIA$_2$ scheme [9] lies in that in the latter scheme the L-calls can pass immediately to the H-queue. Formulas for the generating functions of the call queue lengths of both types and the time of H-call waiting on the queue, as well as their moments, were developed in [6-10]. Additionally, the mean time of waiting in the queue of L-calls was determined.

In [5-10] queuing models with infinite buffers are investigated. So, they have little applicability in the real communication networks. In particular, real communication networks have finite buffer capacity. Secondly, investigated JP are defined by state-independent probabilities. Since they cannot be adapted for real situations depending on loads of heterogeneous calls.

Different approach to study queuing models with dynamical vs state HOL-JP can be found in the papers [11–14] and in chapter 5 of the book [15] where new type of randomized state-dependent JP for continuous-time queuing systems with finite buffers was proposed. In papers [11, 12] models with separate buffers for heterogeneous calls have been examined while in paper [13, 14] models with common buffer are investigated. They make it possible pass to from the L-queue into the H-queue only at the instants of arrival of the L-calls, the probability of such transitions depend only on the number of L-calls in the system. In chapter 5 of the book [15] models with separate buffers which jump priorities depending only on the number of H-calls in the system were examined. In the indicated works [11–14] methods of calculation of main QoS metrics of the investigated models are proposed. To the best of our knowledge, models in which JP depends on the number of both types of calls in the system are not examined. In this paper we investigate such kinds of models.

At the end of this section it's worth noting that in [16] queueing models with finite common buffer and two priority classes of calls are investigated, where it is assumed that H-calls can preempt the service of L-calls. Futhermore in [16] various congestion control mechanisms are also proposed. In order to calculate the steady-state probabilities of the investigated models, new calculation approach of the original method based on the theory of generalized invariant subspace is developed. Unlike [16] preemption of H-calls from the services of L-calls is not allowed in our paper. However L-call can jump H-buffer in order to served as H-calls.

The rest of the paper is organized as follows. In section 2 model with separate buffers and state-dependent JP is examined and both exact and approximate methods of calculation its QoS metrics are developed. Similar problems for model with common buffer are investigated in section 3. Section 4 is about numerical results of models with both separate buffers and common buffer. In numerical experiments, we investigate different schemas of changing elements of JP-matrix. Conclusion remarks are given in section 5.

2. Jump priorities in model with separate buffers

The structure scheme of the studied queuing system is depicted in Figure 1. In the single server queueing system two Poisson traffic of heterogeneous calls have different arrival rate $\lambda_i, i = 1, 2$. We determined first type of calls as high priority calls (H-calls) while second type of calls are treated as low priority calls (L-calls). By default H-call from the buffer dominating to be served by the idle server; only in the case of absence of H-call in the buffer, L-calls can be served. If there isn't any call in the buffer, then the channels becomes free. Service intensity of the server is the same for both type of the call where it is determined as μ obeying exponential distribution.

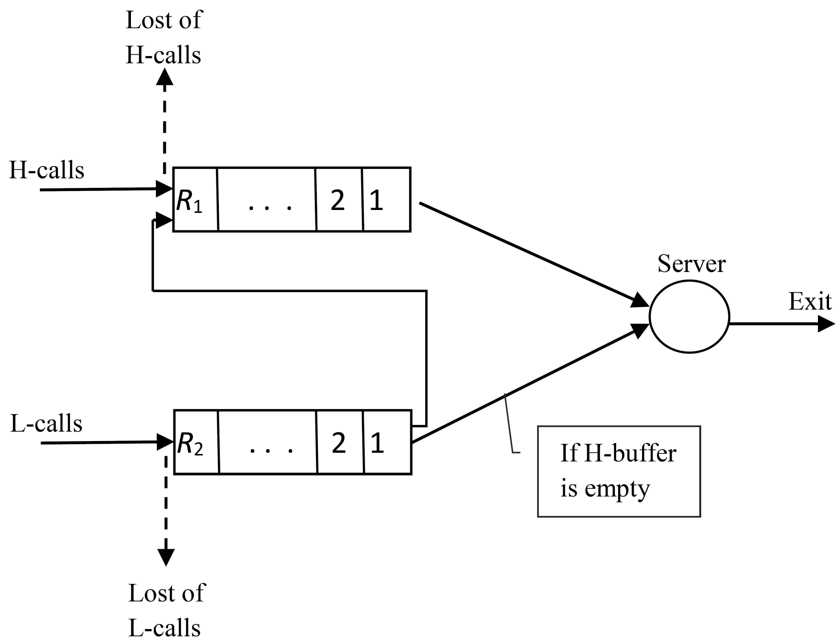


Figure 1: Structure of the queuing system with jump priorities

First let us consider the model with separate buffers, i.e. it is assumed that there are two isolated buffers – H-buffer (for waiting H-calls) and L-buffer (for waiting L-calls) with size of R_1 and$R_2 (0 < R_i < \infty, i = 1, 2)$ respectively.

Decision epochs (i.e. jumping moments of L-calls to H-buffer) coincide with the arrival moments of L-calls. In this model state-dependent HOL-JP is defined as follows:

- High priority calls are always accepted to the H-buffer with the probability 1 if there is a free place in this buffer. If the H-buffer is full then arrived H-call will be dropped with probability 1.

- If upon arrival L-call number of calls of this type equals $i, i < R_2$, and number of H-call equals $j, j < R_1$, then L-call joins the H-buffer with probability $\alpha_i(j)$ and in future it will be served as H-call; and arrived L-call joins the L-buffer with probability $1 - \alpha_i(j)$.

- If upon arrival L-call number of H-call equals R_1, then L-call joins the L-buffer if there is free place in this buffer; otherwise, arrived L-call will be dropped with probability 1.

- If upon arrival L-call L-buffer is full and number of H-call equals $j, j < R_1$, then L-call joins the H-buffer with probability $\alpha_{R_2}(j)$; and arrived L-call will be dropped with probability $1 - \alpha_{R_2}(j)$.

In other words, to define JP matrix with dimension $(R_2 + 1)$ x R_1 is introduced. Entities of this matrix (JP-matrix) are $\alpha_i(j), i = 0, 1, \ldots, R_2, j = 0, 1, \ldots, R_1 - 1$. Thus in this scheme entities of JP-matrix depends on both number of heterogeneous calls in the appropriate buffers.

Let us note some important special schemes regarding the jump priorities mentioned above.

1) *The uniform schemas.* In this schemas, the elements of JP-matrix does not depend on the number of heteregeneous calls in the buffers. So, if the elements of JP-matrix does not depend on the number of H-calls in the buffer, i.e., $\alpha_i(j) = \alpha_i$ for any $j = 0, 1, \ldots, R_1 - 1$, then we obtain JP-schema which was proposed in [11, 12]. Here in the special case $\alpha_i = 0$, we obtain the classical HOL-priorities. Alternative case is that the elements of JP-matrix do not depend on the number of L-calls in the buffer, i.e., $\alpha_i(j) = \alpha(j)$ for any $i = 0, 1, \ldots, R_2$. Such kind of JP has been investigated in [15]. In last scheme in the special case $\alpha(j) = 1$ for any $i = 0, 1, \ldots, R_2$, we obtain the classical queuing system with single traffic.

2) *The threshold-based schemas.* In these schemas, the threshold parameters $T_i, 0 \leq T_i \leq R_1 - 1, i = 0, 1, \ldots, R_2$, are introduced, and elements of JP-matrix are defined as follows:

$$\alpha_i(j) = \begin{cases} \alpha_{i1}, & \text{if } 0 \leq j \leq T_i, \\ \alpha_{i2}, & \text{if } j > T_i. \end{cases}$$

The probabilities $\alpha_{ij},\ i = 0, 1, \ldots, R_2,\ j = 1, 2$, can be defined in various ways. In special cases, i.e. when these probabilities equal either 0 or 1 we obtain different non-randomized threshold-based JP-schemas.

The problem is finding the QoS metrics for this model. The main QoS metrics are the following: the stationary probability of lossing the calls of the ith type (CLP_i), the mean number of the ith type calls in the buffers (L_i) and the mean call transmission delay of the ith type calls (CTD_i), i=1, 2.

2.1. Exact method for model with separate buffers

The state of the system is defined by two dimensional vectors $\boldsymbol{n} = (n_1, n_2)$ where the first component indicates the number of H-calls and the second one the number of L-calls respectively. In other words, operation of this system is described by the two-dimensional Markov Chain (2-D MC) with the following state space:

$$S = \{\boldsymbol{n}: \ n_i = 0, 1, \ldots, R_i, \ i = 1, 2\}. \tag{2.1}$$

Transition intensity from state $\boldsymbol{n} \in S$ to state $\boldsymbol{n'} \in S$ are denoted by $q(\boldsymbol{n}, \boldsymbol{n'})$. Then nonnegative elements of the generating matrix (Q-matrix) of the given 2-D MC can be calculated as below (see Figure 2):

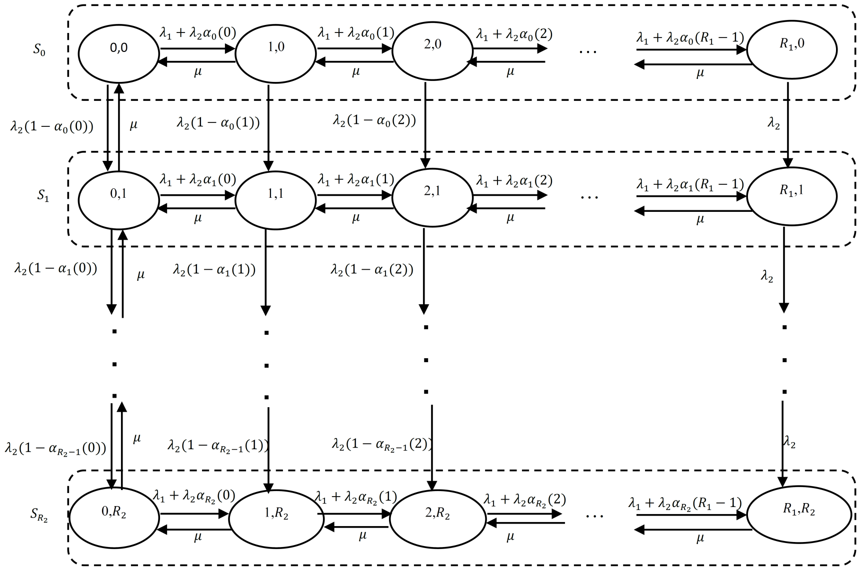


Figure 2: State diagram of the model with separate buffers

$$q(\boldsymbol{n}, \boldsymbol{n'}) = \begin{cases} \lambda_1 + \lambda_2 \alpha_{n_2}(n_1), & \text{if } n_1 < R_1, \ \boldsymbol{n'} = \boldsymbol{n} + \boldsymbol{e}_1, \\ \lambda_2 (1 - \alpha_{n_2}(n_1)), & \text{if } n_1 < R_1, \ n_2 < R_2, \ \boldsymbol{n'} = \boldsymbol{n} + e_2, \\ \lambda_2, & \text{if } n_1 = R_1, \ \boldsymbol{n'} = \boldsymbol{n} + e_2, \\ \mu, & \text{if } n_1 > 0, \ \boldsymbol{n'} = \boldsymbol{n} - \boldsymbol{e}_1 \text{ or} \\ & \quad n_1 = 0, \ \boldsymbol{n'} = \boldsymbol{n} - \boldsymbol{e}_2, \\ 0 & \text{in other cases,} \end{cases} \tag{2.2}$$

where $\boldsymbol{e}_1 = (1, 0)$, $\boldsymbol{e_2} = (0, 1)$.

For any positive values of the parameters of incoming traffics, all states of the investigated finite MC are communicating (see Figure 2), and consequently, stationary distribution of the investigated 2-D MC exists.

The stationary probability of state $\boldsymbol{n} \in S$ is denoted by $p(\boldsymbol{n})$. Construction and solution of the corresponding system of balance equations (SBE) for the given

2-D MC is the standard way for determining the stationary state probabilities. It is constructed with regard to (2.2) and has the following form:

Case $n_1 < R_1,\ n_2 < R_2$:

$$\begin{aligned}&p\left(n_1, n_2\right)\left(\lambda_1+\lambda_2+\mu I\left(n_1+n_2>0\right)\right)= \\ &p\left(n_1-1, n_2\right) I\left(n_1>0\right)\left(\lambda_1+\lambda_2 \alpha_{n_2}\left(n_1-1\right)\right)+ \\ &p\left(n_1, n_2-1\right) I\left(n_2>0\right) \lambda_2\left(1-\alpha_{n_2-1}\left(n_1\right)\right)+ \\ &p\left(n_1+1, n_2\right) \mu+p\left(n_1, n_2+1\right) \mu I\left(n_1=0\right)\end{aligned} \tag{2.3}$$

Case $n_1 = R_1,\ n_2 < R_2$:

$$\begin{aligned}&p\left(R_1, n_2\right)\left(\mu+\lambda_2\right)= \\ &p\left(R_1-1, n_2\right)\left(\lambda_1+\lambda_2 \alpha_{n_2}\left(R_1-1\right)\right)+p\left(R_1, n_2-1\right) I\left(n_2>0\right) \lambda_2;\end{aligned} \tag{2.4}$$

Case $n_2 = R_2$:

$$\begin{aligned}&p\left(n_1, R_2\right)\left(\mu+\lambda_1+\lambda_2 \alpha_{R_2}\left(n_1\right) I\left(n_1<R_1\right)\right)= \\ &p\left(n_1-1, R_2\right) \mu I\left(n_1>0\right)+p\left(n_1, R_2-1\right)_2\left(1-\alpha_{R_2-1}\left(n_1\right)\right)\ .\end{aligned} \tag{2.5}$$

The SBE (2.3)–(2.5) should be completed with normalizing condition over state space (2.1):

$$\sum_{\boldsymbol{n} \in S} p(\boldsymbol{n})=1. \tag{2.6}$$

After determining the state probabilities from SBE (2.3)–(2.6), one can establish its QoS metrics. As indicated above, H-calls are lost if upon their arrivals H-buffer is full. Hence, the loss probability for H-calls (CLP_1) can be determined as follows:

$$CLP_1=\sum_{i=0}^{R_2} p\left(R_1, i\right). \tag{2.7}$$

Similarly, we conclude that L-calls are lost in the following cases: (2.1) at the time an L-call arrives, both buffers are full (in such case L-call is lost with probability 1); (2.2) at the time an L-call arrives, L-buffer is full but there is free place in H-buffer (in such case L-call is lost with probability$1-\alpha_{R_2}(i)$). Thus the loss probability of L-calls (CLP_2) is given by

$$CLP_2=p\left(R_1, R_2\right)+\sum_{i=0}^{R_1-1} p\left(i, R_2\right)\left(1-\alpha_{R_2}(i)\right). \tag{2.8}$$

The first and second terms of the sum in the formula (2.8) denote the probability of events (2.1) and (2.2), respectively.

The mean numbers of the H-calls (L_1) and L-calls (L_2) in the queue are determined as the expected values of appropriate discrete random variables:

$$L_1 = \sum_{i=1}^{R_1} i \sum_{j=0}^{R_2} p\,(i,j); \tag{2.9}$$

$$L_2 = \sum_{i=1}^{R_2} i \sum_{j=0}^{R_1} p\,(j,i). \tag{2.10}$$

Further, formulas (2.3)–(2.6) and modified Little's formula can be used to evaluate the mean times of transmission delay for the heterogeneous:

$$CTD_1 = \frac{L_1}{\lambda_1^{(c)}}\ ; \tag{2.11}$$

$$CTD_2 = \frac{L_2}{\lambda_1^{(c)} + \lambda_2^{(c)}}, \tag{2.12}$$

where $\lambda_1^{(c)}$ and $\lambda_2^{(c)}$ are carried loads of H-calls and L-calls, respectively. These parameters are calculated as follows:

$$\lambda_1^{(c)} = \lambda_1 \left(1 - \sum_{j=0}^{R_2} p\,(R_1, j)\right) + \lambda_2 \sum_{i=0}^{R_1-1} \sum_{j=0}^{R_2} p\,(i,j)\,\alpha_j\,(i);$$

$$\lambda_2^{(c)} = \lambda_2 \sum_{i=0}^{R_1} \sum_{j=0}^{R_2-1} p\,(i,j)\,(1-\alpha_j\,(i)).$$

Thus, to find QoS metrics (2.7)–(2.12), it is necessary to determine the steady-state probabilities of the model from the corresponding SBE (2.3)–(2.6). By implementation of programming languages it is possible to solve the SBE (2.3)–(2.6) for the steady-state probabilities $p\,(\boldsymbol{n})$, $\boldsymbol{n} \in S$ with a help of numerical methods of the linear algebra. This method of calculation of QoS metrics is called the exact (precise) method. In cases of moderate capacity of state space (2.1) this methods is reasonable to calculate QoS metrics of the system. But for large scale system (i.e. when system has large buffers) it isn't suitable. Therefore, we need to find out a more efficient method to calculate the QoS metrics of the models with large buffers.

2.2. Approximate method for model with separate buffers

Below we consider asymptotic analysis of the QoS metrics for large scale models, i.e. when R_1 and R_2 take large values. The developed approximate method has high accuracy for heavy traffic regime of H-calls. In other words, below we consider asymptotic analysis of the large scale model with heavy loads of H-calls, i.e. it is assumed that $\nu_1 \gg \nu_2$,where $\nu_i = \lambda_i/\mu, i = 1, 2$. Note that this assumption is not something extraordinary because introduction of the jump priorities for the L-calls makes sense, namely in the systems with heavy loads of H-calls.

Consider the following splitting of the state space (2.1):

$$S = \bigcup_{i=0}^{R_2} S_i, S_i \bigcap S_j = \emptyset, i \neq j, \tag{2.13}$$

where $S_i = \{\boldsymbol{n} \in S : n_2 = i\}\,,\ i = 0, 1, 2, \ldots, R_2$.

We notice that the assumption made about the relation of the loads of the heterogeneous calls enables one to satisfy the condition for correct use of the algorithms of state space merging of the 2-D MC (see [3, Appendix]): transition intensities within classes $S_i,\ i = 0, 1, \ldots, R_2$, are essentially higher than those between states of different classes.

The classes of microstates S_i are united into individual merged states $< i >$, and in the original state space S the following merge function is defined:

$$U\left(\boldsymbol{n}\right) =< i >,\ \text{if } \boldsymbol{n} \in S_i\ . \tag{2.14}$$

The function (2.14) defines a merged model with the state space

$$\Omega = \{< i >: i = 0, 1, 2, \ldots, R_2\}\,.$$

Let us consider the problem of calculation of state probabilities inside the splitting models. The stationary probability of the state (k, i) in the split model with the state space S_i is denoted by $\rho_i\left(k\right),\ i = 0, 1, 2, \ldots, R_2,\ k = 0, 1, 2, \ldots, R_1$.

Each split model with state space S_i is a one-dimensional birth and death process with the parameters that are calculated as follows (see Figure 2):

$$q_i\left(k_1, k_2\right) = \begin{cases} \lambda_1 + \lambda_2 \alpha_i\left(k_1\right), & \text{if } k_2 = k_1 + 1 \\ \mu, & \text{if } k_2 = k_1 - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2.15}$$

Consequently, we have

$$\rho_i\left(k\right) = \prod_{j=0}^{k-1} (\nu_1 + \nu_2 \alpha_i\left(j\right))\rho_i\left(0\right),\ k = 1, \ldots, R_1, \tag{2.16}$$

where $\rho_i\left(0\right) = \left(1 + \sum_{k=1}^{R_1} \prod_{j=0}^{k-1} \left(\nu_1 + \nu_2 \alpha_i\left(j\right)\right)\right)^{-1}$.

The elements of the Q-matrix of the merged model are denoted by

$$q\left(< k >, < k' >\right), < k >, < k' >\in \Omega.$$

According to the algorithm of state space merging of the 2-D MC (see [3, Appendix]) these elements are given by

$$q\left(< k >, < k' >\right) = \sum_{\substack{\boldsymbol{n} \in \boldsymbol{S_k} \\ \boldsymbol{n'} \in \boldsymbol{S_{k'}}}} q\left(\boldsymbol{n}, \boldsymbol{n'}\right) \rho_{n_1}\left(n_2\right). \tag{2.17}$$

So, by using (2.2), (2.16) and (2.17) after some mathematical transformations the following formula are obtained (see Figure 2):

$$q\left(<k>,<k'>\right)= \begin{cases} \lambda_2\left(\rho_k\left(R_1\right)+\sum_{i=0}^{R_1-1}\left(1-\alpha_k\left(i\right)\right)\rho_k\left(i\right)\right), & \text{if } k'=k+1, \\ \mu\rho_k\left(0\right), & \text{if } k'=k-1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.18}$$

From (2.18) we can calculate the probabilities of the merged states

$$\pi\left(<k>\right), <k>\in \Omega$$

as follows:

$$\pi\left(<k>\right)={\nu_2}^k\prod_{j=0}^{k-1}\Lambda_j\pi\left(<0>\right), \; k=1,\dots,R_2, \tag{2.19}$$

where

$$\pi\left(<0>\right)=\left(1+\sum_{k=1}^{R_2}{\nu_2}^k\prod_{j=0}^{k-1}\Lambda_j\right)^{-1},$$

$$\Lambda_j=\frac{\rho_j\left(R_1\right)+\sum_{i=0}^{R_1-1}\left(1-\alpha_j\left(i\right)\right)\rho_j\left(i\right)}{\rho_{j+1}\left(0\right)}, \quad j=0,\dots,R_2-1.$$

The state probabilities of the initial 2-D MC are determined approximately as follows (see [3, Appendix]):

$$p\left(i,j\right)\approx\rho_j\left(i\right)\pi\left(<j>\right). \tag{2.20}$$

By taking into account (2.16), (2.19) and (2.20) we can calculate approximate values of state probabilities of initial 2-D MC, and omitting the intermediate mathematical calculations the following approximate formulae to calculate the QoS metrics (2.7)-(2.10) are obtained:

$$CLP_1\approx\sum_{i=0}^{R_2}\rho_i\left(R_1\right)\pi\left(<i>\right); \tag{2.21}$$

$$CLP_2\approx\pi\left(<R_2>\right)\left(\rho_{R_2}\left(R_1\right)+\sum_{i=0}^{R_1-1}\rho_{R_2}\left(i\right)\left(1-\alpha_{R_2}\left(i\right)\right)\right); \tag{2.22}$$

$$L_1\approx\sum_{i=1}^{R_1}i\sum_{k=0}^{R_2}\rho_k\left(i\right)\pi\left(<k>\right); \tag{2.23}$$

$$L_2 \approx \sum_{k=1}^{R_2} k \ \pi\left(< k >\right). \tag{2.24}$$

The QoS metrics CTD_k are determined from (2.11) and (2.12) after the calculation of the parameters L_k and $\lambda_k^{(c)}$, $k = 1, 2$.

3. Jump priorities in models with common buffer

Now let us consider the model with common buffer. In this case it is assumed that there is single buffer with size R, $0 < R < \infty$, for both types of calls. In this model the state-dependent HOL-JP is defined as follows.

- High priority calls are always accepted to the common buffer with the probability 1 if there is a free place in the buffer. If the common buffer is full then arriving H-call will be dropped with probability 1.
- If upon arrival of L-call the number of H-calls equals n_1 and number of L-calls equals n_2, where $n_1 + n_2 < R$, then arriving L-call becomes H-call with probability $\beta_{n_2}(n_1)$ or it will be accepted to the buffer as L-call with probability $1 - \beta_{n_2}(n_1)$.
- If upon arrival of L-call common buffer is full (i.e. in case $n_1 + n_2 = R$) it will be dropped with probability 1.

The main QoS metrics of the model are the same with the model with separate buffers.

3.1. Exact method for model with common buffer

As it is mentioned in section 2, the 2-D vector $\boldsymbol{n} = (n_1, n_2)$ is used to describe the state of the system and state space for this model is determined as follows (see Figure 3):

$$S = \{\boldsymbol{n} : \ n_i = 0, 1, \ldots, R, \ i = 1, 2; \ n_1 + n_2 = R\} \ . \tag{3.1}$$

Note 1. Hereinafter, for simplicity, we use the same notations for the state spaces, state probabilities etc. in different models. This will not lead to misunderstanding since from the context it will be clear which model is being considered.

In this case the nonnegative elements of the Q-matrix of the given 2-D MC can be calculated as follows (see Figure 3):

$$q\left(\boldsymbol{n}, \boldsymbol{n'}\right) = \begin{cases} \lambda_1 + \lambda_2 \alpha_{n_2}\left(n_1\right), & \text{if } n_1 + n_2 < R, \ \boldsymbol{n'} = \boldsymbol{n} + \boldsymbol{e}_1, \\ \lambda_2\left(1 - \alpha_{n_2}\left(n_1\right)\right), & \text{if } n_1 + n_2 < R, \ \boldsymbol{n'} = \boldsymbol{n} + e_2, \\ \mu, & \text{if } n_1 > 0, \ \boldsymbol{n'} = \boldsymbol{n} - \boldsymbol{e}_1 \text{ or} \\ & \quad n_1 = 0, \ \boldsymbol{n'} = \boldsymbol{n} - \boldsymbol{e}_2, \\ 0 & \text{in other cases.} \end{cases} \tag{3.2}$$

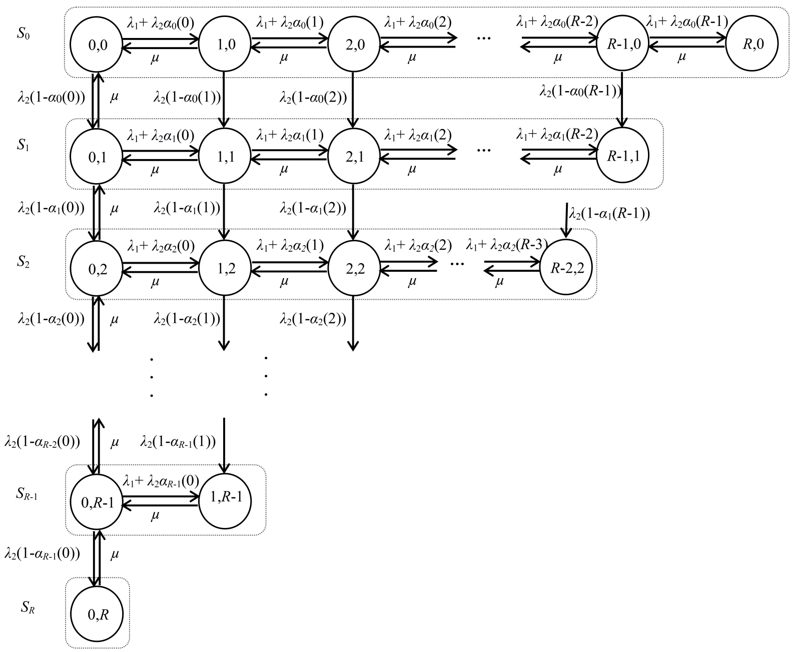


Figure 3: State diagram of the model with common buffer

Loss probabilities of heterogeneous calls in this model equal each other and are given by

$$CLP_1 = CLP_2 = \sum_{\boldsymbol{n} \in S_d} p(\boldsymbol{n}), \tag{3.3}$$

where $S_d = \{\boldsymbol{n} \in S : \ n_1 + n_2 = R\}$ is set of diagonal states (see Figure 3).

In this model the mean number of the H-calls and L-calls in the buffer are determined as follows:

$$L_k = \sum_{i=1}^{R} i \xi_k(i) \tag{3.4}$$

where $\xi_k = \sum_{\boldsymbol{n} \in S} p(\boldsymbol{n}) \, \delta(n_k = i)$, $k = 1, 2$; $\delta(i, j)$ are Kronecker's symbols.

The mean transmission delays of the heterogeneous calls are determined by Eqs. (2.11) and (2.12).

Computational difficulties of the exact approach to solve appropriate SBE (it is constructed with regard to (3.2) and here it is omitted) are observed in the models with large scale too. In order to overcome the above-mentioned computational difficulties to calculate the QoS metrics (3.3), (3.4) an approximate approach to asymptotic analysis can be applied as well.

3.2. Approximate method for model with common buffer

Similar to (2.13) splitting of the state space (3.1) is considered:

$$S = \bigcup_{i=0}^{R} S_i,\ S_i \bigcap S_j = \emptyset,\ i \neq j, \tag{3.5}$$

where $S_i = \{\boldsymbol{n} \in S : n_2 = i\}\,,\ i = 0, 1, \ldots, R.$

Not repeating the stages of the approximate approach, final formula to calculate QoS metrics are given below. In this case stationary distribution of the split model with state space $S_i,\ i = 0, 1, \ldots, R-1,$ is defined as follows:

$$\rho_i\,(k) = \prod_{j=0}^{k-1} (\nu_1 + \nu_2 \beta_i\,(j))\rho_i\,(0)\,, \quad k = 1, \ldots, R-i, \tag{3.6}$$

where $\rho_i\,(0) = \left(1 + \sum_{k=1}^{R-i} \prod_{j=0}^{k-1} (\nu_1 + \nu_2 \beta_i\,(j))\right)^{-1}.$

Note 2. Split model with state space S_R contains only one micro-state $(0, R) \in S$ so we set $\rho_R\,(0) = 1$.

According to the algorithm of state space merging of the 2-D MC, the elements of the Q-matrix of the merged model in this case (see formulae (2.17), (3.2) and (3.6)) are given by

$$q\,(< k >, < k' >) = \begin{cases} \lambda_2 \sum_{i=0}^{R-k-1} \rho_k\,(i)\,(1 - \beta_k\,(i)), & \text{if } k' = k+1, \\ \mu \rho_k\,(0)\,, & \text{if } k' = k-1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

So, the probabilities of the merged states in this model are calculated as follows:

$$\pi\,(< k >) = {\nu_2}^k \prod_{j=0}^{k-1} M_j \pi\,(< 0 >)\,,\ k = 1, \ldots, R, \tag{3.8}$$

where $\pi\,(< 0 >) = \left(1 + \sum_{k=1}^{R} {\nu_2}^k \ \prod_{j=0}^{k-1} M_j\right)^{-1}$, $M_j = \frac{\sum_{i=0}^{R-j-1} \rho_j(i)(1-\beta_j(i))}{\rho_{j+1}(0)}$, $j = 0, \ldots, R-1.$

So, by using (3.6) and (3.8) from (2.20) we can calculate approximate values of steady-state probabilities of initial 2-D MC for the model with common buffer. Consequently, for asymptotic analysis of QoS metrics of the investigated model the following approximate formula are obtained:

$$CLP_1 \approx CLP_2 \approx \sum_{i=0}^{R} \rho_i\,(R-i)\pi\,(< i >)\,; \tag{3.9}$$

$$L_1 \approx \sum_{k=1}^{R} k \sum_{i=0}^{R-k} \rho_i\,(k)\pi\,(< i >)\,; \tag{3.10}$$

$$L_2 \approx \sum_{k=1}^{R} k \ \pi\,(< k >). \tag{3.11}$$

Based on (3.10) and (3.11) approximate values of transmission delays of the heterogeneous calls are calculated by (2.11) and (2.12).

4. Numerical results

The developed approximate formula allow one to carry out an authentic analysis of QoS metrics over any range of change of values of loading parameters of the heterogeneous traffic, satisfying assumption concerning their ration (i.e. when$\nu_1 \gg \nu_2$) and also at any buffers sizes.

Let us first examine the results of the numerical experiments for the model with separate buffers. The following initial data for hypothetical model was selected: $R_1 + R_2 = 110$, $\lambda_1 = 2$, $\lambda_2 = 0.5$, $\mu = 3$, i.e. $\nu_1 = 2/3$, $\nu_2 = 1/6$. The numerical results are analyzed based on the two schemas for changing the elements of JP-matrix. In schema 1 it is assumed that they are changed with respect to both parameters (state-dependent JP) and defined as $\alpha_i\,(j) = \frac{i+1}{i+j+2}$ while in second one we assume that they are constant (state-independent JP), i.e. $\alpha_i\,(j) = 0.5$ for any i and j. In other words, in schema 1 probabilities of jumping to H-buffer are decreasing function with respect to number of H-calls in buffer at fixed values of L-calls but in schema 2 they do not depend on number of heterogeneous calls in buffers.

Figures 4–6 show dependences of QoS metrics on R_1. As it was expected the loss probability of H-calls is positively related to the buffer size of H-buffer (Figures 4) while loss probability of L-calls is increasing function versus R_1. As we see from Figures 4, rate of change of the indicated functions are high enough. Also from this figure we conclude that schema 1 is favorable for the loss probabilities of H-calls while for loss probabilities of L-calls schema 2 is favorable. Moreover, differences between values of loss probabilities of H-calls are essential in different schemas especially at large buffer sizes but values of loss probabilities of L-calls are very close to each other in different schemas.

Let us note that from this graph for both schemas we may find such values of buffer sizes for which difference between loss probabilities of heterogeneous calls is less than given ϵ>0 (such kind of problems are called ϵ-fair servicing policy).

Dependency of length of heterogeneous calls on the H-buffer size is shown in Figures 5. In both schemas the mean queue length of the H-calls positively related to the H-buffer size but length of the L-calls is negatively related to the H-buffer size. From this figure we conclude that schema 1 again is favorable for length of the H-calls while length of the L-calls is invariant to different schemas. Behavior of the indicated QoS metrics is interesting. So, length of the H-calls in both schemas increases with low rates for small values of R_1, and for about $R_1 > 20$ they are almost constant; alternative situation occurs for length of the L-calls in both

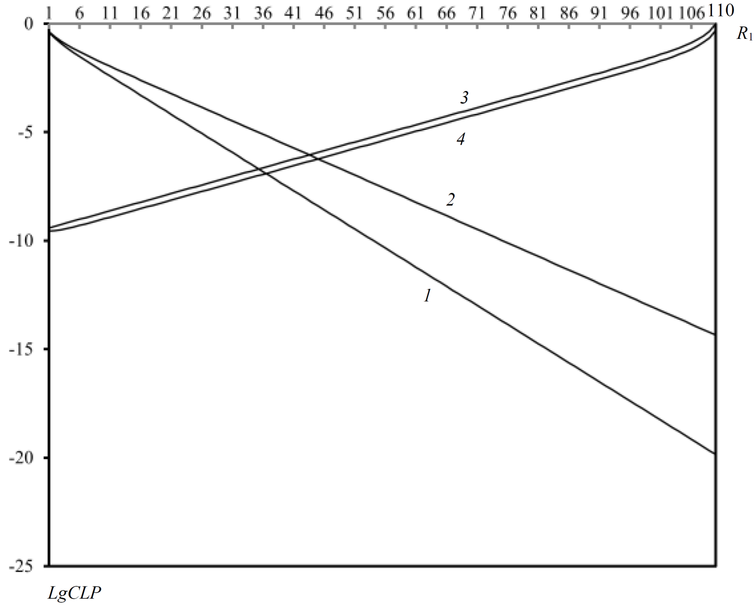


Figure 4: Dependence of loss probabilities versus R_1 in model with separate buffers: 1- CLP_1 in *schema 1*; 2- CLP_1 in *schema 2*; 3- CLP_2 in *schema 1*; 4- CLP_2 in *schema 2*

schemas, i.e. they are almost constant for $R_1 < 100,$ and after that they decrease with low rates.

Behavior of both functions CTD_1 and CTD_2 are very similar to behavior of functions L_1 and L_2 respectively (see Figures 6). In other words, schema 1 again is favorable for CTD_1 while CTD_2 is almost constant in different schemas; in both schemes CTD_1 increases with low rates for small values of R_1, and for about $R_1 > 20$ they are almost constant and CTD_2 in both schemas is almost constant for $R_1 < 100,$ and for $R_1 > 100$ it decreases with low rates.

Let us now consider the results for model with common buffer based on above indicated different schemas of changing of parameters $\alpha_i\left(j\right),\ i = 0, 1, 2, \ldots, R-1,\ j = 0, 1, \ldots, R-i-1$. Loads of this model are unchanged, i.e. we select$\nu_1 = 2/3,\ \ \nu_2 = 1/6$.

Dependency of function CLP (as it was mentioned above loss probabilities of heterogeneous calls in this model equal each other, see (3.2)) on the buffer size is shown in Figures 7. It is seen from this figure that in both schemas the loss probability of calls strictly decreases (with high rate) versus the common buffer size and as it was expected schema 1 is favorable for the loss probabilities. Note that differences between values of loss probabilities in different schemas are increased versus buffer size.

In Figures 8 the dependency of functions L_1 and L_2 on the buffer size is shown.

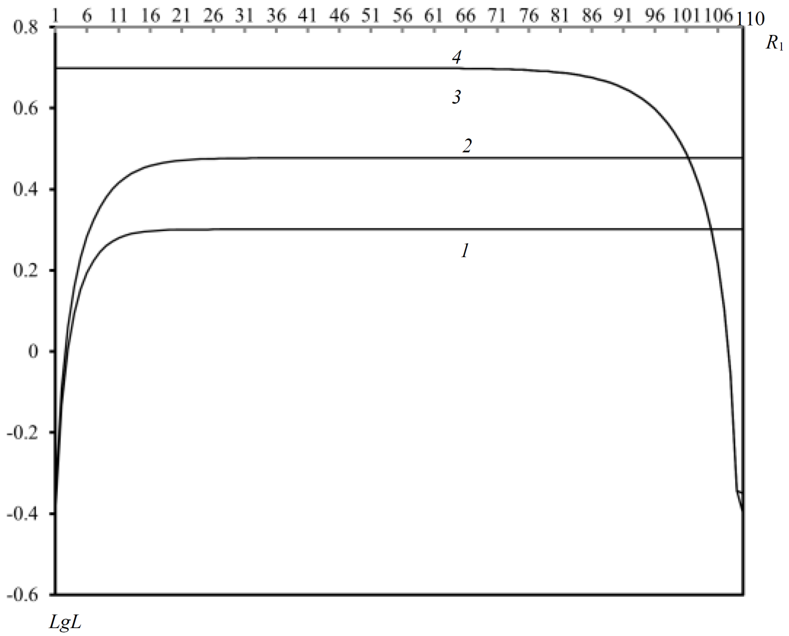


Figure 5: Dependence of mean length of queues versus R_1 in model with separate buffers: 1- L_1 in *schema 1*; 2- L_1 in *schema 2*; 3- L_2 in *schema 1*; 4- L_2 in *schema 2*

In both schemas these functions increase versus the common buffer size. From this figure we conclude that schema 1 is favorable for length of the H-calls while schema 2 is favorable for length of the L-calls. As in case of the model with separate buffers, in this model the rate of change (increasing) of these functions are very small too, i.e. about $R > 15$ they are almost constant.

Again behavior of both functions CTD_1 and CTD_2 is very similar to behavior of functions L_1 and L_2 respectively (see Figures 9). It is interesting that in this case about $R > 15$ values of the function CTD_1 in schema 1 are almost same with values of the function CTD_2 in schema 2.

Presented numerical results allow to take some comparisons proposed in two buffer management mechanisms. So, for instance, values of both functions CLP_1 and CLP_2 in model with separate buffers equal (approximately) $10^{-6.5}$ and this value corresponds to buffers size $R_1 = 35,\ R_2 = 75$ ($R_1 + R_2 = 110$). However, the indicated value for both kinds of calls might be provided in model with common buffer at size $R = 36$. In other words, common buffer is essentially effective buffer management mechanisms for call loss probabilities. Other interesting conclusions with respect to the rest QoS metrics in different buffer management mechanisms might be carried out.

Another goal of performing numerical experiments was the estimation of the

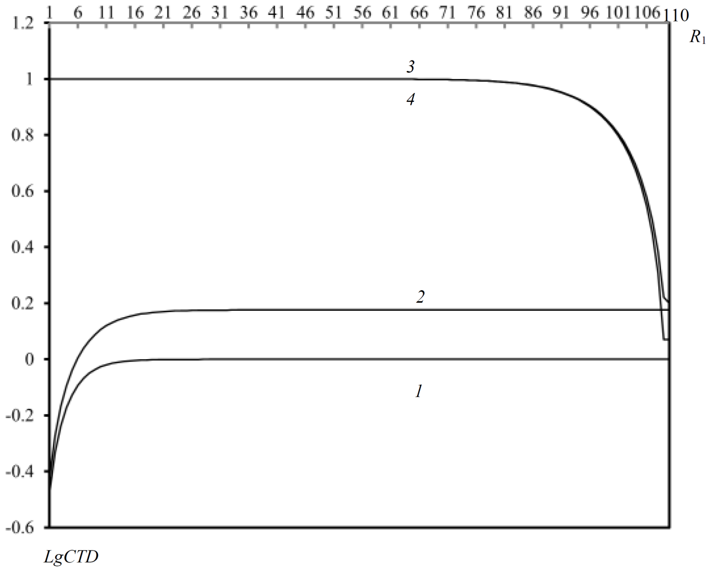


Figure 6: Dependence mean transmission delays versus R_1 in model with separate buffers: 1- CTD_1 in *schema 1*; 2- CTD_1 in *schema 2*; 3- CTD_2 in *schema 1*; 4- CTD_2 in *schema 2*

proposed approximate formulae accuracy. As it was indicated in section 2, the exact values (EV) of the QoS metrics are determined by the appropriate SBE (such an approach allows studying QoS metrics of the model only for small buffer stores).

In order to be short, here in Table 1 and the results only for the model with separate buffers are demonstrated only for the schema 1 (similar results are obtained for the model with common buffer in both schemas as well). Here initial data was selected as above, i.e. $R_1 + R_2 = 110$, $\lambda_1 = 2$, $\lambda_2 = 0.5$, $\mu = 3$.

As it is given in the tables accuracy of the proposed approximate formulae are acceptable for engineering practice. The bigger the ratio $\boldsymbol{v_1/v_2}$, the higher accuracy of approximate value (AV).

5. Conclusion

This paper proposed a new class of state-dependent JP in queueing systems with finite separate buffers and finite common buffer for heterogeneous calls. An exact and effective approximate approaches for calculating the QoS metrics of heterogeneous calls in such systems are developed. The important advantage of approximate approach lies in the use of explicit formulae to calculate the QoS metrics, which enables our approach to be used for models of any dimension. In addition, it is possible to use the proposed formulae to find the optimal (in given sense) values of

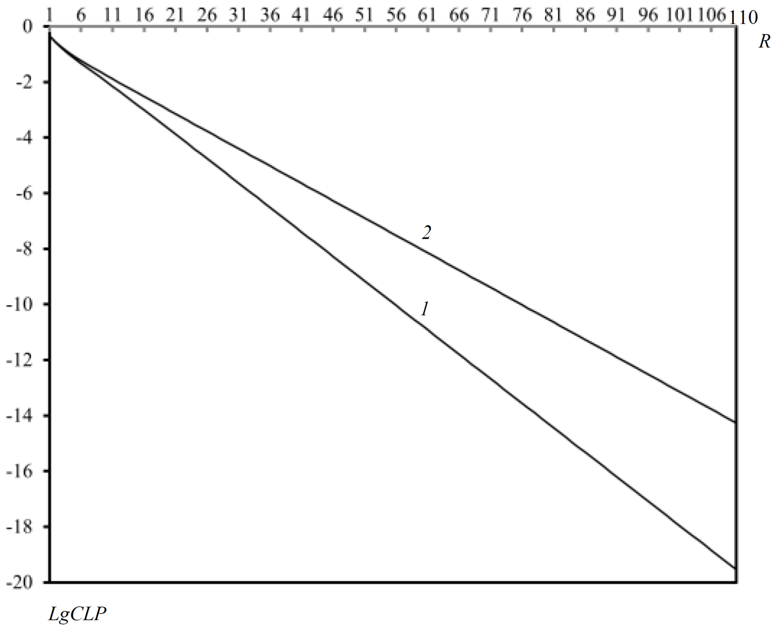


Figure 7: Dependence of loss probabilities versus R in model with common buffer: 1-*schema 1*, 2-*schema 2*

JP-matrix. Latest problems are important especially for the threshold-based non-randomized JP-schemas (see end of section 2) and they are a subject for further study.

Acknowledgements. The publication was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

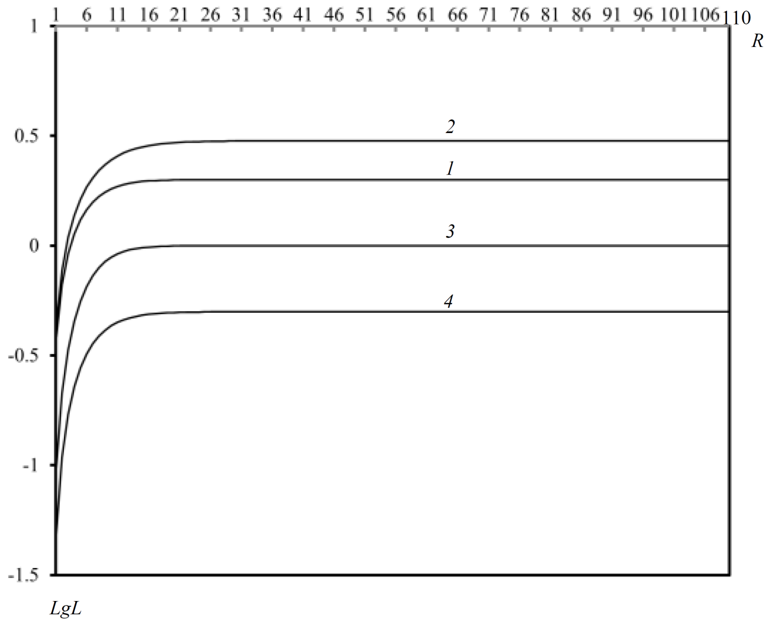


Figure 8: Dependence of mean length of queues versus R in model with common buffer: 1- L_1 in *schema 1*; 2- L_1 in *schema 2*; 3- L_2 in *schema 1*; 4- L_2 in *schema 2*

$\boldsymbol{R_1}$	$\boldsymbol{CLP_1}$		$\boldsymbol{L_1}$	
	$\boldsymbol{EV}$	$\boldsymbol{AV}$	$\boldsymbol{EV}$	$\boldsymbol{AV}$
15	6.76E-03	7.62E-03	1.30929	1.97560
22	1.52E-05	4.46E-05	1.64089	1.99795
29	3.63E-05	2.61E-06	1.75494	1.99984
36	8.87E-06	1.53E-07	1.79172	1.99998
43	2.21E-08	8.93E-09	1.80311	1.99999
50	5.53E-09	5.23E-10	1.80654	2
57	1.43E-10	3.06E-11	1.80756	2
64	3.58E-11	1.79E-12	1.80787	2
71	9.22E-12	1.05E-13	1.80797	2
78	2.39E-13	6.13E-15	1.80808	2
85	6.21E-14	3.59E-16	1.80844	2
92	1.63E-16	2.12E-17	1.80998	2
99	4.46E-17	1.23E-18	1.81776	2
106	1.67E-18	7.24E-20	1.87771	2

Table 1: Comparison for H-calls in model with separate buffers in schema 1

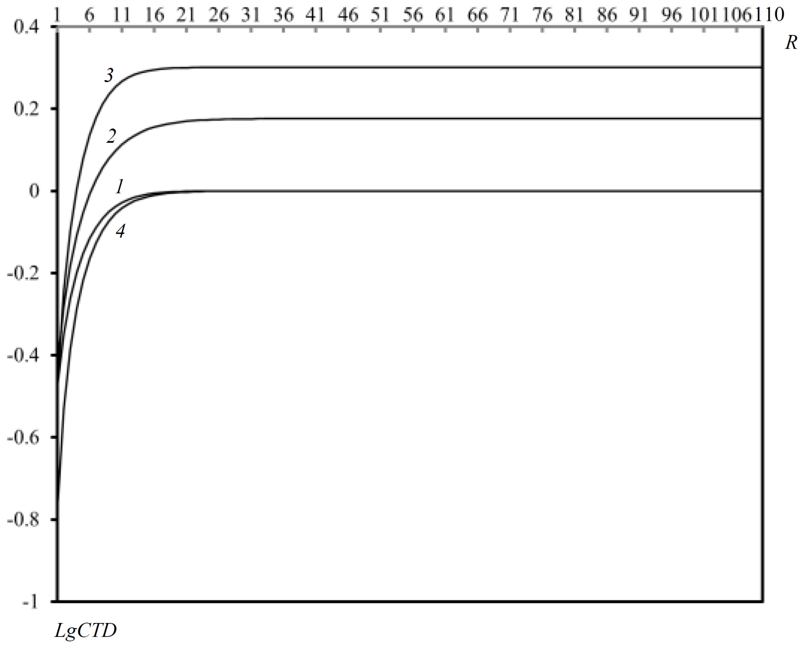


Figure 9: Dependence mean transmission delays versus R in model with common buffer: 1- CTD_1 in *schema 1*; 2- CTD_1 in *schema 2*; 3- CTD_2 in *schema 1*; 4- CTD_2 in *schema 2*

$\boldsymbol{R_1}$	$\boldsymbol{CLP_2}$		$\boldsymbol{L_2}$	
	EV	***AV***	***EV***	***AV***
15	3.21E-10	5.01E-09	4.193002	4.999998
22	6.25E-09	1.79E-08	4.201397	4.999992
29	4.13E-08	6.43E-08	4.196922	4.999974
36	1.93E-07	2.33E-07	4.193955	4.999914
43	7.84E-07	8.25E-07	4.192694	4.999719
50	9.02E-07	2.96E-06	4.192222	4.999098
57	1.15E-06	1.06E-05	4.192017	4.997138
64	4.36E-06	3.89E-05	4.191798	4.991074
71	1.68E-05	1.36E-04	4.191215	4.972766
78	6.61E-05	4.89E-04	4.189308	4.919352
85	2.67E-04	1.76E-03	4.182937	4.770876
92	1.13E-03	6.462E-03	4.161524	4.386067
99	5.26E-03	2.53E-02	3.08775	3.484102
106	1.20E-02	1.34E-01	1.795528	1.640507

Table 2: Comparison for L-calls in model with separate buffers in schema 1

References

[1] Jaiswal HK. Priority queues. New York, *Academic Press* (1968).

[2] Kleinrock L. A delay dependent queue discipline.*Naval Research Logistics Quarterly Journal*, Vol. 11 (1964), 329–341.

[3] Melikov A, Ponomarenko L, Kim CS., Performance Analysis and Optimization of Multi-Traffic on Communication Networks. Heidelberg: Springer (2010).

[4] Wittevrongel S, De Vuyst S, Sys C, Bruneel H., A reservation-based scheduling mechanism for fair QoS provisioning in packet-based networks. *In: Proceeding of the 26th IEEE International Teletraffic Congress*, Karlskrona (2014), 55–62.

[5] Lim Y, Kobza JE., Analysis of delay dependent priority discipline in an integrated multiclass traffic fast packet switch, *IEEE Transactions on Communication*, Vol. 38(5)(1990),659–665.

[6] Maertens T, Walraevens J, Bruneel H., On priority queues with priority jumps, *Performance Evaluation*, Vol. 63(12)(2006): 1235–1252.

[7] Maertens T, Walraevens J, Bruneel H., A modified HOL priority scheduling discipline: performance analysis, *European Journal of Operation Research*, Vol. 180(3)(2007): 1168–1185.

[8] Maertens T, Walraevens J, Moeneclaey M, Bruneel H., A new dynamic priority scheme: performance analysis, *In: Proceeding of the 13th International Conference on Analytical and Stochastic Modeling Techniques and Applications*, Bonn (2006), 74–84.

[9] Maertens T, Walraevens J, Bruneel H., Performance comparison of several priority schemes with priority jumps, *Annals of Operation Research*, Vol.! 162 (2008), 109–125.

[10] Walraevens J, Steyaert B, Bruneel H., Performance analysis of single-server ATM queue with priority scheduling, *Computers and Operation Research*, Vol. 30(12)(2003), 1807–1829.

[11] Melikov AZ, Ponomarenko LA., Kim CS., Algorithmic approach to analysis of queuing system with finite buffers and jump priorities, *Journal of Automation and Information Sciences*, Vol. 44(12)(2012),43–54.

[12] Kim CS, Oh Y, Melikov AZ., A space merging approach to the analysis of the performance of queueing models with buffers and priority jumps, *Industrial Engineering and Management Systems*, Vol. 12(3)(2013),274–280.

[13] Melikov AZ, Ponomarenko LA, Kim CS., Approximate method for analysis of queuing models with jump priorities, *Automation and Remote Control*, Vol. 74(1)(2013), 62–75.

[14] Melikov AZ, Ponomarenko LA, Kim CS., Numerical method for analysis of queuing models with priority jumps, *Cybernetics and System Analysis*, Vol. 49(1)(2013), 55–61.

[15] Melikov A., Ponomarenko L., Multidimensional queueing models in telecommunication networks, *Heidelberg: Springer*, (2014).

[16] Tran HT., Do TV., Pap L., Analysis of a queue with two priority classes and feedback control, *Vietnam Journal of Computer Science*, Vol. 1 (2014), 71–78.

Annales Mathematicae et Informaticae
46 *(2016) pp. 165–173*
http://ami.ektf.hu

Recurrence sequences in the hyperbolic Pascal triangle corresponding to the regular mosaic $\{4,5\}$

László Németh[a], **László Szalay**[ac]

[a]Institute of Mathematics
University of West Hungary, Sopron
nemeth.laszlo@nyme.hu,szalay.laszlo@nyme.hu

[c]Department of Mathematics and Informatics
J. Selye University, Komarno, Slovakia

Submitted April 17, 2016 — Accepted September 5, 2016

Abstract

Recently, a new generalization of Pascal's triangle, the so-called hyperbolic Pascal triangles were introduced. The mathematical background goes back to the regular mosaics in the hyperbolic plane. In this article, we investigate the paths in the hyperbolic Pascal triangle corresponding to the regular mosaic $\{4,5\}$, in which the binary recursive sequences $f_n = \alpha f_{n-1} \pm f_{n-2}$ are represented ($\alpha \in \mathbb{N}^+$).

Keywords: Pascal triangle, hyperbolic Pascal triangle, binary recurrences.

MSC: 11B37, 05A10.

1. Introduction

In the hyperbolic plane there are an infinite number of types of regular mosaics (see, for example [4]), they are assigned by Schläfli's symbol $\{p,q\}$, where the positive integers p and q satisfy $(p-2)(q-2) > 4$. Each regular mosaic induces a so-called hyperbolic Pascal triangle (see [1]), following and generalizing the connection between the classical Pascal's triangle and the Euclidean regular square mosaic $\{4,4\}$. For more details see [1], but here we also collect some necessary information.

There are several approaches to generalize Pascal's arithmetic triangle (see, for instance [3]). The hyperbolic Pascal triangle based on the mosaic $\{p, q\}$ can be figured as a digraph, where the vertices and the edges are the vertices and the edges of a well defined part of the lattice $\{p, q\}$, respectively, further each vertex possesses a value, say label, giving the number of different shortest paths from the fixed base vertex. Figure 1 illustrates the hyperbolic Pascal triangle linked to $\{p, q\} = \{4, 5\}$. Generally, for $\{4, q\}$, the quadrilateral shape cells surrounded by appropriate edges are corresponding to the squares in the mosaic. The base vertex has two edges (both are outgoing), the leftmost and the rightmost vertices have three (one ingoing and two outgoing), the others have q edges (either two ingoing and $q-2$ outgoing (type A) or one ingoing and $q-1$ outgoing (type B)). In other words, apart from the winger elements, vertices of type A have two ascendants and $q-2$ descendants, vertices of type B do one ascendant and $q-1$ descendants. In the figures, we denote the A-type vertices by red circle and B-type vertices by cyan diamond, further the wingers by white diamond. The vertices having distance n from the base vertex are located in row n. The general method of drawing is the following. Going along the vertices of the n^{th} row, according to type of the elements (winger, A, B), we draw appropriate number of edges downward (2, $q-2$, $q-1$, respectively). Neighbor edges of two neighbor vertices of the n^{th} row meet in the $(n+1)^{th}$ row, constructing a vertex of type A. The other descendants of row n in row $n+1$ have type B, except the two wingers. In the sequel, $\left\rangle {n \atop k} \right\langle$ denotes the k^{th} element in row n, which is either the sum of the labels of its two ascendants or coincide the label of its unique ascendant. For instance, if $\{p, q\} = \{4, 5\}$, then

$$\left\rangle {4 \atop 6} \right\langle = 5 = 2 + 3 = \left\rangle {3 \atop 2} \right\langle + \left\rangle {3 \atop 3} \right\langle \qquad \text{and} \qquad \left\rangle {4 \atop 5} \right\langle = 2 = \left\rangle {3 \atop 2} \right\langle$$

hold (see Figure 1). We note, that the hyperbolic Pascal triangle has the property of vertical symmetry.

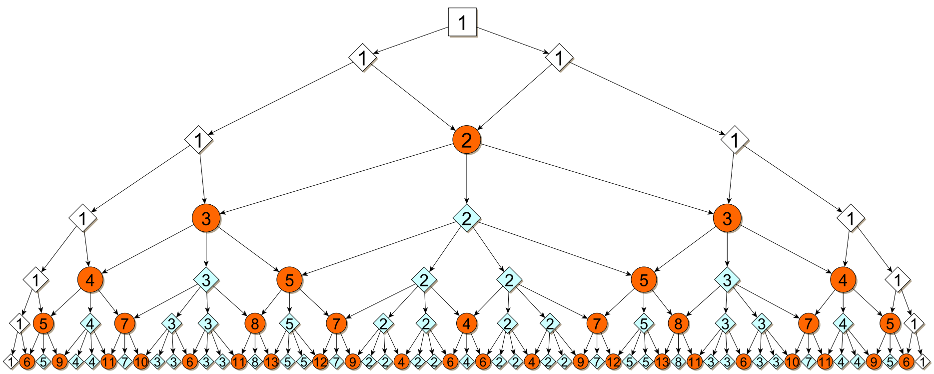


Figure 1: Hyperbolic Pascal triangle linked to $\{4, 5\}$ up to row 6

2. Recurrence sequences linked to $\{4,5\}$

Let $\{p,q\} = \{4,5\}$ be fixed, further we let $\mathcal{HPT}_{45}$ denote the hyperbolic Pascal triangle corresponding to the mosaic $\{4,5\}$. It was showed in [1] that all the binary recurrence sequences $(f_i)_{i\geq 0}$ which are defined by

$$f_i = \eta f_{i-1} + f_{i-2}, \quad (n \geq 2), \tag{2.1}$$

where η and $f_0 < f_1$ are positive integers, appear in $\mathcal{HPT}_{45}$.

In the following we describe paths corresponding to further positive integer binary recurrence sequences. We remark that although we restrict ourselves to $\mathcal{HPT}_{45}$, the methods and the results have been worked out can be fitted to other hyperbolic Pascal triangles with $p = 4,\ q \geq 6$.

Taking a vertex of type A in row n, it has exactly two descendants of type A in the row $n+1$. In order to reach and distinguish them, we denote the left-down step and right-down step (along the appropriate edge of the graph) by L and R, respectively. For the sake of brevity, the sequence of $\ell + r$ consecutive steps

$$\underbrace{LL\cdots L}_{\ell}\underbrace{RR\cdots R}_{r}$$

will be denoted by $L^\ell R^r$. Till the end of this work, such a path is always considered on vertices of type A. Generally, we are interested in the labels of these vertices, therefore sometimes we call them elements (as the elements or terms of a sequence), but if it is necessary we determine the location of the element, too.

This paper will use the next theorem (Theorem 5 in [1]), which states that any two positive integers can be found next to each other somewhere in $\mathcal{HPT}_{45}$.

Theorem 2.1. *Given $u, v \in \mathbb{N}^+$, then there exist $n, k \in \mathbb{N}^+$ such that $u = \genfrac{\rangle}{\langle}{0pt}{}{n}{k}$ and $v = \genfrac{\rangle}{\langle}{0pt}{}{n}{k+1}$.*

Using Theorem 2.1, Corollary 2.2 provides an immediate consequence of the properties of $\mathcal{HPT}_{45}$.

Corollary 2.2. *If $u = \genfrac{\rangle}{\langle}{0pt}{}{n}{k} < v = \genfrac{\rangle}{\langle}{0pt}{}{n}{k+1}$ holds for some positive integers u and v, then $\genfrac{\rangle}{\langle}{0pt}{}{n}{k+2} = v - u$, moreover the type of $\genfrac{\rangle}{\langle}{0pt}{}{n}{k+1}$ is A, while the types of $\genfrac{\rangle}{\langle}{0pt}{}{n}{k}$ and $\genfrac{\rangle}{\langle}{0pt}{}{n}{k+2}$ are not A (i.e., either B or winger).*

Remark 2.3. Clearly, by the symmetry we also have the construction $u = \genfrac{\rangle}{\langle}{0pt}{}{n}{k} > v = \genfrac{\rangle}{\langle}{0pt}{}{n}{k+1}$ and $\genfrac{\rangle}{\langle}{0pt}{}{n}{k-1} = u - v$. Further, the type of $\genfrac{\rangle}{\langle}{0pt}{}{n}{k}$ is A.

2.1. Recurrence sequences and paths

Let $(f_i)_{i\geq 0}$ be a recurrence sequence defined by

$$f_i = \alpha f_{i-1} - f_{i-2}, \qquad (i \geq 2), \tag{2.2}$$

where $\alpha \in \mathbb{N}^+$, $\alpha \geq 2$, and $f_0 < f_1$ are positive integers with $\gcd(f_0, f_1) = 1$. If $\alpha = 2$ then $(f_i)_{i\geq 0}$ is an arithmetic progression given by $f_i = f_{i-1} + (f_1 - f_0)$.

From Theorem 2.1 and Corollary 2.2 we know that in case of any positive integers $f_0 < f_1$, there exist an element in $\mathcal{HPT}_{45}$ with value f_1, and with neighbors in the same row valued by f_0 and $f_1 - f_0$. In Theorem 2.4 we give a path in $\mathcal{HPT}_{45}$ (analogously to Theorem 6 in [1]) contains all the elements of (2.2).

Theorem 2.4. *There exists a path in $\mathcal{HPT}_{45}$ crossing vertices of type A, such that the vertices are labelled with the terms of $(f_i)_{i\geq 1}$ as follows. Assume that $\rangle{}^{n}_{k}\langle = f_1$, and $\rangle{}^{n}_{k-1}\langle = f_1 - f_0$. Then the first element of the path is f_1 and the pattern of the steps from f_{i-1} to f_i $(i \geq 2)$ is $LR^{\alpha-2}$.*

Proof. According to Theorem 2.1, any f_1 and $f_1 - f_0$ can be neighbours in $\mathcal{HPT}_{45}$, where type of f_1 is A (and the type of $f_1 - f_0$ is not A).

If $\alpha = 2$, then the statment is easy to show, since no R steps. Indeed, the difference of an element type A, and its immediate left descendant having type A is the constant $f_1 - f_0$.

Assume now $\alpha \geq 3$. By the construction rule of $\mathcal{HPT}_{45}$, we can follow the way from any f_{i-1} to f_i $(i \geq 2)$ in Figure 2, which justifies the theorem (the type of the rectangle shaped elements is A). In the last row of the figure we use, among others, the equality $f_i - f_{i-1} = (\alpha - 1)f_{i-1} - f_{i-2}$. $\square$

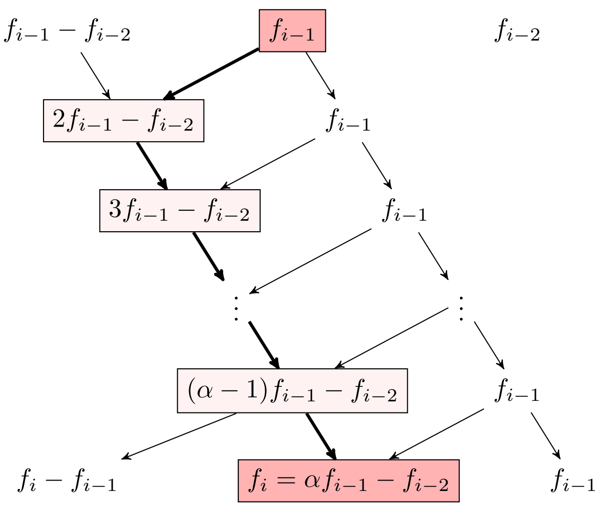


Figure 2: Path $LR^{\alpha-2}$ between f_{i-1} and f_i

Remark 2.5. Theorem 2.4 can be extended for the whole sequence $(f_i)_{i\geq 0}$ if and only if $(\alpha - 1)f_0 < f_1 < \alpha f_0$. Under these conditions one can follow the path back from the bottom of Figure 2 to the top, from f_1 to f_0.

The path showed on the right hand side of Figure 4 (cf. Figure 1) is an example for the binary recurrence $f_i = 4f_{i-1} - f_{i-2}$ with $f_0 = 1$, $f_1 = 2$.

Theorem 2.4 finds a path to the sequence (2.2). Considering the opposite direction, now we decribe the sequence corresponding to a given pattern of steps. The expression “corner element” means a labelled vertex where the direction of the sequence of steps changes. For example, the first corner element of the path L^3R^2 is the vertex reached after three left steps, the second corner element comes after further two right steps, etc.

Theorem 2.6. *Suppose that the A-type vertex* $\left|{n \atop k}\right| = U_1 = u_1$ *is a starting point of the path* $L^\ell R^r$. *We let* U_i, *and* u_i $(i = 1, \dots)$ *denote the label of the corner elements, and the label of every second corner elements of the path, respectively. Then we have*

$$u_i = (\ell r + 2)u_{i-1} - u_{i-2}, \quad (i \geq 3). \tag{2.3}$$

Moreover, if $\ell = r$, *then*

$$U_i = \ell U_{i-1} + U_{i-2}, \quad (i \geq 3).$$

Obviously, $u_i = U_{2i-1}$ holds. The proof of Theorem 2.6 applies the following lemma (see [1], Remark 1 linked to Lemma 4).

Lemma 2.7. *Let* x_0, y_0, *further* a_j *and* b_j $(j = 1, 2)$ *be complex numbers such that* $a_2b_1 \neq 0$. *Assume that for* $i \geq i_0$ *the terms of the sequences* (x_i) *and* (y_i) *satisfy*

$$\begin{aligned} x_{i+1} &= a_1x_i + b_1y_i, \\ y_{i+1} &= a_2x_i + b_2y_i. \end{aligned}$$

Then for both sequences

$$z_{i+2} = (a_1 + b_2)z_{i+1} + (-a_1b_2 + a_2b_1)z_i$$

holds $(i \geq i_0)$.

Proof of Theorem 2.6. Suppose that v_1 is the left ascendant of u_1. By Figure 3, which demonstrates the path precisely from u_i to u_{i+2} $(i \geq 1)$ along vertices type A in $\mathcal{HPT}_{45}$, we gain the system of the recursive equations

$$\begin{aligned} u_{i+1} &= (r + 1)u_i + (\ell + r(\ell - 1))\, v_i, \\ v_{i+1} &= ru_i + (\ell + (r - 1)(\ell - 1))\, v_i. \end{aligned} \tag{2.4}$$

Using Lemma 2.7 we receive that both u_i and v_i satisfy the equation

$$z_{i+2} = (\ell r + 2)z_{i+1} - z_i.$$

If $\ell = r$, then we simply obtain

$$\begin{aligned} U_{i+1} &= U_i + \ell V_i, \\ V_{i+1} &= U_i + (\ell - 1)V_i. \end{aligned} \tag{2.5}$$

Now Lemma 2.7 results that U_i and V_i satisfy the equation

$$Z_{i+2} = \ell Z_{i+1} + Z_i. \qquad \square$$

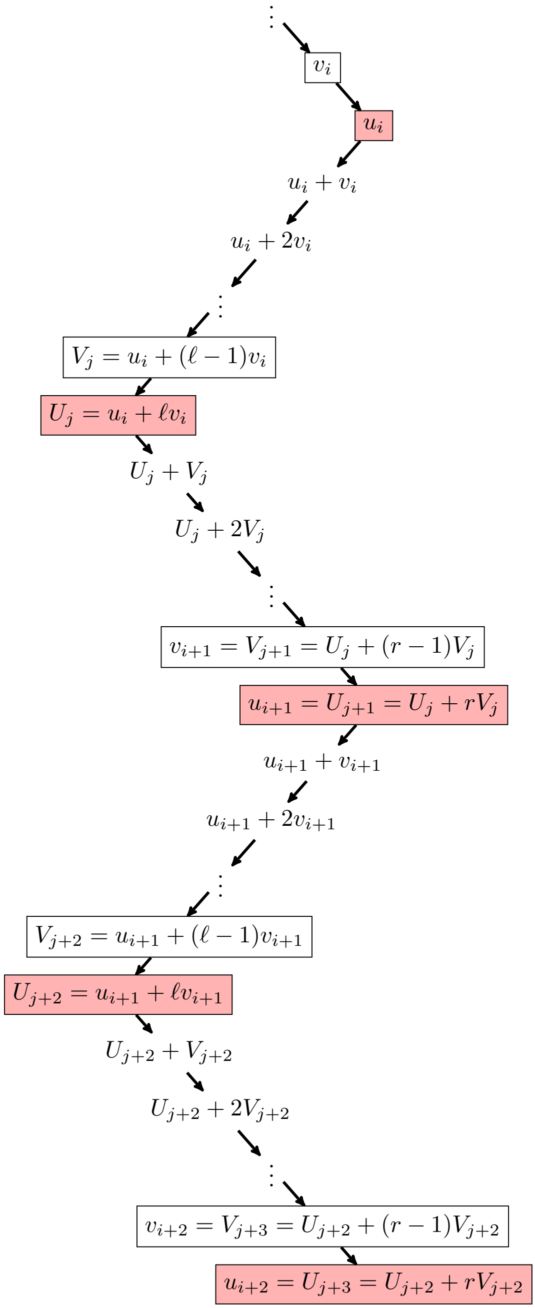


Figure 3: Path $L^\ell R^r$ from u_i to u_{i+2}

Remark 2.8. Let $\rangle{n \atop k}\langle = f_1$ be the initial element, and consider the path $L^\ell R^r$. Since every second corner element of the path satisfies the recurrence equation (2.2) with $\alpha = \ell r + 2$, the number of different paths belonging to different patterns but corresponding to the linear recurrence $(f_i)_{i=1}^{\infty}$ is the number of the divisors of $\ell r = \alpha - 2$.

Figure 4 gives an example for the case when $\alpha - 2 = 2 = 2 \cdot 1 = 1 \cdot 2$ and $u_1 = f_1 = 2$, $u_2 = f_2 = 7$. Clearly, the patters are L^2R and LR^2.

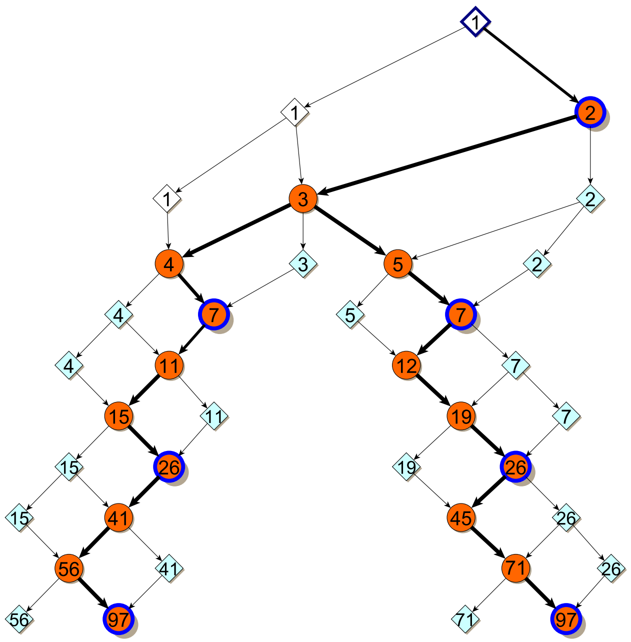


Figure 4: $f_0 = 1$, $f_1 = 2$, $f_i = 4f_{i-1} - f_{i-2}$

Now we describe the intermediate sequences located in the path given by $v_1, \rangle{n \atop k}\langle = u_1$ and by $L^\ell R^r$. The labels of the elements having distance $(\ell + r)t$ $(t \in \mathbb{N})$ from the base element $\rangle{n \atop k}\langle$ are given by suitable sequences $\{w_i\}$.

Theorem 2.9. *Put $w_i = u_i + mv_i$, where $0 \le m < \ell$, or let $w_i = U_{2i} + mV_{2i} = (m+1)u_i + (\ell + m(\ell - 1))v_i$, where $0 \le m < r$. Then the terms of the sequence (w_i) satisfy*

$$w_i = (\ell r + 2)w_{i-1} - w_{i-2}, \quad (i \ge 3).$$

Proof. Consider again Figure 3 to show the statement for the first type of sequences. One can observe the labels of the path described by $w_i = u_i + mv_i$, where $0 \le m < \ell$ and $i \ge 1$. From (2.3) we see

$$w_{i+2} = u_{i+2} + mv_{i+2} = (\ell r + 2)(u_{i+1} + kv_{i+1}) - (u_i + mv_i)$$

$$= (\ell r + 2)w_{i+1} - w_i.$$

The second part of the proof is analoguous. In Figure 3 the equation $j = 2i$ holds, but generally it does not. □

Corollary 2.10. *In case of $\ell = r$, $W_j = U_j + mV_j$ $(0 \le m < \ell)$ satisfy the equation*

$$W_j = \ell W_{j-1} + W_{j-2}, \quad (j \ge 3). \tag{2.6}$$

In Figure 5, according to Corollary 2.10 we give two examples for the representation of elements of recurrence sequence $f_i = 3f_{i-1} + f_{i-2}$. The pattern of both paths is R^3L^3, moreover, $u_1 = 3$, $v_1 = 2$, $m = 2$ and $u_1 = 4$, $v_1 = 3$, $m = 1$, respectively.

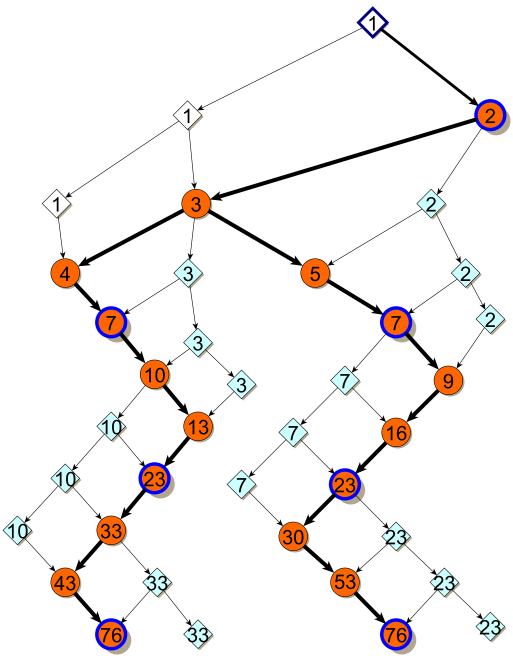


Figure 5: $f_0 = 1$, $f_1 = 2$, $f_i = 3f_{i-1} + f_{i-2}$

Theorem 2.11. *Consider the sequence* (2.2)*. If $\ell r = \alpha - 2$, and*

$$m = \frac{f_{j+1} - (r+1)f_j}{\ell + r(\ell - 1)}$$

is an integer for some $j \ge 1$, further $m < f_j$ holds, then the elements f_i $(i \ge j)$ can be represented in $\mathcal{HPT}_{45}$ by every second corner elements of a paths given by the the pattern $L^\ell R^r$, and by $u_1 = f_j$ and $v_1 = m$.

Proof. Let $u_1 = f_j$ and $u_2 = f_{j+1}$. Then equation (2.4) yields $v_1 = (f_{j+1} - (r + 1)f_j)/(\ell + r(\ell - 1))$. Since the integers v_1 and u_1 are neigbours in a suitable row of $\mathcal{HPT}_{45}$, therefore there is a path with the pattern $L^\ell R^r$ from v_1 and u_1 such that every second corner elements are f_{i+1} $(i \geq j)$. □

Figure 4 gives examples on the paths of $f_i = 4f_{i-1} - f_{i-2}$ with initial elements $u_1 = f_2 = 7$ and $u_2 = f_3 = 26$, moreover $v_1 = 4$ and $v_1 = 5$, respectively, where $\alpha - 2 = 2 = 2 \cdot 1 = 1 \cdot 2 = lr$, and the patters are L^2R and LR^2.

Theorem 2.12. *Consider now the sequence* (2.1)*. If $\ell^2 = \eta - 2$, and $m = (f_{j+1} - f_j)/\ell$ is an integer, further $m < f_j$, then the elements f_i $(i \geq j \geq 1)$ can be represented in $\mathcal{HPT}_{45}$ by every corner elements of the paths given by the pattern $L^\ell R^r$, and by $u_1 = f_j$ and $v_1 = m$.*

Proof. The proof is similar to the proof of Theorem 2.11. Using (2.5), from $u_1 = f_j$ and $u_2 = f_{j+1}$ we gain $v_1 = (f_{j+1} - f_j)/\ell$. □

References

[1] H. Belbachir, L. Németh, L. Szalay, Hyperbolic Pascal triangles, *Appl. Math. Comp.*, **273** (2016), 453–464.

[2] L. Németh, L. Szalay, Alternating sums in hyperbolic Pascal triangles, *Miskolc Mathematical Notes*, (accepted).

[3] H. Belbachir, L. Szalay, On the arithmetic triangles, *Šiauliai Math. Sem.*, **9** (17) (2014), 15–26.

[4] H. S. M. Coxeter, Regular honeycombs in hyperbolic space, in *Proc. Int. Congress Math.*, Amsterdam, Vol. III., 1954, pp. 155–169.

Annales Mathematicae et Informaticae
46 (2016) pp. 175–193
http://ami.ektf.hu

Generalized r-Whitney numbers of the first kind

José L. Ramírez[a], Mark Shattuck[b]

[a]Departamento de Matemáticas
Universidad Sergio Arboleda, Bogotá, Colombia
josel.ramirez@ima.usergioarboleda.edu.co

[b]Department of Mathematics
University of Tennessee, Knoxville, TN, USA
shattuck@math.utk.edu

Submitted May 15, 2016 — Accepted December 1, 2016

Abstract

In this paper, we consider a (p, q)-generalization of the r-Whitney number sequence of the first kind that reduces to it when $p = q = 1$. We obtain generalizations of some earlier results for the r-Whitney sequence, including recurrence and generating function formulas. We develop a combinatorial interpretation for our generalized numbers in terms of a pair of statistics on the set of r-permutations in which the elements within cycles of a permutation are assigned colors according to certain rules. This allows one to provide combinatorial proofs of various identities, including orthogonality relations. Finally, we consider the (p, q)-Whitney matrix of the first kind and find various factorizations for it.

Keywords: q-generalization, r-Whitney numbers, Stirling numbers, Whitney matrix

MSC: 05A19, 05A15, 60C05, 15A23

1. Introduction

We will make use of the following notation. If m and n are positive integers, then let $[m, n] = \{m, m + 1, \ldots, n\}$ if $m \leq n$, with $[m, n] = \varnothing$ if $m > n$. We will denote the special case $[1, n]$ by $[n]$. Given a positive integer k and an indeterminate q, let

$[k]_q = 1 + q + \cdots + q^{k-1}$, with $[0]_q = 0$. Throughout, empty sums will assume the value 0, and empty products the value 1.

Suppose $r \geq 0$ and $m \geq 1$ are given integers. Let $w(n,k) = w(n,k;r,m)$ denote the r-Whitney numbers of the first kind (see, e.g., [7, 15]), which are defined as connection constants in the polynomial identities

$$m^n x(x-1)\cdots(x-n+1) = \sum_{k=0}^{n} w(n,k)(mx+r)^k, \quad n \geq 0.$$

See also [1, 18] for further properties of the $w(n,k)$. Equivalently, the $w(n,k)$ array is determined by the recurrence

$$w(n,k) = w(n-1,k-1) - (r + m(n-1))w(n-1,k), \quad n,k \geq 1,$$

with initial values $w(n,0) = (-1)^n \prod_{i=0}^{n-1}(r+mi)$ and $w(0,k) = \delta_{k,0}$ for $n,k \geq 0$. Note that $w(n,k;0,1) = s(n,k)$ and $w(n,k;1,1) = s(n+1,k+1)$, where $s(n,k)$ is the Stirling number of the first kind. In [10], a combinatorial interpretation for $w(n,k)$ when $r = 0$ is given in terms of the coefficients of the characteristic polynomial for the rank function on the Dowling lattice of rank n over a finite group of order m. An interpretation for $w(n,k)$ was given in [7] for arbitrary r.

Here, we consider a polynomial generalization of the r-Whitney numbers of the first kind, which we will denote by $w_{p,q}(n,k) = w_{p,q}(n,k;r,m)$. It is defined by the recurrence

$$w_{p,q}(n,k) = w_{p,q}(n-1,k-1) - ([r]_p + m[n-1]_q)w_{p,q}(n-1,k), \quad n,k \geq 1, \tag{1.1}$$

with initial values $w_{p,q}(n,0) = (-1)^n \prod_{i=0}^{n-1}([r]_p + m[i]_q)$ and $w_{p,q}(0,k) = \delta_{k,0}$ for $n,k \geq 0$. Note that $w_{1,1}(n,k) = w(n,k)$ for all n and k. The numbers $w_{p,q}(n,k)$ when $p = 1$ differ slightly from the (q,r)-Whitney numbers $w_{m,r,q}(n,k)$ studied in [12] due to the absence of extra factors of q in the defining recurrence. In contrast to [12], where identities for $w_{m,r,q}(n,k)$ were shown by algebraic methods using q-boson operators, we provide a combinatorial interpretation for our $w_{p,q}(n,k)$ which allows one to explain identities bijectively. Moreover, on account of its simpler recurrence, it seems that the numbers $w_{p,q}(n,k)$ provide a more natural combinatorial generalization of the r-Whitney numbers than those studied in [12] which arose in a physical setting. Furthermore, the $w(n,k)$ form an orthogonal pair with the generalized version of the r-Whitney numbers of the second kind considered in [13]. This orthogonality relation generalizes one between the r-Whitney numbers of the first and second kind (see [7]).

In the next section, we give several algebraic properties satisfied by $w_{p,q}(n,k)$, including a connection between it and the elementary symmetric functions. We then develop in the third section a combinatorial interpretation for $w_{p,q}(n,k)$ in terms of statistics on a structure $\mathcal{A}$ enumerated by $|w(n,k)|$ and use this to provide bijective proofs of identities satisfied by $w_{p,q}(n,k)$, including orthogonality relations. The p- and q- variables will be seen here to play different combinatorial roles, with the former marking a statistic on $\mathcal{A}$ related only to those cycles containing the elements of $[r]$ no two of which are to belong to the same cycle, while the latter marks a statistic on $\mathcal{A}$ related to the positions of the elements of $[r+1, r+n]$ within all cycles. In the case $r = 0$ and $m = 1$, these statistics appear to be new on the set of permutations. In the final section, we consider the (p, q)-Whitney matrix of the first kind and find some factorizations of this matrix in analogy with the results of [16].

Let $W(n, k) = W(n, k; r, m)$ denote the r-Whitney number of the second kind (see [7]). We now recall a (p, q)-generalization of $W(n, k)$ from [13] defined by the recurrence

$$W_{p,q}(n, k) = W_{p,q}(n-1, k-1) + ([r]_p + m[k]_q)W_{p,q}(n-1, k), \quad n, k \geq 1,$$

with initial values $W_{p,q}(n, 0) = [r]_p^n$ and $W_{p,q}(0, k) = \delta_{k,0}$. Among the results, it was shown that the $W_{p,q}(n, k)$ are determined by the identities

$$(mx + [r]_p)^n = \sum_{k=0}^{n} W_{p,q}(n, k) m^k [x]_q^{\underline{k}}, \quad n \geq 0, \tag{1.2}$$

where

$$[x]_q^{\underline{n}} = \begin{cases} x(x-[1]_q)\cdots(x-[n-1]_q), & \text{if } n \geq 1; \\ 1, & \text{if } n = 0, \end{cases}$$

or equivalently by the generating function

$$\sum_{n \geq k} W_{p,q}(n, k) x^n = \frac{x^k}{(1-([r]_p + m[0]_q)x)\cdots(1-([r]_p + m[k]_q)x)}, \quad k \geq 0. \tag{1.3}$$

Various connections will be made between $w_{p,q}(n, k)$ and $W_{p,q}(n, k)$ in the next two sections. Note that $W_{p,q}(n, k)$ reduces to $W(n, k)$ when $p = q = 1$. Let us denote by $s_q(n, k)$ and $S_q(n, k)$ the $r = 0, m = p = 1$ case of $w_{p,q}(n, k)$ and $W_{p,q}(n, k)$, respectively. The $s_q(n, k)$ and $S_q(n, k)$ are q-Stirling polynomials of the first and second kind which were originally considered by Carlitz [5, 6] and have since been studied [20, 21] (see also [8, 9] for a related generalization).

2. Identities of the generalized r-Whitney numbers

In this section, we prove various algebraic properties of the array $w_{p,q}(n, k)$. We first show that the $w_{p,q}(n, k)$ serve as connection constants as follows.

Theorem 2.1. *If $n \geq 0$, then*

$$m^n [x]_q^{\underline{n}} = \sum_{k=0}^{n} w_{p,q}(n, k)(mx + [r]_p)^k. \tag{2.1}$$

Proof. We proceed by induction on n. The equality clearly holds for $n = 0$. Now assume that the claim holds for n, and let us prove it for $n + 1$. From recurrence (1.1), we have

$$\begin{aligned}
&\sum_{k=0}^{n+1} w_{p,q}(n+1,k)(mx+[r]_p)^k = \sum_{k=0}^{n} w_{p,q}(n+1,k)(mx+[r]_p)^k + (mx+[r]_p)^{n+1}\\
&= \sum_{k=0}^{n} w_{p,q}(n,k-1)(mx+[r]_p)^k - ([r]_p + m[n]_q)\sum_{k=0}^{n} w_{p,q}(n,k)(mx+[r]_p)^k\\
&\quad + (mx+[r]_p)^{n+1}\\
&= \sum_{k=0}^{n-1} w_{p,q}(n,k)(mx+[r]_p)^{k+1} - ([r]_p + m[n]_q)m^n[x]_q^{\underline{n}} + (mx+[r]_p)^{n+1}\\
&= \sum_{k=0}^{n} w_{p,q}(n,k)(mx+[r]_p)^{k+1} - ([r]_p + m[n]_q)m^n[x]_q^{\underline{n}}\\
&= (mx+[r]_p)m^n[x]_q^{\underline{n}} - ([r]_p + m[n]_q)m^n[x]_q^{\underline{n}}\\
&= m^{n+1}[x]_q^{\underline{n+1}},
\end{aligned}$$

which completes the induction. □

We next give the generating function of the array $w_{p,q}(n,k)$ for fixed n.

Theorem 2.2. *If $n \geq 0$, then*

$$\sum_{k=0}^{n} w_{p,q}(n,n-k)x^k = \prod_{k=0}^{n-1}(1-([r]_p + m[k]_q)x). \tag{2.2}$$

Proof. We proceed by induction on n. The equality clearly holds for $n = 0$. Now assume that the claim holds for n, and let us prove it for $n + 1$. From recurrence (1.1), we have

$$\begin{aligned}
&\sum_{k=0}^{n+1} w_{p,q}(n+1,n+1-k)x^k\\
&= \sum_{k=0}^{n} w_{p,q}(n,n-k)x^k - ([r]_p + m[n]_q)\sum_{k=0}^{n+1} w_{p,q}(n,n+1-k)x^k\\
&= \prod_{k=0}^{n-1}(1-([r]_p + m[k]_q)x) - ([r]_p + m[n]_q)\sum_{k=0}^{n} w_{p,q}(n,n-k)x^{k+1}\\
&= \prod_{k=0}^{n-1}(1-([r]_p + m[k]_q)x) - ([r]_p + m[n]_q)x\prod_{k=0}^{n-1}(1-([r]_p + m[k]_q)x)\\
&= \prod_{k=0}^{n}(1-([r]_p + m[k]_q)x),
\end{aligned}$$

which completes the induction. □

Given a set of variables $x_1, x_2, \ldots, x_n$, the k-th elementary and complete symmetric functions are defined, respectively, by

$$e_k(x_1, x_2, \ldots, x_n) = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}, \quad 1 \le k \le n,$$
$$h_k(x_1, x_2, \ldots, x_n) = \sum_{1 \le i_1 \le i_2 \le \cdots \le i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}, \quad k \ge 1,$$

with initial conditions $e_0(x_1, x_2, \ldots, x_n) = h_0(x_1, x_2, \ldots, x_n) = 1$. Note that $e_k(x_1, x_2, \ldots, x_n) = 0$ if $k > n$. The generating functions for the e_k and h_k are given by

$$\sum_{k=0}^{n} e_k(x_1, x_2, \ldots, x_n) z^k = \prod_{i=1}^{n} (1 + x_i z),$$
$$\sum_{k \ge 0} h_k(x_1, x_2, \ldots, x_n) z^k = \prod_{i=1}^{n} \frac{1}{1 - x_i z}.$$

Using (2.2) and (1.3), it is not difficult to show that the (p, q)-Whitney numbers are the specializations of the elementary and complete symmetric functions given by

$$\begin{aligned} & w_{p,q}(n+1, n+1-k) \\ & = (-1)^k e_k([r]_p, m[1]_q + [r]_p, m[2]_q + [r]_p, \ldots, m[n]_q + [r]_p), \end{aligned} \tag{2.3}$$

$$\begin{aligned} & W_{p,q}(n+k, n) \\ & = h_k([r]_p, m[1]_q + [r]_p, m[2]_q + [r]_p, \ldots, m[n]_q + [r]_p). \end{aligned} \tag{2.4}$$

In particular, the q-Stirling numbers of the first and second kind satisfy

$$s_q(n+1, n+1-k) = (-1)^k e_k([1]_q, [2]_q, \ldots, [n]_q),$$
$$S_q(n+k, n) = h_k([1]_q, [2]_q, \ldots, [n]_q).$$

Lemma 2.3 (Merca [14])**.** *Let k and n be positive integers. Then*

$$f_k(t + x_1, t + x_2, \ldots, t + x_n) = \sum_{i=0}^{k} \binom{n - c_i}{k - i} f_i(x_1, x_2, \ldots, x_n) t^{k-i}$$

and

$$f_k(x_1, x_2, \ldots, x_n) = \sum_{i=0}^{k} \binom{n - c_i}{k - i} f_i(t + x_1, t + x_2, \ldots, t + x_n)(-t)^{k-i},$$

where $t, x_1, x_2, \ldots, x_n$ are variables, f_i is either the i-th elementary or complete symmetric function for all i, and

$$c_i = \begin{cases} i, & \text{if } f_i = e_i; \\ 1-k, & \text{if } f_i = h_i. \end{cases}$$

Using the prior lemma, one can obtain the following formulas for the (p,q)-Whitney numbers.

Proposition 2.4. *If $r \geq s \geq 0$, then*

$$\begin{aligned} w_{p,q}(n,k;r,m) &= \sum_{i=k}^{n} \binom{i}{k} w_{p,q}(n,i;s,m)([s]_p - [r]_p)^{i-k} \qquad (2.5)\\ &= \sum_{i=k}^{n} (-1)^{i-k} p^{s(i-k)} \binom{i}{k} w_{p,q}(n,i;s,m)[r-s]_p^{i-k} \end{aligned}$$

and

$$\begin{aligned} W_{p,q}(n,k;r,m) &= \sum_{i=k}^{n} \binom{n}{i} W_{p,q}(i,k;s,m)([r]_p - [s]_p)^{n-i} \qquad (2.6)\\ &= \sum_{i=k}^{n} p^{s(n-i)} \binom{n}{i} W_{p,q}(i,k;s,m)[r-s]_p^{n-i}. \end{aligned}$$

Proof. By (2.3), (2.4), and Lemma 2.3, we have

$$\begin{aligned} &w_{p,q}(n,n-k;r,m) = (-1)^k e_k([r]_p, m[1]_q + [r]_p, \ldots, m[n-1]_q + [r]_p)\\ &= (-1)^k \sum_{i=0}^{k} \binom{n-i}{k-i} e_i([s]_p, m[1]_q + [s]_p, \ldots, m[n-1]_q + [s]_p)([r]_p - [s]_p)^{k-i}\\ &= \sum_{i=0}^{k} \binom{n-i}{k-i} (-1)^i e_i([s]_p, m[1]_q + [s]_p, \ldots, m[n-1]_q + [s]_p)([s]_p - [r]_p)^{k-i}\\ &= \sum_{i=0}^{k} \binom{n-i}{k-i} w_{p,q}(n,n-i;s,m)([s]_p - [r]_p)^{k-i} \end{aligned}$$

and

$$\begin{aligned} &W_{p,q}(n+k,n;r,m) = h_k([r]_p, m[1]_q + [r]_p, \ldots, m[n]_q + [r]_p)\\ &= \sum_{i=0}^{k} \binom{n+k}{k-i} h_i([s]_p, m[1]_q + [s]_p, \ldots, m[n]_q + [s]_p)([r]_p - [s]_p)^{k-i}\\ &= \sum_{i=0}^{k} \binom{n+k}{k-i} W_{p,q}(n+i,n;s,m)([r]_p - [s]_p)^{k-i}, \end{aligned}$$

from which the identities may be obtained. $\square$

Taking $p = q = 1$ in (2.6) gives the following identity which was shown previously by a different method using Riordan matrix groups.

Corollary 2.5 (Cheon and Jung [7]). *If $r \geq s \geq 0$, then*

$$W(n,k;r,m) = \sum_{i=k}^{n} \binom{n}{i} W(i,k;s,m)(r-s)^{n-i}.$$

The following proposition shows how to express the (p,q)-Whitney numbers of both kinds in terms of the q-Stirling numbers and vice-versa.

Proposition 2.6. *If $n, k \geq 0$, then*

$$w_{p,q}(n,k) = \sum_{i=k}^{n} m^{n-i} \binom{i}{k} (-[r]_p)^{i-k} s_q(n,i), \tag{2.7}$$

$$s_q(n,k) = \frac{1}{m^{n-k}} \sum_{i=k}^{n} \binom{i}{k} [r]_p^{i-k} w_{p,q}(n,i), \tag{2.8}$$

$$W_{p,q}(n,k) = \sum_{i=k}^{n} m^{i-k} \binom{n}{i} [r]_p^{n-i} S_q(i,k), \tag{2.9}$$

$$S_q(n,k) = \frac{1}{m^{n-k}} \sum_{i=k}^{n} \binom{n}{i} (-[r]_p)^{n-i} W_{p,q}(i,k). \tag{2.10}$$

Proof. To show (2.9), note that

$$\begin{aligned}
W_{p,q}(n+k,n) &= h_k([r]_p, m[1]_q + [r]_p, \ldots, m[n]_q + [r]_p) \\
&= \sum_{i=0}^{k} \binom{n+k}{k-i} h_i(m[1]_q, m[2]_q, \ldots, m[n]_q)[r]_p^{k-i} \\
&= \sum_{i=0}^{k} \binom{n+k}{k-i} m^i h_i([1]_q, [2]_q, \ldots, [n]_q)[r]_p^{k-i} \\
&= \sum_{i=0}^{k} \binom{n+k}{k-i} m^i S_q(n+i,n)[r]_p^{k-i},
\end{aligned}$$

and then replace n by $n-k$, k by $n-k$, and i by $i-k$ in that order.

For (2.10), observe that

$$\begin{aligned}
S_q(n+k,n) &= h_k([1]_q, [2]_q, \ldots, [n]_q) \\
&= \frac{1}{m^k} h_k(m[1]_q, m[2]_q, \ldots, m[n]_q) \\
&= \sum_{i=0}^{k} \frac{1}{m^k} \binom{n+k}{k-i} h_i([r]_p, m[1]_q + [r]_p, \ldots, m[n]_q + [r]_p)(-[r]_p)^{k-i}
\end{aligned}$$

$$= \sum_{i=0}^{k} \frac{1}{m^k} \binom{n+k}{k-i} W_{p,q}(n+i,n)(-[r]_p)^{k-i}.$$

The proofs of (2.7) and (2.8) are similar. □

Let $s(n,k)$ and $S(n,k)$ denote the Stirling numbers of the first and second kind, respectively. Taking $p = q = 1$ in (2.7) and (2.9) gives the following formulas.

Corollary 2.7 (Cheon and Jung [7]). *If $n, k \geq 0$, then*

$$w(n,k) = \sum_{i=k}^{n} m^{n-i} \binom{i}{k} (-r)^{i-k} s(n,i)$$

and

$$W(n,k) = \sum_{i=k}^{n} m^{i-k} \binom{n}{i} r^{n-i} S(i,k).$$

The $p = q = r = 1$ case of (2.8) is also previously known.

Corollary 2.8 (Benoumhani [2]). *If $n, k \geq 0$, then*

$$s(n,k) = \frac{1}{m^{n-k}} \sum_{i=k}^{n} \binom{i}{k} w(n,i;1,m).$$

3. Combinatorial interpretation and properties

In this section, we develop a combinatorial interpretation for the array $w_{p,q}(n,k)$ and use it to explain various relations that it satisfies, including its orthogonality with $W_{p,q}(n,k)$. We first recall the concept of an r-permutation, see [3].

Definition 3.1. Given $0 \leq r \leq m$, by an r-permutation of $[m]$, it is meant a member of $\mathcal{S}_m$ in which the elements of $[r]$ belong to distinct cycles. If $n, k, r \geq 0$, then let $\Omega_r(n,k)$ denote the set of all r-permutations of $[n+r]$ having exactly $k+r$ cycles and let $\Omega_r(n) = \cup_{k=0}^{n} \Omega_r(n,k)$.

When $r = 0$, a member of $\Omega_r(n)$ is the same as an ordinary permutation of $[n]$. Note that the cardinality of $\Omega_r(n,k)$ is given by the (signless) r-Stirling number of the first kind (see, e.g., [3]), while the cardinality of $\Omega_r(n)$ is seen to be $(r+1)^{\overline{n}}$.

Within a member of $\Omega_r(n,k)$, we will refer to the cycles containing an element of $[r]$ as *special* and to the remaining cycles comprised exclusively of elements of $I = [r+1, r+n]$ as *non-special.* (The members of $[r]$ themselves will also at times be described as *special.*) In addition, we will refer to an element within a member of $\Omega_r(n,k)$ that is the smallest within its cycle as *minimal*, and to all other elements as *non-minimal.* Throughout, we will assume that members of $\Omega_r(n,k)$ are expressed in *standard cycle form*, i.e., minimal elements first within each cycle, with cycles arranged left-to-right in ascending order of minimal elements.

We now consider a certain subset of the elements within a permutation expressed in standard form.

Definition 3.2. Suppose $\sigma \in \Omega_r(n)$ is in standard cycle form and $i \in I$, with i not the first element of a cycle of σ. Consider the word w obtained by writing all elements of the cycle C containing i, except for the first, left-to-right as they appear within C. Then we will say that i is a left-to-right cycle minimum (l-r cycle min) if i is a left-to-right minimum within w in the usual sense.

For example, let

$$\sigma = (1, 7, 13, 12, 4, 15)(2, 6, 10, 8, 5)(3, 9)(11, 14) \in \Omega_3(12, 1).$$

Then the first three cycles are special, the final cycle is non-special, and the l-r cycle min are $7, 4, 6, 5, 9, 14$. Note that the second element and the second smallest element within a cycle are always l-r cycle min, by definition. We now allow for certain elements within an r-permutation to be colored.

Definition 3.3. Given a positive integer m, let $\Omega_{r,m}(n, k)$ denote the set of r-permutations of $[n + r]$ having $k + r$ cycles in which elements within the following two classes are each assigned one of m colors: (i) non-minimal elements within non-special cycles, and (ii) non-minimal elements within special cycles that do not correspond to left-to-right cycle minima. Define $\Omega_{r,m}(n) = \cup_{k=0}^{n} \Omega_{r,m}(n, k)$.

Within the permutation σ above, the elements that would be assigned colors are (i) 14 and (ii) $13, 12, 15, 10, 8$.

Let $v(n, k) = v(n, k; r, m) = |\Omega_{r,m}(n, k)|$; note that $v(n, k) = (-1)^{n-k} w(n, k)$ upon comparing recurrences and initial values. See also Mihoubi and Rahmani [17] for an interpretation of $v(n, k; r, m)$ in terms of their partial r-Bell polynomials. In the formulation above, one may also regard m as an indeterminate marking the statistic on $\Omega_r(n)$ that records the sum of the number of non-minimal elements in non-special cycles and the number of non-minimal elements in special cycles not corresponding to l-r cycle min. For a combinatorial interpretation of $w(n, k)$ in terms of Dowling lattices, the reader is referred to [7, Section 2].

Definition 3.4. Suppose $\sigma \in \Omega_{r,m}(n)$ and that $i \in I$ belongs to cycle C of σ, with i not the first element of C. Then the predecessor of i is the first element of I to the left of i in C and smaller than i, provided such an element exists, which we will denote by *pred*(i). Define S_σ to be the set of all $i \in I$ that have a predecessor (possibly empty).

Observe that all non-minimal elements in non-special cycles of $\sigma \in \Omega_{r,m}(n)$ have predecessors, whereas only non-minimal elements not corresponding to l-r cycle min in special cycles have them. For example, if

$$\sigma = (1, 6, 4, 5)(2, 8, 7, 9)(3, 11, 13, 14)(10, 12)(15) \in \Omega_{3,1}(12, 2),$$

then we have $S_\sigma = \{5, 9, 12, 13, 14\}$. Given $\sigma \in \Omega_{r,m}(n)$ and $1 \le i \le r$, let ℓ_i denote the number of l-r cycle min within the i-th special cycle of σ. In the last example, we have $r = 3$, with $\ell_1 = \ell_2 = 2$ and $\ell_3 = 1$.

We now introduce a pair of statistics on the set $\Omega_{r,m}(n)$.

Definition 3.5. Define the statistics v_1 and v_2 on $\Omega_{r,m}(n)$ by letting

$$v_1(\sigma) = \sum_{i=1}^{r}(i-1)\ell_i$$

and

$$v_2(\sigma) = \sum_{i \in S_\sigma} (pred(i) - r - 1).$$

Note that the statistic v_1 appears to be new even in the case $r = 0$ and $m = 1$, though in this case it has the same distribution on $\mathcal{S}_n$ as a certain type of inversion statistic originally considered by Carlitz [6] and later studied [20]. We found no special cases in the literature of the statistic v_2. We now consider a (p,q)-generalization of the r-Whitney number of the first kind in terms of these statistics.

Definition 3.6. Define $v_{p,q}(n,k) = v_{p,q}(n,k;r,m)$ as the joint distribution polynomial for the v_1 and v_2 statistics on the set $\Omega_{r,m}(n,k)$, that is,

$$v_{p,q}(n,k) = \sum_{\sigma \in \Omega_{r,m}(n,k)} p^{v_1(\sigma)} q^{v_2(\sigma)}, \quad n,k \geq 0.$$

The $v_{p,q}(n,k)$ are determined recursively as follows.

Proposition 3.7. *The array $v_{p,q}(n,k)$ satisfies the recurrence*

$$v_{p,q}(n,k) = v_{p,q}(n-1,k-1) + ([r]_p + m[n-1]_q)v_{p,q}(n-1,k), \quad n,k \geq 1, \tag{3.1}$$

and has initial values $v_{p,q}(n,0) = \prod_{i=0}^{n-1}([r]_p + m[i]_q)$ and $v_{p,q}(0,k) = \delta_{k,0}$ for $n,k \geq 0$.

Proof. The initial condition $v_{p,q}(0,k) = \delta_{k,0}$ is clear from the definitions. To show $v_{p,q}(n,0) = \prod_{i=0}^{n-1}([r]_p + m[i]_q)$, we add the elements of I sequentially to the special cycles starting with $r+1$. The element $r+i$ contributes a factor of $[r]_p + m[i-1]_q$, upon considering whether it is inserted directly following a member of $[r]$ or a member of $[r+1, r+i-1]$; note that there are $1+p+\cdots+p^{r-1} = [r]_p$ possibilities in the former case where $r+i$ would correspond to a l-r cycle min and $m(1+q+\cdots+q^{i-2}) = m[i-1]_q$ possibilities in the latter. To show (3.1), first observe that the weight of all members of $\Omega = \Omega_{r,m}(n,k)$ in which $n+r$ belongs to its own cycle is $v_{p,q}(n-1,k-1)$, since neither the v_1 nor the v_2 statistic values are changed by its addition in this case. On the other hand, the weight of all members of Ω in which $n+r$ is a l-r cycle min within a special cycle is given by $[r]_p v_{p,q}(n-1,k)$. Finally, members of Ω in which $n+r$ directly follows some member of $[r+1, r+n-1]$ within a cycle are seen to have weight $m[n-1]_q v_{p,q}(n-1,k)$. Observe that the addition of $n+r$ to a cycle does not affect the predecessors of smaller elements already occupying the cycle. Combining the three previous cases gives (3.1). $\square$

Note that $w_{p,q}(n,k) = (-1)^{n-k} v_{p,q}(n,k)$, upon comparing recurrences. One has the following further recurrence satisfied by $w_{p,q}(n,k)$.

Proposition 3.8. *If $n, k \geq 1$, then*

$$w_{p,q}(n,k) = \sum_{j=k}^{n} (-1)^{n-j} w_{p,q}(j-1, k-1) \prod_{i=0}^{n-j-1} ([r]_p + m[j+i]_q).$$

Proof. We show, equivalently, the relation

$$v_{p,q}(n,k) = \sum_{j=k}^{n} v_{p,q}(j-1, k-1) \prod_{i=0}^{n-j-1} ([r]_p + m[j+i]_q). \tag{3.2}$$

To do so, consider the smallest element, $r + j$, within the k-th non-special cycle of a member of $\Omega_{r,m}(n,k)$; note that $k \leq j \leq n$. Then the elements of $[r+j-1]$ may be positioned according to any member of $\Omega_{r,m}(j-1, k-1)$, and thus there are $v_{p,q}(j-1, k-1)$ possibilities concerning their arrangement. After placing the element $r+j$ in its own cycle, we insert the members of $[r+j+1, r+n]$ one-by-one starting with $r+j+1$. For $1 \leq i \leq n-j$, there are $[r]_p + m[j+i-1]_q$ possibilities concerning the placement of the element $r+j+i$, upon considering whether it directly follows a member of $[r]$ or a member of $[r+1, r+j+i-1]$. Thus, there are $\prod_{i=1}^{n-j}([r]_p + m[j+i-1]_q)$ possibilities concerning the placement of elements of $[r+j+1, r+n]$. Summing over j gives (3.2) and completes the proof. $\square$

Using our combinatorial interpretation for $w_{p,q}(n,k)$, it is possible to prove bijectively the formulas for $w_{p,q}(n,k)$ and $s_q(n,k)$ given above in Proposition 2.6.

Combinatorial proofs of (2.7) *and* (2.8) *in Proposition 2.6.* We first prove formula (2.7), rewritten in the form

$$v_{p,q}(n,k) = \sum_{j=k}^{n} m^{n-j} \binom{j}{k} [r]_p^{j-k} c_q(n,j), \quad n \geq k \geq 0, \tag{3.3}$$

where $c_q(n,j) = (-1)^{n-j} s_q(n,j)$. To show (3.3), we count members of $\Omega_{r,m}(n,k)$ according to the number, $j-k$, of l-r cycle min in special cycles, where $k \leq j \leq n$. To form a member of $\Omega_{r,m}(n,k)$ for which the number of l-r cycle min in special cycles is $j-k$, we first consider $\rho \in \Omega_{0,1}(n,j)$ in standard cycle form, i.e., ρ is a permutation of $[n]$ having j cycles, and count all such ρ according to the value of the v_2 statistic. Note that there are $c_q(n,j)$ possibilities for ρ, each of whose $n-j$ non-minimal elements is assigned one of m colors. Next, we add r to all of the letters of ρ. We then select $j-k$ of the j cycles of ρ, remove the enclosing parentheses, and let $w_1, w_2, \ldots, w_{j-k}$ denote the resulting words, where $\min(w_1) < \min(w_2) < \cdots < \min(w_{j-k})$.

We insert the words w_i into r urns labeled $1, 2, \ldots, r$. Assign the weight of p^{i-1} to each word added to the i-th urn for $1 \leq i \leq r$, which we multiply to obtain the total weight. Thus, there are $[r]_p^{j-k}$ possibilities concerning the placement of the words w_i. Within urns, words are ordered from left-to-right in descending order of first (= smallest) elements and then concatenated, with the number labeling the urn written at the beginning. That is, if $w_{i_1}, \ldots, w_{i_s}$, with $i_1 < \cdots < i_s$, are the words in urn j, we form the long word $jw_{i_s} \cdots w_{i_1}$. The contents of urn j then becomes that of the j-th special cycle. Note that the first letter of each w_i becomes a l-r cycle min, by the ordering of words within urns. Taken together with the k cycles of ρ that were not selected, we obtain a member $\pi \in \Pi_{r,m}(n,k)$ in which the l-r cycle min in special cycles number $j-k$. Upon considering all possible ρ, the weight of such members of $\Pi_{r,m}(n,k)$ is seen to be $m^{n-j}\binom{j}{k}[r]_p^{j-k}c_q(n,j)$. Summing over all j then gives (3.3).

We illustrate the above procedure for transforming ρ into π, where $n = 20$, $k = 2$, $r = 4$ and $m = 1$. Let $j = 8$ and

$$\rho = (1,7,3)(2)(4,18,9,5)(6)(8,10,20,11)(12,16,13,19)(14)(15,17) \in \Omega_{0,1}(20,8).$$

Increase each element of ρ by $r = 4$ and then, in the resulting permutation of $[5,24]$, select $j-k = 6$ of the cycles, shown below:

$$(5,11,7), \quad (8,22,13,9), \quad (10), \quad (12,14,24,15), \quad (16,20,17,23), \quad (18),$$

which will be denoted by w_i, $1 \le i \le 6$, from left to right. Insert these words randomly into four urns U_i as shown:

$$\begin{array}{c|c|c|c} \mathbf{U_1} & \mathbf{U_2} & \mathbf{U_3} & \mathbf{U_4} \\ w_6, w_2 & & w_4, w_3, w_1 & w_5. \end{array}$$

From this arrangement, we form the cycles $(1, w_6, w_2) = (1,18,8,22,13,9)$, (2), $(3, w_4, w_3, w_1) = (3,12,14,24,15,10,5,11,7)$ and $(4, w_5) = (4,16,20,17,23)$. Considering these cycles together with the two that were not selected, one obtains $\pi \in \Pi_{4,1}(20,2)$ given by

$$\pi = (1,18,8,22,13,9)(2)(3,12,14,24,15,10,5,11,7)(4,16,20,17,23)(6)(19,21).$$

We now prove formula (2.8), rewritten in the form

$$m^{n-k}c_q(n,k) = \sum_{j=k}^{n}(-1)^{j-k}\binom{j}{k}[r]_p^{j-k}v_{p,q}(n,j), \quad n \ge k \ge 0. \tag{3.4}$$

To show (3.4), let $\Omega'_{r,m}(n,k)$ denote the set obtained from members of $\Omega_{r,m}(n,k)$ by marking some subset of the l-r cycle min belonging to special cycles. Define the sign $\lambda \in \Omega'_{r,m}(n,k)$ to be $(-1)^{j-k}$, where $j-k$ denotes the number of marked l-r cycle min of λ, and define the weight of λ as we did before for members of $\Omega_{r,m}(n,k)$. We first show that the right-hand side of (3.4) gives the total (signed) weight of all the members of $\Omega'_{r,m}(n,k)$. To do so, it is enough to show that the weight of the members of $\Omega'_{r,m}(n,k)$ in which there are $j-k$ marked cycle min is $\binom{j}{k}[r]_p^{j-k}v_{p,q}(n,j)$ for $k \le j \le n$.

To form such members of $\Omega'_{r,m}(n,k)$, we first choose $\tau \in \Omega_{r,m}(n,j)$ for some $j \geq k$ and select exactly $j-k$ of the j non-special cycles of τ. We insert the contents of these $j-k$ cycles into the special cycles of τ as follows. Let $b = b_1 b_2 \cdots b_s$ denote the contents of a selected cycle in the order that the letters appear within the cycle. We will insert b into one of the r special cycles of τ so that b_1 will become a l-r cycle min. Let $C = j w_1 w_2 \cdots w_\ell$ denote the contents of the cycle in which we are to insert b, where $j \in [r]$ and w_i denotes all of the letters from the i-th largest cycle min of C up to but not including the $(i+1)$-st largest cycle min. That is, we have $\min(w_1) > \min(w_2) > \cdots > \min(w_\ell)$, with $\min(w_i)$ also the first letter of the subword w_i for each i.

If $b_1 > \min(w_1)$ or if C contains only j, then we write the letters in b directly after the letter j in C. Otherwise, let i_0 be the index $i \in [\ell]$ such that $\min(w_i) > b_1 > \min(w_{i+1})$, where $\min(w_{\ell+1}) = 0$. We then write the letters of b between the subwords w_{i_0} and w_{i_0+1} in C if $i_0 < \ell$ or after w_ℓ if $i_0 = \ell$. Next, we mark the letter b_1; note that b_1 is a cycle min, as are still the first letters of each of the w_i. Repeat the above procedure for each of the $j-k$ selected cycles, where cycles are inserted one after another, sequentially, and we consider also the subwords arising from previously inserted cycles when deciding on the position of the current cycle. Since the first letter of each selected cycle becomes a l-r cycle min, there are $[r]_p^{j-k}$ possibilities concerning the insertion of these cycles. Furthermore, since the predecessors of the non-minimal elements within the selected non-special cycles of τ remain the same once their contents have been added to the special cycles as described, the contribution of these non-minimal elements towards the q-weight (and also the m-weight) remains the same.

We illustrate the procedure described above for creating members of $\Omega'_{r,m}(n,k)$, where $n = 21$, $k = 2$, $r = 3$ and $m = 1$. Let $j = 6$ and $\tau \in \Omega_{3,1}(21,6)$ given by

$$\tau = (1,7,5,19)(2)(3,18,12,4)(6,9)(8,13,10)(11)(14,22,16)(15,24,21,17)(20,23).$$

Suppose now that we select the four non-special cycles $(6,9)$, (11), $(15,24,21,17)$ and $(20,23)$, and stipulate that $(6,9)$ and $(20,23)$ go in the first and second special cycle of τ, respectively, while the other two go in the third. This yields the permutation $\lambda \in \Omega'_{3,1}(21,2)$ given by

$$\lambda = (1,7,\underline{6},9,5,19)(2,\underline{20},23)(3,18,\underline{15},24,21,17,12,\underline{11},4)(8,13,10)(14,22,16),$$

where the marked cycle min are underlined. Upon considering the marked letters of λ, the above process is seen to be reversible. Allowing τ to vary thus yields all members of $\Omega'_{r,m}(n,k)$ having exactly $j-k$ marked cycle min, which are seen to have weight $\binom{j}{k}[r]_p^{j-k} v_{p,q}(n,j)$, as desired.

Now consider the smallest l-r cycle min belonging to a special cycle within a member of $\Omega'_{r,m}(n,k)$. Either mark it if it is unmarked, or remove the marking from it. For example, this would entail underlining the element 4 in the permutation λ above. This operation is a sign-changing, weight-preserving involution of $\Omega'_{r,m}(n,k)$, which is not defined whenever all of the special cycles are singletons. The sign of each such member of $\Omega'_{r,m}(n,k)$ is positive, and the weight of all such members is seen to be $m^{n-k}c_q(n,k)$, which implies (3.4) and completes the proof. $\square$

We now provide a combinatorial proof of the orthogonality relations between $w_{p,q}(n,k)$ and $W_{p,q}(n,k)$. Before doing so, we first recall a combinatorial interpretation for the array $W_{p,q}(n,k)$ from [13]. Given $0 \leq r \leq m$, by an *r-partition* of $[m]$, we will mean a partition of the set $[m]$ in which the elements of $[r]$ belong to distinct blocks. If $n,k,r \geq 0$, then let $\Pi_r(n,k)$ denote the set of all r-partitions of $[n+r]$ having $k+r$ blocks and let $\Pi_r(n) = \cup_{k=0}^{n}\Pi_r(n,k)$. Note that when $r=0$, an r-partition of $[m]$ is the same as an ordinary partition. We will apply the terms *special* and *minimal* with regard to the members of $\Pi_r(n,k)$ in a manner completely analogous to how those terms were applied above towards members of $\Omega_r(n,k)$ (with "cycles" replaced by "blocks" at the appropriate points in the definitions).

Elements of r-partitions will be assigned colors in the following manner.

Definition 3.9 (Mansour et al. [13])**.** Given an integer $m \geq 1$, let $\Pi_{r,m}(n,k)$ denote the set of r-partitions of $[n+r]$ having $k+r$ blocks wherein within each non-special block, every non-minimal element is assigned one of m colors, and let $\Pi_{r,m}(n) = \cup_{k=0}^{n}\Pi_{r,m}(n,k)$.

Upon making a comparison of the recurrences and initial values, we see that $|\Pi_{r,m}(n,k)| = W(n,k;r,m)$ for all r and m. We now recall a couple of statistics on $\Pi_{r,m}(n,k)$.

Definition 3.10 (Mansour et al. [13])**.** Suppose $\pi \in \Pi_{r,m}(n,k)$ is represented as

$$\pi = A_1/A_2/\cdots/A_r/B_1/B_2/\cdots/B_k,$$

where A_i denotes the special block containing the element i for $i \in [r]$ and non-special blocks are denoted by B_j, with $\min(B_1) < \min(B_2) < \cdots < \min(B_k)$. Define the statistics w_1 and w_2 on $\Pi_{r,m}(n,k)$ by letting

$$w_1(\pi) = \sum_{i=1}^{r}(i-1)(|A_i|-1)$$

and

$$w_2(\pi) = \sum_{i=1}^{k}(i-1)(|B_i|-1).$$

In [13], it was shown that

$$W_{p,q}(n,k) = \sum_{\pi \in \Pi_{r,m}(n,k)} p^{w_1(\pi)}q^{w_2(\pi)}, \quad n,k \geq 0.$$

Note that $W_{p,q}(n,k)$ reduces to $W(n,k)$ when $p=q=1$. Using (1.2) and (2.1), one can obtain orthogonality relations between the arrays $w_{p,q}(n,k)$ and $W_{p,q}(n,k)$. Here, we give bijective proofs by making use of our combinatorial interpretations for these arrays.

Theorem 3.11. *If $n \geq k \geq 0$, then*

$$\sum_{j=k}^{n} W_{p,q}(n,j)w_{p,q}(j,k) = \sum_{j=k}^{n} w_{p,q}(n,j)W_{p,q}(j,k) = \delta_{n,k}. \tag{3.5}$$

Proof. To show the first relation in (3.5), we consider sets $\mathcal{A}_j$ where $k \leq j \leq n$ of ordered pairs (α, β) in which $\alpha \in \Pi_{r,m}(n,j)$ and β is an arrangement of the blocks of α according to some member of $\Omega_{r,m}(j,k)$. Within β, blocks of α are ordered according to the size of their smallest elements, with the special blocks of α (i.e., those containing a member of $[r]$) regarded as special elements of β. Thus, the special cycles of β are those that contain a special block of α. Define the sign of $(\alpha, \beta) \in \mathcal{A}_j$ by $(-1)^{j-k}$ and the weight by $p^{w_1(\alpha)+v_1(\beta)}q^{w_2(\alpha)+v_2(\beta)}$. Let $\mathcal{A} = \cup_{j=k}^{n}\mathcal{A}_j$. For example, if $n = 10$, $k = 1$, $r = 2$, $m = 1$ and $j = 4$, then $(\alpha, \beta) \in \mathcal{A}_4$, where

$$\alpha = \{1,3,5\}, \{2,4,8\}, \{6\}, \{7,11\}, \{9\}, \{10,12\}$$

and

$$\beta = (\{1,3,5\})(\{2,4,8\}, \{9\}, \{6\})(\{7,11\}, \{10,12\}),$$

has sign $(-1)^{4-1} = -1$ and weight $p^{2+2}q^{4+1} = p^4q^5$. The first sum in (3.5) then gives the total (signed) weight of all the members of $\mathcal{A}$. To complete the proof, we define a sign-changing, weight-preserving involution of $\mathcal{A}$.

In order to do so, given $(\alpha, \beta) \in \mathcal{A}$, let x be the largest $i \in I$ such that it is *not* the case that a cycle of the form $(\{i\})$ containing only the block $\{i\}$ occurs in β. Let B be the block of α containing the element x. Note that B cannot have smallest element x, lest B be a singleton. If $|B| \geq 2$, then break off x and form the singleton $\{x\}$ to directly follow $B - \{x\}$ within its cycle of β. Observe that if $\{x\}$ occurs as a block of α, then it cannot be first within its cycle of β, by the ordering of blocks of α within β and the assumption on x (note that all $i > x$ must occur within β as 1-cycles of the form $(\{i\})$). Thus, if $\{x\}$ occurs, one may move it to the block within its cycle of β that directly precedes it. Combining the two previous mappings defines an involution of $\mathcal{A}$ if $n > k$ since at least one cycle of β in this case always contains at least two elements of $[n + r]$ altogether, with at least one member of I belonging to a block within such a cycle. If $n = k$, then $\mathcal{A}$ contains only a single element having weight 1.

Clearly, the involution defined in the previous paragraph changes the sign since the number of (non-special) blocks of α changes by one. We now show that it always preserves the weight. First suppose that B belongs to a non-special cycle of β. Then moving x as described in the first mapping preserves the sum $w_2(\alpha)+v_2(\beta)$ since if x belonged to the i-th non-special block of α to start with, then breaking off $\{x\}$ reduces $w_2(\alpha)$ by $i - 1$ but increases $v_2(\beta)$ by the same amount since $\{x\}$ has predecessor $B - \{x\}$, which is now the i-th smallest non-special element of β. Note that $\{x\}$ then becomes the largest element within its cycle of β, and hence $\{x\}$ following $B - \{x\}$ does not affect a possible contribution to $v_2(\beta)$ from a block succeeding B in this cycle. Moreover, since all $i > x$ occur as singletons in

α, reordering the blocks of α after forming $\{x\}$ does not further affect the $w_2(\alpha)$ value. Note that the value of $w_1(\alpha)+v_1(\beta)$ is unaffected since neither statistic is. Finally, the color that the element x would have been assigned being a non-minimal element of a non-special block of α is transferred to the color assigned the block $\{x\}$ for having a predecessor. Thus, the weight of (α,β) is preserved by the involution in this case.

Now suppose that the block B belongs to a special cycle of β (i.e., one that has a block of α containing a member of $[r]$). If B is a non-special block of α that does not correspond to a l-r cycle min of β, then one may use the reasoning of the prior paragraph to show that the weight is preserved. The same also applies if B is indeed a l-r cycle min of β. Finally, suppose B is a special block of α. Then breaking off $\{x\}$ reduces the $w_1(\alpha)$ value by $\ell-1$ for some $\ell\in[r]$, while it increases $v_1(\beta)$ by the same amount since $\{x\}$ in this case becomes a l-r cycle min within the ℓ-th special cycle of β. Thus, the sum $w_1(\alpha)+v_1(\beta)$ is preserved. There is also no change in $w_2(\alpha)+v_2(\beta)$ since neither statistic is affected in this case, with neither the element x in α nor the block $\{x\}$ in β being assigned a color. Thus, the weight of (α,β) is once again preserved, which implies the first relation in (3.5).

A similar proof applies to the second relation in (3.5). We describe the main differences. Here, one would consider ordered pairs (γ,δ) in which $\gamma\in\Omega_{r,m}(n,j)$ and δ is an arrangement of the cycles of γ according to some member of $\Pi_{r,m}(j,k)$. The sign of (γ,δ) would be $(-1)^{n-j}$ and the weight $p^{v_1(\gamma)+w_1(\delta)}q^{v_2(\gamma)+w_2(\delta)}$. Note that a special block of δ is one that has an element of $[r]$ contained within one of its cycles.

To define the involution in this case, suppose that the blocks of δ are arranged from left-to-right in ascending order of smallest elements contained therein (the special blocks then being first). Consider the leftmost block of δ that contains at least two elements of $[n+r]$ altogether within its cycles. Let C denote this block and u and v be the smallest and second smallest elements of $[n+r]$ in C, respectively. If u and v belong to the same cycle of γ within C, then we split this cycle at v and form a new cycle starting with v, which we write directly following the cycle containing u in C. If u and v belong to different cycles of γ, whence v starts a cycle of γ, then we merge them into one large cycle. Upon considering whether or not C is a special block of δ, one may verify that this mapping is always a sign-changing, weight-preserving involution, which completes the proof. □

4. The (p,q)-Whitney matrix of the first kind

In [13], the (p,q)-Whitney matrix of the second kind was introduced and several properties of this matrix are proven. In this section, we introduce the (p,q)-Whitney matrix of the first kind and find some factorizations of it in analogy with the results of Mező and Ramírez [16].

Definition 4.1. The (p,q)-Whitney matrix of the first kind is the $n\times n$ matrix defined by

$$\mathcal{L}_{p,q}(n) := \mathcal{L}_{p,q}^{(m,r)}(n) = \left[w_{p,q}(i,j;r,m)\right]_{0\le i,j\le n-1}.$$

For example, $\mathcal{L}_{p,q}(4)$ is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -[r]_p & 1 & 0 & 0 \\ [r]_p^2 + m[r]_p & -m-2[r]_p & 1 & 0 \\ -[r]_p^3 - (2+q)m[r]_p^2 - (1+q)m^2[r]_p & (1+q)m^2 + 2(2+q)m[r]_p + 3[r]_p^2 & -(2+q)m - 3[r]_p & 1 \end{bmatrix}.$$

In particular, if $p = q = 1$, we obtain the r-Whitney matrix of the first kind [16]. If $m = p = 1$ and $r = 0$, we obtain the q-Stirling matrix of the first kind $S_{q,n}^{(1)} := [s_q(i,j)]_{0\le i,j\le n-1}$; see, e.g., [11, 19].

Recall the generalized $n\times n$ Pascal matrix $P_n[x]$ (see [4]) defined as

$$P_n[x] := \left[\binom{i}{j} x^{i-j}\right]_{0\le i,j\le n-1}.$$

If $x = 1$, we obtain the Pascal matrix P_n of order n. Moreover,

$$P_n^{-1}[x] = P_n[-x] = \left[(-1)^{i-j}\binom{i}{j} x^{i-j}\right]_{0\le i,j\le n-1}.$$

From identity (2.7), we have the following factorization:

$$\mathcal{L}_{p,q}(n) = S_{q,n}^{(1)}[m] P_n[-[r]_p], \quad n \ge 1, \tag{4.1}$$

where $S_{q,n}^{(1)}[x] := [s_q(i,j)x^{i-j}]_{0\le i,j\le n-1}$.

For example,

$$\begin{aligned} \mathcal{L}_{p,q}(4) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -m & 1 & 0 \\ 0 & m^2(1+q) & -(2+q)m & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 \\ -[r]_p & 1 & 0 & 0 \\ [r]_p^2 & -2[r]_p & 1 & 0 \\ -[r]_p^3 & 3[r]_p^2 & -3[r]_p & 1 \end{bmatrix} \\ &= S_{q,4}^{(1)}[m]\, P_4[-[r]_p]. \end{aligned}$$

Moreover, from identity (2.5), we obtain

$$\mathcal{L}_{p,q}^{(m,r)}(n) = \mathcal{L}_{p,q}^{(m,s)}(n) P_n[-p^s[r-s]_p], \quad 0 \le s \le r.$$

Given $n \ge 1$, let $S_n[x]$ be the $n\times n$ matrix defined by $S_n[x] := \left[x^{i-j}\right]_{0\le j\le i\le n-1}$. The following factorization of the generalized Pascal matrix is due to Zhang [22, Theorem 1]:

$$P_n[x] = G_n[x]G_{n-1}[x]\cdots G_1[x], \quad n \ge 1, \tag{4.2}$$

where $G_n[x] = S_n[x]$ and $G_k[x] = I_{n-k} \oplus S_k[x]$ for $1 \le k \le n-1$ with $\oplus$ denoting the matrix direct sum.

From the preceding, we obtain the following factorization of the (p,q)-Whitney matrix of the first kind.

Proposition 4.2. *If $n \geq 2$, then*

$$\begin{aligned}&\mathcal{L}_{p,q}(n)\\&= \overline{P}_1[-mq^{n-2}]\cdots\overline{P}_{n-2}[-mq]\overline{P}_{n-1}[-m]G_n[-[r]_p]G_{n-1}[-[r]_p]\cdots G_1[-[r]_p],\end{aligned} \tag{4.3}$$

where

$$\overline{P}_k[x] = I_{n-k} \oplus P_k[x].$$

Proof. By (4.1), we have

$$\mathcal{L}_{p,q}(n) = S^{(1)}_{q,n}[m]P_n[-[r]_p].$$

The matrix $P_n[-[r]_p]$ can be factorized by means of (4.2), while the matrix $S^{(1)}_{q,n}[m]$ can be factorized by a result of Ernst [11] as

$$S^{(1)}_{q,n}[m] = \overline{P}_1[-mq^{n-2}]\cdots\overline{P}_{n-2}[-mq]\overline{P}_{n-1}[-m],$$

which implies (4.3). □

5. Conclusion

In this paper, we have introduced a (p,q)-generalization $w_{p,q}(n,k)$ of the r-Whitney numbers of the first kind that reduces to these numbers when $p = q = 1$. In addition to forming an orthogonal pair with a prior generalization of the r-Whitney numbers of the second kind, the $w_{p,q}(n,k)$ arise as the joint distribution for two statistics on a set of generalized permutations. When $r = 0$ and $m = 1$, these statistics are seen to be new on the usual set of permutations and the statistic marked by the q-variable has the same distribution on $\mathcal{S}_n$ as an earlier statistic considered by Carlitz. Since our $w_{p,q}(n,k)$ when $p = 1$ are closely related to the $w_{m,r,q}(n,k)$ studied in [12], which arose in a physical context, modifying slightly our combinatorial interpretation for $w_{p,q}(n,k)$ furnishes a comparable interpretation for the $w_{m,r,q}(n,k)$. Thus, one may obtain, via combinatorial arguments, generalizations of identities from [12].

Furthermore, using Theorem 2.2 and a generalized version of the central limit theorem, it is possible to show that the v_1 and v_2 statistics follow an asymptotically normal distribution as n increases without bound for all $r \geq 2$ and $m \geq 1$. In addition, from Theorem 2.2, it is seen that the array $w_{p,q}(n,k)$ when p and q are real is log-concave by Newton's criterion since the polynomial in identity (2.2) is real-rooted in that case. On the other hand, we seek a general asymptotic formula for $w_{p,q}(n,k)$ when p and q are positive. Non-trivial combinatorial (or algebraic) generalizations of the sequence satisfying recurrence (3.1) are also sought, as such generalizations would likely yield new statistics on the set of permutations. Finally, it would be interesting to find connections between the array $w_{p,q}(n,k)$ and other combinatorial structures.

References

[1] H. Belbachir and I. E. Bousbaa, Translated Whitney and r-Whitney numbers: a combinatorial approach, *J. Integer Seq.* **16** (2013), Art. 13.8.6.

[2] M. Benoumhani, On Whitney numbers of Dowling lattices, *Discrete Math.* **159** (1996), 13–33.

[3] A. Z. Broder, The r-Stirling numbers, *Discrete Math.* **49** (1984), 241–259.

[4] G. S. Call and D. J. Velleman, Pascal's matrices, *Amer. Math. Monthly* **100** (1993), 372–376.

[5] L. Carlitz, q-Bernoulli numbers and polynomials, *Duke Math. J.* **15** (1948), 987–1000.

[6] L. Carlitz, Generalized Stirling numbers, *Combinatorial Analysis Notes*, Duke University (1968), 1–15.

[7] G.-S. Cheon and J.-H. Jung, r-Whitney numbers of Dowling lattices, *Discrete Math.* **312** (2012), 2337–2348.

[8] R. B. Corcino and C. Barrientos, Some theorems on the q-analogue of the generalized Stirling numbers, *Bull. Malays. Math. Sci. Soc.*, Series 2, **34(3)** (2011), 487–501.

[9] R. B. Corcino and C. B. Montero, A q-analogue of Rucinski-Voigt numbers, *ISRN Discrete Math.* (2012), Art. ID 592818.

[10] T. A. Dowling, A class of geometric lattices based on finite groups, *J. Combin. Theory Ser. B* **14** (1973), 61–86. (erratum: *J. Combin. Theory Ser. B* **15** (1973), 211.)

[11] T. Ernst, q-Leibniz functional matrices with applications to q-Pascal and q-Stirling matrices, *Adv. Stud. Contemp. Math.* **22(4)** (2012), 537–555.

[12] M. M. Mangontarum and J. Katriel, On q-boson operators and q-analogues of the r-Whitney and r-Dowling numbers, *J. Integer Seq.* **18** (2015), Art. 15.9.8.

[13] T. Mansour, J. L. Ramírez, and M. Shattuck, A generalization of the r-Whitney numbers of the second kind, *J. Comb.* **8(1)** (2017), 29–55.

[14] M. Merca, A note on the r-Whitney numbers of Dowling lattices, *C. R. Acad. Sci. Paris, Ser. I* **351** (2013), 649–655.

[15] I. Mező, A new formula for the Bernoulli polynomials, *Results Math.* **58** (2010), 329–335.

[16] I. Mező and J. L. Ramírez, The linear algebra of the r-Whitney matrices, *Integral Transforms Spec. Funct.* **26(3)** (2015), 213–225.

[17] M. Mihoubi and M. Rahmani, The partial r-Bell polynomials, arXiv:1308.0863, (2013).

[18] M. Mihoubi and M. Tiachachat, Some applications of the r-Whitney numbers, *C. R. Acad. Sci. Paris, Ser. I* **352** (2014), 965–969.

[19] H. Oruç and H. K. Akmaz, Symmetric functions and the Vandermonde matrix, *J. Comput. Appl. Math.* **172** (2004), 49–64.

[20] M. Shattuck, Parity theorems for statistics on permutations and Catalan words, *Integers* **5** (2005), #A07.

[21] C. Wagner, Generalized Stirling and Lah numbers, *Discrete Math.* **160** (1996), 199–218.

[22] Z. Zhang, The linear algebra of the generalized Pascal matrix, *Linear Algebra Appl.* **250** (1997), 51–60.

Annales Mathematicae et Informaticae
46 *(2016) pp. 195–204*
http://ami.ektf.hu

On the (s,t)-Pell and (s,t)-Pell-Lucas numbers by matrix methods*

Somnuk Srisawat, Wanna Sriprad

Department of Mathematics and computer science,
Faculty of Science and Technology,
Rajamangala University of Technology Thanyaburi,
Pathum Thani 12110,Thailand.
somnuk_s@rmutt.ac.th
wanna_sriprad@rmutt.ac.th

Submitted March 17, 2016 — Accepted September 5, 2016

Abstract

In this paper, we investigate some generalization of Pell and Pell-Lucas numbers, which is called (s,t)-Pell and (s,t)-Pell-Lucas numbers, and we define the 2×2 matrix W, which satisfy the relation $W^2 = 2sW + tI$. After that, we establish some identities of (s,t)-Pell and (s,t)-Pell-Lucas numbers and some sum formulas for (s,t)-Pell and (s,t)-Pell-Lucas numbers by using this matrix.

Keywords: Fibonacci number; Lucas number; Pell number; Pell-Lucas number; (s,t)-Pell number; (s,t)-Pell-Lucas number.

MSC: 11B37; 15A15.

1. Introduction

For over several years, there are many recursive sequences that have been studied in the literatures. The famous examples of these sequences are Fibonacci, Lucas, Pell and Pell-Lucas, because they are extensively used in various research areas such as Engineering, Architecture, Nature and Art (for examples see: [2, 3, 4, 5, 6, 7]). For $n \geq 2$, the classical Fibonacci $\{F_n\}$, Lucas $\{L_n\}$, Pell $\{P_n\}$ and Pell-Lucas $\{Q_n\}$

*This research was supported by faculty of science and technology, Rajamangala University of Technology Thanyaburi (RMUTT), Thailand.

sequences are defined by $F_n = F_{n-1}+F_{n-2}$, $L_n = L_{n-1}+L_{n-2}$, $P_n = 2P_{n-1}+P_{n-2}$, and $Q_n = 2Q_{n-1}+Q_{n-2}$, with the initial conditions $F_0 = 0$, $F_1 = 1$, $L_0 = 2$, $L_1 = 1$, $P_0 = 0$, $P_1 = 1$ and $Q_0 = Q_1 = 2$, respectively. For more detialed information about Fibonacci, Lucas, Pell, Pell-Lucas sequences can be found in [2, 3].

Recently, Fibonacci, Lucas, Pell and Pell-Lucas were generalized and studied by many authors in the different ways to derive many identities. In 2012, Gulec and Taskara [1] introduced a new generalization of Pell and Pell-Lucas sequence which is called (s,t)-Pell and (s,t)-Pell-Lucas sequences as in the definition 1.1 and by considering these sequences, they introduced the matrix sequences which have elements of (s,t)-Pell and (s,t)-Pell-Lucas sequences. Further, they obtained some properties of (s,t)-Pell and (s,t)-Pell-Lucas matrices sequences by using elementary matrix algebra.

Definition 1.1. [1] Let s,t be any real number with $s^2+t>0$, $s>0$ and $t\neq 0$. Then the (s,t)-Pell sequences $\{\mathcal{P}_n(s,t)\}_{n\in\mathbb{N}}$ and the (s,t)-Pell-Lucas sequences $\{\mathcal{Q}_n(s,t)\}_{n\in\mathbb{N}}$ are defined respectively by

$$\mathcal{P}_n(s,t) = 2s\mathcal{P}_{n-1}(s,t) + t\mathcal{P}_{n-2}(s,t), \text{ for } n\geq 2, \tag{1.1}$$

$$\mathcal{Q}_n(s,t) = 2s\mathcal{Q}_{n-1}(s,t) + t\mathcal{Q}_{n-2}(s,t), \text{ for } n\geq 2, \tag{1.2}$$

with initial conditions $\mathcal{P}_0(s,t)=0$, $\mathcal{P}_1(s,t)=1$ and $\mathcal{Q}_0(s,t)=2$, $\mathcal{Q}_1(s,t)=2s$.

In particular, if $s=\frac{1}{2}, t=1$, then the classical Fibonacci and Lucas sequence are obtained, and if $s=t=1$, then the classical Pell and Pell-Lucas sequences are obtained. From the definition 1.1, we have that the characteristic equation of (1.1) and (1.2) are in the form

$$x^2 = 2sx+t \tag{1.3}$$

and the root of equation (1.3) are $\alpha = s+\sqrt{s^2+t}$ and $\beta = s-\sqrt{s^2+t}$. Note that $\alpha+\beta = 2s$, $\alpha-\beta = 2\sqrt{s^2+t}$ and $\alpha\beta=-t$. Moreover, it can be seen that [1]

$$\mathcal{Q}_n(s,t) = 2s\mathcal{P}_n(s,t) + 2t\mathcal{P}_{n-1}(s,t), \text{ for all } n\geq 0. \tag{1.4}$$

In this paper, we introduce the 2×2 matrix W which satisfy the relation $W^2 = 2sW+tI$. After that, we establish some identities of (s,t)-Pell and (s,t)-Pell-Lucas numbers and some sum formulas for (s,t)-Pell and (s,t)-Pell-Lucas numbers by using this matrix. Now, we first define (s,t)-Pell and (s,t)-Pell-Lucas numbers for negative subscript as follows:

$$\mathcal{P}_{-n}(s,t) = \frac{-\mathcal{P}_n(s,t)}{(-t)^n}, \text{ and } \mathcal{Q}_{-n}(s,t) = \frac{\mathcal{Q}_n(s,t)}{(-t)^n}, \tag{1.5}$$

for all $n\geq 1$. In the rest of this paper, for convenience we will use the symbol $\mathcal{P}_n$ and $\mathcal{Q}_n$ instead of $\mathcal{P}_n(s,t)$ and $\mathcal{Q}_n(s,t)$ respectively.

2. Main results

We begin this section with the following Lemma.

Lemma 2.1. *If X is a square matrix with $X^2 = 2sX + tI$, then*

$$X^n = \mathcal{P}_n X + t\mathcal{P}_{n-1} I \quad \text{for all} \quad n \in \mathbb{N}.$$

Proof. If $n = 0$, then the proof is obvious. It can be shown by induction that $X^n = \mathcal{P}_n X + t\mathcal{P}_{n-1} I$ for all $n \in \mathbb{N}$. Now, we will show that $X^{-n} = \mathcal{P}_{-n} X + t\mathcal{P}_{-n-1} I$ for all $n \in \mathbb{N}$. Let $Y = 2sI - X = -tX^{-1}$. Then we have

$$Y^2 = (2sI - X)^2 = 2s(2sI - X) + tI = 2sY + tI.$$

It implies that $Y^n = \mathcal{P}_n Y + t\mathcal{P}_{n-1} I$. That is $(-tX^{-1})^n = \mathcal{P}_n(2sI - X) + t\mathcal{P}_{n-1} I$. Thus

$$\begin{aligned}(-t)^n X^{-n} &= 2s\mathcal{P}_n I - \mathcal{P}_n X + t\mathcal{P}_{n-1} I \\ &= -\mathcal{P}_n X + (2s\mathcal{P}_n + t\mathcal{P}_{n-1}) I \\ &= -\mathcal{P}_n X + \mathcal{P}_{n+1} I.\end{aligned}$$

Therefore, $X^{-n} = -\frac{\mathcal{P}_n}{(-t)^n} X + \frac{\mathcal{P}_{n+1}}{(-t)^n} I = \mathcal{P}_{-n} X + t\mathcal{P}_{-(n+1)} I = \mathcal{P}_{-n} X + t\mathcal{P}_{-n-1} I$. This complete the proof. □

By using Lemma 2.1, we obtain the Binet's formula for (s,t)-Pell and (s,t)-Pell-Lucas numbers.

Corollary 2.2 (Binet's formula)**.** *The n^{th} (s,t)-Pell and (s,t)-Pell-Lucas number are given by*

$$\mathcal{P}_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \quad \text{and} \quad \mathcal{Q}_n = \alpha^n + \beta^n, \quad \text{for all} \quad n \in \mathbb{Z},$$

where $\alpha = s + \sqrt{s^2 + t}$ and $\beta = s - \sqrt{s^2 + t}$ are the roots of the characteristic equation (1.3)*.*

Proof. Take $X = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}$, then $X^2 = 2sX + tI$. By Lemma 2.1, we have $X^n = \mathcal{P}_n X + t\mathcal{P}_{n-1} I$. It follows that

$$\begin{bmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{bmatrix} = \begin{bmatrix} \alpha\mathcal{P}_n + t\mathcal{P}_{n-1} & 0 \\ 0 & \beta\mathcal{P}_n + t\mathcal{P}_{n-1} \end{bmatrix}.$$

Thus, $\alpha^n = \alpha\mathcal{P}_n + t\mathcal{P}_{n-1}$ and $\beta^n = \beta\mathcal{P}_n + t\mathcal{P}_{n-1}$, which implies that

$$\mathcal{P}_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \quad \text{and} \quad \mathcal{Q}_n = \alpha^n + \beta^n, \quad \text{for all} \quad n \in \mathbb{Z}.$$ □

Let us define the 2×2 matrix W as follows:

$$W = \begin{bmatrix} s & 2(s^2 + t) \\ \frac{1}{2} & s \end{bmatrix}. \tag{2.1}$$

Then it easy to see that $W^2 = 2sW + tI$. From this fact and Lemma 2.1, we get the following Lemma.

Lemma 2.3. *Let W be a matrix as in (2.1). Then* $W^n = \begin{bmatrix} \frac{1}{2}\mathcal{Q}_n & 2(s^2+t)\mathcal{P}_n \\ \frac{1}{2}\mathcal{P}_n & \frac{1}{2}\mathcal{Q}_n \end{bmatrix}$ *for all $n \in \mathbb{Z}$.*

Proof. Since $W^2 = 2sW + tI$, the proof follows from Lemma 2.1 and using $\mathcal{Q}_n = 2s\mathcal{P}_n + 2t\mathcal{P}_{n-1}$. □

Now, by using the matrix W, we obtain some identities of (s,t)-Pell and (s,t)-Pell-Lucas numbers.

Lemma 2.4. *Let m, n be any integers. Then the following results hold.*

(i) $\mathcal{Q}_n^2 - 4(s^2+t)\mathcal{P}_n^2 = 4(-t)^n$,

(ii) $2\mathcal{Q}_{m+n} = \mathcal{Q}_m\mathcal{Q}_n + 4(s^2+t)\mathcal{P}_m\mathcal{P}_n$,

(iii) $2\mathcal{P}_{m+n} = \mathcal{P}_m\mathcal{Q}_n + \mathcal{Q}_m\mathcal{P}_n$,

(iv) $2(-t)^n\mathcal{Q}_{m-n} = \mathcal{Q}_m\mathcal{Q}_n - 4(s^2+t)\mathcal{P}_m\mathcal{P}_n$,

(v) $2(-t)^n\mathcal{P}_{m-n} = \mathcal{P}_m\mathcal{Q}_n - \mathcal{Q}_m\mathcal{P}_n$,

(vi) $\mathcal{Q}_m\mathcal{Q}_n = \mathcal{Q}_{m+n} + (-t)^n\mathcal{Q}_{m-n}$,

(vii) $\mathcal{P}_m\mathcal{Q}_n = \mathcal{P}_{m+n} + (-t)^n\mathcal{P}_{m-n}$.

Proof. Since $\det(W^n) = (\det(W))^n = (-t)^n$ and $\det(W^n) = \frac{1}{4}\mathcal{Q}_n^2 - (s^2+t)\mathcal{P}_n^2$, we get that $\mathcal{Q}_n^2 - 4(s^2+t)\mathcal{P}_n^2 = 4(-t)^n$ and then (i) immediately seen. Since $W^{m+n} = W^mW^n$, we obtain

$$\begin{aligned}&\begin{bmatrix} \frac{1}{2}\mathcal{Q}_{m+n} & 2(s^2+t)\mathcal{P}_{m+n} \\ \frac{1}{2}\mathcal{P}_{m+n} & \frac{1}{2}\mathcal{Q}_{m+n} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{4}\big(\mathcal{Q}_m\mathcal{Q}_n + 4(s^2+t)\mathcal{P}_m\mathcal{P}_n\big) & (s^2+t)(\mathcal{Q}_m\mathcal{P}_n + \mathcal{P}_m\mathcal{Q}_n) \\ \frac{1}{4}(\mathcal{P}_m\mathcal{Q}_n + \mathcal{Q}_m\mathcal{P}_n) & \frac{1}{4}\big(4(s^2+t)\mathcal{P}_m\mathcal{P}_n + \mathcal{Q}_m\mathcal{Q}_n\big) \end{bmatrix}.\end{aligned}$$

Thus, identities (ii) and (iii) are easily seen. Next, we note that $W^{m-n} = W^m(W^{-n}) = W^m(W^n)^{-1}$. Thus, we get that

$$\begin{aligned}&\begin{bmatrix} \frac{1}{2}\mathcal{Q}_{m-n} & 2(s^2+t)\mathcal{P}_{m-n} \\ \frac{1}{2}\mathcal{P}_{m-n} & \frac{1}{2}\mathcal{Q}_{m-n} \end{bmatrix} \\ &= \frac{1}{(-t)^n}\begin{bmatrix} \frac{1}{4}\big(\mathcal{Q}_m\mathcal{Q}_n - 4(s^2+t)\mathcal{P}_m\mathcal{P}_n\big) & (s^2+t)(-\mathcal{Q}_m\mathcal{P}_n + \mathcal{P}_m\mathcal{Q}_n) \\ \frac{1}{4}(\mathcal{P}_m\mathcal{Q}_n - \mathcal{Q}_m\mathcal{P}_n) & \frac{1}{4}\big(-4(s^2+t)\mathcal{P}_m\mathcal{P}_n + \mathcal{Q}_m\mathcal{Q}_n\big) \end{bmatrix}.\end{aligned}$$

Therefore, the identities (iv) and (v) can be derived directly. The proof of (vi) and (vii) goes on in the same fashion as above by using the property
$W^{m+n} + (-t)^nW^{m-n} = W^m(W^n + (-t)^nW^{-n})$. □

Next, we give the following Lemma for using in the next Theorems.

Lemma 2.5. *Let W be a matrix as in (2.1). Then*

$$H = W + tW^{-1} = \begin{bmatrix} 0 & 4(s^2+t) \\ 1 & 0 \end{bmatrix}.$$

Proof. Since $\det(W) = -t$, we get that $W^{-1} = -\frac{1}{t}\begin{bmatrix} s & -2(s^2+t) \\ -\frac{1}{2} & s \end{bmatrix}$. Thus, $H = \begin{bmatrix} 0 & 4(s^2+t) \\ 1 & 0 \end{bmatrix}$. □

Finally, by using matrices W and H, we obtain some sum formulas for (s,t)-Pell and (s,t)-Pell-Lucas numbers.

Theorem 2.6. *Let $n \in \mathbb{N}$ and $m, k \in \mathbb{Z}$ with $(-t)^m - \mathcal{Q}_m \neq -1$. Then*

$$\sum_{j=0}^{n} \mathcal{Q}_{mj+k} = \frac{\mathcal{Q}_k - \mathcal{Q}_{mn+m+k} + (-t)^m\big(\mathcal{Q}_{mn+k} - \mathcal{Q}_{k-m}\big)}{1 + (-t)^m - \mathcal{Q}_m}$$

and

$$\sum_{j=0}^{n} \mathcal{P}_{mj+k} = \frac{\mathcal{P}_k - \mathcal{P}_{mn+m+k} + (-t)^m\big(\mathcal{P}_{mn+k} - \mathcal{P}_{k-m}\big)}{1 + (-t)^m - \mathcal{Q}_m}$$

Proof. It is know that

$$I - (W^m)^{n+1} = (I - W^m)\sum_{j=0}^{n}(W^m)^j.$$

By Lemma 2.4 (i), we have

$$\det(I - W^m) = (1 - \frac{1}{2}\mathcal{Q}_m)^2 - (s^2+t)\mathcal{P}_m^2 = 1 + (-t)^m - \mathcal{Q}_m.$$

Since $\det(I - W^m) \neq 0$, we can write

$$\begin{aligned}(I - W^m)^{-1}\big(I - (W^m)^{n+1}\big)W^k &= \sum_{j=0}^{n} W^{mj+k} \\ &= \begin{bmatrix} \frac{1}{2}\sum_{j=0}^{n}\mathcal{Q}_{mj+k} & 2(s^2+t)\sum_{j=0}^{n}\mathcal{P}_{mj+k} \\ \frac{1}{2}\sum_{j=0}^{n}\mathcal{P}_{mj+k} & \frac{1}{2}\sum_{j=0}^{n}\mathcal{Q}_{mj+k} \end{bmatrix}. \quad (2.2)\end{aligned}$$

Since

$$(I - W^m)^{-1} = \frac{1}{1 + (-t)^m - \mathcal{Q}_m}\begin{bmatrix} 1 - \frac{1}{2}\mathcal{Q}_m & 2(s^2+t)\mathcal{P}_m \\ \frac{1}{2}\mathcal{P}_m & 1 - \frac{1}{2}\mathcal{Q}_m \end{bmatrix}$$

$$= \frac{1}{1+(-t)^m - \mathcal{Q}_m}\Big((1-\frac{1}{2}\mathcal{Q}_m)I + \frac{1}{2}\mathcal{P}_m H\Big),$$

we have

$$\begin{aligned}
&(I-W^m)^{-1}\big(I-(W^m)^{n+1}\big)W^k \\
&= \frac{\Big((1-\frac{1}{2}\mathcal{Q}_m)I + \frac{1}{2}\mathcal{P}_m H\Big)(W^k - W^{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} \\
&= \frac{\Big((1-\frac{1}{2}\mathcal{Q}_m)(W^k - W^{mn+m+k}) + \frac{1}{2}\mathcal{P}_m H(W^k - W^{mn+m+k})\Big)}{1+(-t)^m - \mathcal{Q}_m} \\
&= (1-\frac{1}{2}\mathcal{Q}_m)\begin{bmatrix} \frac{\frac{1}{2}(\mathcal{Q}_k - \mathcal{Q}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} & \frac{2(s^2+t)(\mathcal{P}_k - \mathcal{P}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} \\ \frac{\frac{1}{2}(\mathcal{P}_k - \mathcal{P}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} & \frac{\frac{1}{2}(\mathcal{Q}_k - \mathcal{Q}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} \end{bmatrix} \\
&\quad + \frac{1}{2}\mathcal{P}_m \begin{bmatrix} \frac{2(s^2+t)(\mathcal{P}_k - \mathcal{P}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} & \frac{2(s^2+t)(\mathcal{Q}_k - \mathcal{Q}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} \\ \frac{\frac{1}{2}(\mathcal{Q}_k - \mathcal{Q}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} & \frac{2(s^2+t)(\mathcal{P}_k - \mathcal{P}_{mn+m+k})}{1+(-t)^m - \mathcal{Q}_m} \end{bmatrix}
\end{aligned} \tag{2.3}$$

Using (2.2) and (2.3), we obtain

$$\begin{aligned}
&\sum_{j=0}^{n} \mathcal{Q}_{mj+k} \\
&= \frac{\Big((1-\frac{1}{2}\mathcal{Q}_m)(\mathcal{Q}_k - \mathcal{Q}_{mn+m+k}) + 2(s^2+t)\mathcal{P}_m(\mathcal{P}_k - \mathcal{P}_{mn+m+k})\Big)}{1+(-t)^m - \mathcal{Q}_m}.
\end{aligned} \tag{2.4}$$

By Lemma 2.4 (*iv*), (2.4) becomes

$$\sum_{j=0}^{n} \mathcal{Q}_{mj+k} = \frac{\mathcal{Q}_k - \mathcal{Q}_{mn+m+k} + (-t)^m\big(\mathcal{Q}_{mn+k} - \mathcal{Q}_{k-m}\big)}{1+(-t)^m - \mathcal{Q}_m}.$$

On the other hand, using (2.2) and (2.3) we get

$$\sum_{j=0}^{n} \mathcal{P}_{mj+k} = \frac{\Big((1-\frac{1}{2}\mathcal{Q}_m)(\mathcal{P}_k - \mathcal{P}_{mn+m+k}) + \frac{1}{2}\mathcal{P}_m(\mathcal{Q}_k - \mathcal{Q}_{mn+m+k})\Big)}{1+(-t)^m - \mathcal{Q}_m}. \tag{2.5}$$

By Lemma 2.4 (*v*), (2.5) becomes

$$\sum_{j=0}^{n} \mathcal{P}_{mj+k} = \frac{\mathcal{P}_k - \mathcal{P}_{mn+m+k} + (-t)^m\big(\mathcal{P}_{mn+k} - \mathcal{P}_{k-m}\big)}{1+(-t)^m - \mathcal{Q}_m}. \qquad \square$$

Theorem 2.7. *Let $n \in \mathbb{N}$ and $m, k \in \mathbb{Z}$ with $(-t)^m + \mathcal{Q}_m \neq -1$. If n is even, then*

$$\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k} = \frac{\mathcal{Q}_k + \mathcal{Q}_{mn+m+k} + (-t)^m\big(\mathcal{Q}_{mn+k} + \mathcal{Q}_{k-m}\big)}{1 + (-t)^m + \mathcal{Q}_m}$$

and

$$\sum_{j=0}^{n}(-1)^j \mathcal{P}_{mj+k} = \frac{\mathcal{P}_k + \mathcal{P}_{mn+m+k} + (-t)^m\big(\mathcal{P}_{mn+k} + \mathcal{P}_{k-m}\big)}{1 + (-t)^m + \mathcal{Q}_m}$$

Proof. Let n is an even natural number. Then we have

$$I + (W^m)^{n+1} = (I + W^m)\sum_{j=0}^{n}(-1)^j (W^m)^j.$$

By Lemma 2.4 (i), we have

$$\det(I + W^m) = (1 + \frac{1}{2}\mathcal{Q}_m)^2 - (s^2 + t)\mathcal{P}_m^2 = 1 + \mathcal{Q}_m + (-t)^m.$$

Since $\det(I + W^m) \neq 0$, we can write

$$\begin{aligned}
&(I + W^m)^{-1}\big(I + (W^m)^{n+1}\big)W^k \\
&= \sum_{j=0}^{n}(-1)^j W^{mj+k} \\
&= \begin{bmatrix} \frac{1}{2}\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k} & 2(s^2+t)\sum_{j=0}^{n}(-1)^j \mathcal{P}_{mj+k} \\ \frac{1}{2}\sum_{j=0}^{n}(-1)^j \mathcal{P}_{mj+k} & \frac{1}{2}\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k} \end{bmatrix}.
\end{aligned} \tag{2.6}$$

Since

$$\begin{aligned}
(I + W^m)^{-1} &= \frac{1}{1 + \mathcal{Q}_m + (-t)^m}\begin{bmatrix} 1 + \frac{1}{2}\mathcal{Q}_m & -2(s^2+t)\mathcal{P}_m \\ -\frac{1}{2}\mathcal{P}_m & 1 + \frac{1}{2}\mathcal{Q}_m \end{bmatrix} \\
&= \frac{1}{1 + \mathcal{Q}_m + (-t)^m}\Big((1 + \frac{1}{2}\mathcal{Q}_m)I - \frac{1}{2}\mathcal{P}_m H\Big),
\end{aligned}$$

we have

$$\begin{aligned}
&(I + W^m)^{-1}\big(I + (W^m)^{n+1}\big)W^k \\
&= \frac{\Big((1 + \frac{1}{2}\mathcal{Q}_m)I - \frac{1}{2}\mathcal{P}_m H\Big)(W^k + W^{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} \\
&= \frac{\Big((1 + \frac{1}{2}\mathcal{Q}_m)(W^k + W^{mn+m+k}) - \frac{1}{2}\mathcal{P}_m H(W^k + W^{mn+m+k})\Big)}{1 + \mathcal{Q}_m + (-t)^m}
\end{aligned}$$

$$= (1 + \frac{1}{2}\mathcal{Q}_m) \begin{bmatrix} \frac{\frac{1}{2}(\mathcal{Q}_k + \mathcal{Q}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} & \frac{2(s^2+t)(\mathcal{P}_k + \mathcal{P}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} \\ \frac{\frac{1}{2}(\mathcal{P}_k + \mathcal{P}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} & \frac{\frac{1}{2}(\mathcal{Q}_k + \mathcal{Q}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} \end{bmatrix}$$
$$- \frac{1}{2}\mathcal{P}_m \begin{bmatrix} \frac{2(s^2+t)(\mathcal{P}_k + \mathcal{P}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} & \frac{2(s^2+t)(\mathcal{Q}_k + \mathcal{Q}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} \\ \frac{\frac{1}{2}(\mathcal{Q}_k + \mathcal{Q}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} & \frac{2(s^2+t)(\mathcal{P}_k + \mathcal{P}_{mn+m+k})}{1 + \mathcal{Q}_m + (-t)^m} \end{bmatrix}. \quad (2.7)$$

Using (2.6) and (2.7), we obtain

$$\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k}$$
$$= \frac{\Big((1 + \frac{1}{2}\mathcal{Q}_m)(\mathcal{Q}_k + \mathcal{Q}_{mn+m+k}) - 2(s^2+t)\mathcal{P}_m(\mathcal{P}_k + \mathcal{P}_{mn+m+k})\Big)}{1 + \mathcal{Q}_m + (-t)^m}. \quad (2.8)$$

By Lemma 2.4 (*iv*), (2.8) becomes

$$\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k} = \frac{\mathcal{Q}_k + \mathcal{Q}_{mn+m+k} + (-t)^m\big(\mathcal{Q}_{k-m} + \mathcal{Q}_{mn+k}\big)}{1 + (-t)^m + \mathcal{Q}_m}.$$

Similarly it can be easily seen that

$$\sum_{j=0}^{n}(-1)^j \mathcal{P}_{mj+k} = \frac{\mathcal{P}_k + \mathcal{P}_{mn+m+k} + (-t)^m\big(\mathcal{P}_{k-m} + \mathcal{P}_{mn+k}\big)}{1 + (-t)^m + \mathcal{Q}_m}. \qquad \square$$

Theorem 2.8. *Let $n \in \mathbb{N}$ and $m, k \in \mathbb{Z}$ with $(-t)^m + \mathcal{Q}_m \neq -1$. If n is odd, then*

$$\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k} = \frac{\mathcal{Q}_k - \mathcal{Q}_{mn+m+k} + (-t)^m\big(\mathcal{Q}_{k-m} - \mathcal{Q}_{mn+k}\big)}{1 + (-t)^m + \mathcal{Q}_m}$$

and

$$\sum_{j=0}^{n}(-1)^j \mathcal{P}_{mj+k} = \frac{\mathcal{P}_k - \mathcal{P}_{mn+m+k} + (-t)^m\big(\mathcal{P}_{k-m} - \mathcal{P}_{mn+k}\big)}{1 + (-t)^m + \mathcal{Q}_m}$$

Proof. Let n is an odd natural number. Then we get

$$\sum_{j=0}^{n}(-1)^j \mathcal{Q}_{mj+k} = \sum_{j=0}^{n-1}(-1)^j \mathcal{Q}_{mj+k} - \mathcal{Q}_{mn+k}.$$

Since n is an odd natural number then $n-1$ is even. By Thorem 2.7, we have

$$\sum_{j=0}^{n-1}(-1)^j \mathcal{Q}_{mj+k} = \frac{\mathcal{Q}_k + \mathcal{Q}_{mn+k} + (-t)^m\big(\mathcal{Q}_{mn+k-m} + \mathcal{Q}_{k-m}\big)}{1 + (-t)^m + \mathcal{Q}_m}$$

and

$$\sum_{j=0}^{n}(-1)^j\mathcal{Q}_{mj+k}$$
$$=\frac{\mathcal{Q}_k+(-t)^m\big(\mathcal{Q}_{mn+k-m}+\mathcal{Q}_{k-m}\big)-(-t)^m\mathcal{Q}_{mn+k}-\mathcal{Q}_{mn+k}\mathcal{Q}_m}{1+(-t)^m+\mathcal{Q}_m}. \tag{2.9}$$

Using Lemma 2.4 (vi) in (2.9), we get

$$\sum_{j=0}^{n}(-1)^j\mathcal{Q}_{mj+k}=\frac{\mathcal{Q}_k-\mathcal{Q}_{mn+m+k}+(-t)^m\big(\mathcal{Q}_{k-m}-\mathcal{Q}_{mn+k}\big)}{1+(-t)^m+\mathcal{Q}_m}.$$

In a similar way, it can be seen that

$$\sum_{j=0}^{n}(-1)^j\mathcal{P}_{mj+k}=\sum_{j=0}^{n-1}(-1)^j\mathcal{P}_{mj+k}-\mathcal{P}_{mn+k}.$$

By Theorem 2.7, it follows that

$$\sum_{j=0}^{n}(-1)^j\mathcal{P}_{mj+k}$$
$$=\frac{\mathcal{P}_k+(-t)^m\big(\mathcal{P}_{mn+k-m}+\mathcal{P}_{k-m}\big)-(-t)^m\mathcal{P}_{mn+k}-\mathcal{P}_{mn+k}\mathcal{Q}_m}{1+(-t)^m+\mathcal{Q}_m}. \tag{2.10}$$

Using Lemma 2.4 (vii) in (2.10), we obtain

$$\sum_{j=0}^{n}(-1)^j\mathcal{P}_{mj+k}=\frac{\mathcal{P}_k-\mathcal{P}_{mn+m+k}+(-t)^m\big(\mathcal{P}_{k-m}-\mathcal{P}_{mn+k}\big)}{1+(-t)^m+\mathcal{Q}_m}. \qquad \square$$

Acknowledgements. The authors would like to thank the faculty of science and technology, Rajamangala University of Technology Thanyaburi (RMUTT), Thailand for the financial support. Moreover, the authors would like to thank the referees for their valuable suggestions and comments which helped to improve the quality and readability of the paper.

References

[1] H.H. Gulec, N. Taskara, On the (s,t)-Pell and (s,t)-Pell Lucas sequences and their matrix representations, Applied Mathematics Letters, Vol. 25 (2012), 1554–1559.

[2] T. Koshy, Pell and Pell-Lucas Numbers with Applications. Springer, Berlin (2014)

[3] T. Koshy, Fibonacci and Lucas Numbers with Applications, John Wiley and Sons Inc., New York (2001).

[4] M.S. El Naschie, The Fibonacci Code behind Super Strings and P-Branes: An Answer to M. Kaku's Fundamental Question. Chaos, Solitons & Fractals, Vol. 31 (2007), 537–547.

[5] M.S. El Naschie, Notes on Superstrings and the Infinite Sums of Fibonacci and Lucas Numbers. Chaos, Solitons & Fractals, Vol. 12 (2001), 1937–1940.

[6] A.P. Stakhov, Fibonacci Matrices: A Generalization of the "Cassini Formula" and a New Coding Theory. Chaos, Solitons Fractals, Vol. 30 (2006), 56–66.

[7] A.P. Stakhov, The Generalized Principle of the Golden Section and Its Applications in Mathematics, Science and Engineering. Chaos, Solitons & Fractals, Vol. 26 (2005), 263–289.

Annales Mathematicae et Informaticae
46 (2016) pp. 205–216
http://ami.ektf.hu

Convolution of second order linear recursive sequences I.

Tamás Szakács

Eszterházy Károly University
szakacs.tamas@uni-eszterhazy.hu

Submitted October 28, 2016 — Accepted December 1, 2016

Abstract

In this paper, we deal with convolutions of second order linear recursive sequences and give some special convolutions for Fibonacci-, Pell-, Jacobsthal- and Mersenne-sequences and their associated sequences.

Keywords: convolution, Fibonacci, generating function

MSC: 11B37, 11B39

1. Introduction

Let A, B be given real numbers with $AB \neq 0$. A second order linear recursive sequence $\{G_n\}_{n=0}^{\infty}$ is defined by the recursion

$$G_n = AG_{n-1} + BG_{n-2} \quad (n \geq 2),$$

where the initial terms G_0, G_1 are fixed real numbers with $|G_0| + |G_1| \neq 0$. For brevity, we use the following notation $G_n(G_0, G_1, A, B)$, too. The polynomial

$$p(x) = x^2 - Ax - B \tag{1.1}$$

is said to be the characteristic polynomial of the sequence $\{G_n\}_{n=0}^{\infty}$. If $D = A^2 + 4B \neq 0$ then the Binet formula of $\{G_n\}_{n=0}^{\infty}$ is

$$G_n = \frac{G_1 - \beta G_0}{\alpha - \beta}\alpha^n - \frac{G_1 - \alpha G_0}{\alpha - \beta}\beta^n,$$

where α, β are distinct roots of the characteristic polynomial. If $G_0 = 0$ and $G_1 = 1$ then $\{G_n\}_{n=0}^{\infty}$ is known as R-sequence $\{R_n\}_{n=0}^{\infty}$ with it's Binet formula

$$R_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}. \tag{1.2}$$

If $G_0 = 2$ and $G_1 = A$ then the sequence is known as associated-R, or R-Lucas sequence $\{V_n\}_{n=0}^{\infty}$ with it's Binet formula

$$V_n = \alpha^n + \beta^n. \tag{1.3}$$

In the following sections, we will use the generating function and partial-fraction decomposition for the proofs. The generating function of $\{G_n\}_{n=0}^{\infty}$ (which can easily be verified by the well known methods) is

$$g(x) = \frac{G_0 + (G_1 - AG_0)x}{1 - Ax - Bx^2}. \tag{1.4}$$

The following table contains some special, well-known sequences with their initial terms, characteristic polynomial and generating function, where P-Lucas, J-Lucas and M-Lucas sequences are the associated sequences of Pell, Jacobsthal and Mersenne sequences, respectively.

Name	$G_n(G_0, G_1, A, B)$	Characteristic polynomial	Gen. function
Fibonacci	$F_n(0,1,1,1)$	$p(x) = x^2 - x - 1$	$g(x) = \frac{x}{1-x-x^2}$
Pell	$P_n(0,1,2,1$	$p(x) = x^2 - 2x - 1$	$g(x) = \frac{x}{1-2x-x^2}$
Jacobsthal	$J_n(0,1,1,2)$	$p(x) = x^2 - x - 2$	$g(x) = \frac{x}{1-x-2x^2}$
Mersenne	$M_n(0,1,3,-2)$	$p(x) = x^2 - 3x + 2$	$g(x) = \frac{x}{1-3x+2x^2}$
Lucas	$L_n(2,1,1,1)$	$p(x) = x^2 - x - 1$	$g(x) = \frac{2-x}{1-x-x^2}$
P-Lucas	$p_n(2,2,2,1)$	$p(x) = x^2 - 2x - 1$	$g(x) = \frac{2-2x}{1-2x-x^2}$
J-Lucas	$j_n(2,1,1,2)$	$p(x) = x^2 - x - 2$	$g(x) = \frac{2-x}{1-x-2x^2}$
M-Lucas	$m_n(2,3,3,-2)$	$p(x) = x^2 - 3x + 2$	$g(x) = \frac{2-3x}{1-3x+2x^2}$

Table 1: Named sequences

For further generating functions for second order linear recursive sequences see the paper of Mező [3].

We consider the sequence $\{c(n)\}_{n=0}^{\infty}$ given by the convolution of two different second order linear recursive sequences $\{G_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=0}^{\infty}$:

$$c(n) = \sum_{k=0}^{n} G_k H_{n-k}.$$

Griffiths and Bramham [1] investigated the convolution of Lucas- and Jacobsthal-numbers and got the result:

$$c(n) = j_{n+1} - L_{n+1},$$

which can be found in the OEIS [2] with the following id: A264038.

In this paper, we deal with convolution of two different sequences, where all of the roots are distinct and the sequences are R-sequences or R-Lucas sequences. The convolution of sequences with themselves was investigated by Zhang W., Zhang Z., He P., Feng H. and many others. In [5], Feng and Zhang Z. generalized the previous results, i.e. they evaluated the following summation:

$$\sum_{a_1+a_2+\cdots+a_k=n} W_{ma_1}W_{ma_2}\cdots W_{ma_k}.$$

For example, the convolution of Fibonacci numbers with themselves was given as a corollary in [4] by Zhang W.:

$$\sum_{a+b=n} F_aF_b = \frac{1}{5}\left[(n-1)F_n + 2nF_{n-1}\right], \quad n \geq 1.$$

2. Results

In this section, we present three theorems and give formulas for $\{c(n)\}_{n=0}^{\infty}$, where the formulas depend only on the initial terms and the roots of the characteristic polynomials. After each theorem, we show the special cases of the theorem in corollaries using the named sequences (Fibonacci, Pell, Jacobsthal, Mersenne, Lucas, P-Lucas, J-Lucas, M-Lucas).

In this paper –for brevity–, we use the following notations:

$$\begin{aligned} a &= (A_1 - A_2)\alpha + B_1 - B_2, \\ b &= (A_1 - A_2)\beta + B_1 - B_2, \\ c &= (A_2 - A_1)\gamma + B_2 - B_1, \\ d &= (A_2 - A_1)\delta + B_2 - B_1, \end{aligned} \tag{2.1}$$

where $abcd \neq 0$, α, β and γ, δ are distinct roots of the characteristic polynomial of $\{G_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=0}^{\infty}$, respectively. We suppose that all the roots are real numbers and the characteristic polynomials have no common roots.

In the following theorem, we deal with the convolution of two different R-sequences.

Theorem 2.1. *The convolution of* $G_n(0, 1, A_1, B_1)$ *and* $H_n(0, 1, A_2, B_2)$ *is*

$$c(n) = \sum_{k=0}^{n} G_kH_{n-k} = \frac{\frac{\alpha^{n+1}}{a} - \frac{\beta^{n+1}}{b}}{\alpha - \beta} + \frac{\frac{\gamma^{n+1}}{c} - \frac{\delta^{n+1}}{d}}{\gamma - \delta}.$$

For the well-known sequences, listed in Table 1, we can get special convolution formulas:

Corollary 2.2. *Using Theorem 2.1 the convolution of Fibonacci and Pell numbers is:*

$$c(n) = \sum_{k=0}^{n} F_k P_{n-k} = P_n - F_n.$$

Remark 2.3. In [2], (A106515) it can be found that

$$c(n) = \sum_{k=0}^{n} F_{n-k-1} P_{k+1} = P_n - F_n + P_{n+1},$$

where because of the different indices the term P_{n+1} occures, as well.

Corollary 2.4. *Using Theorem 2.1 the convolution of Fibonacci and Jacobsthal numbers is:*

$$c(n) = \sum_{k=0}^{n} F_k J_{n-k} = J_{n+1} - F_{n+1}.$$

Remark 2.5. In [2], (A094687) the formula

$$c(n) = \sum_{k=0}^{n} F_k J_{n-k} = c(n-1) + 2c(n-2) + F_{n-1}$$

can be found. After a short calculation one can easily verify that the two formulas for $c(n)$ are the same ones.

Corollary 2.6. *Using Theorem 2.1 the convolution of Fibonacci and Mersenne numbers is:*

$$c(n) = \sum_{k=0}^{n} F_k M_{n-k} = m_{n+1} - F_{n+4}.$$

Corollary 2.7. *Using Theorem 2.1 the convolution of Pell and Jacobsthal numbers is:*

$$c(n) = \sum_{k=0}^{n} P_k J_{n-k} = \frac{P_{n+1} + P_n - J_{n+2}}{2}.$$

Corollary 2.8. *Using Theorem 2.1 the convolution of Pell and Mersenne numbers is:*

$$c(n) = \sum_{k=0}^{n} P_k M_{n-k} = \frac{P_{n+2} + P_{n+1} - M_{n+2}}{2}.$$

In the following theorem, we deal with the convolution of an R-sequence and an R-Lucas sequence.

Theorem 2.9. *The convolution of $G_n(0,1,A_1,B_1)$ and $H_n(2,A_2,A_2,B_2)$ is*

$$c(n) = \sum_{k=0}^{n} G_k H_{n-k} =$$

$$= \frac{\frac{\alpha^{n+1}(2\alpha-A_2)}{a} - \frac{\beta^{n+1}(2\beta-A_2)}{b}}{\alpha-\beta} + \frac{\frac{\gamma^{n+1}(2\gamma-A_2)}{c} - \frac{\delta^{n+1}(2\delta-A_2)}{d}}{\gamma-\delta}.$$

For the well-known sequences, listed in Table 1, we can get special convolution formulas:

Corollary 2.10. *Using Theorem 2.9 the convolution of Fibonacci and P-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} F_k p_{n-k} = p_n - 2F_{n-1}.$$

Corollary 2.11. *Using Theorem 2.9 the convolution of Fibonacci and J-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} F_k j_{n-k} = j_{n+1} - L_{n+1}.$$

Remark 2.12. This our convolution has the same form as of Griffiths and Bramham in [1].

Corollary 2.13. *Using Theorem 2.9 the convolution of Fibonacci and M-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} F_k m_{n-k} = M_{n+1} - F_{n+1}.$$

Remark 2.14. For the sequence $a(n)$ (A228078 in [2]), where $a(n+1)$ is the sum of n-th row of the Fibonacci-Pascal triangle in A228074, we get that

$$c(n) = a(n+1).$$

Corollary 2.15. *Using Theorem 2.9 the convolution of Pell and Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} P_k L_{n-k} = P_n + p_n - L_n.$$

Corollary 2.16. *Using Theorem 2.9 the convolution of Pell and J-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} P_k j_{n-k} = \frac{8P_{n+1} + p_{n+1} - 2j_{n+2}}{4}.$$

Corollary 2.17. *Using Theorem 2.9 the convolution of Pell and M-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} P_k m_{n-k} = \frac{4P_{n+2} + p_{n+1} - 2m_{n+2}}{4}.$$

Corollary 2.18. *Using Theorem 2.9 the convolution of Jacobsthal and Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} J_k L_{n-k} = j_{n+1} - L_{n+1}.$$

Remark 2.19. The convolution of Lucas and Jacobsthal numbers was also investigated by Griffiths and Bramham in [1], the two formulas are the same ones.

Corollary 2.20. *Using Theorem 2.9 the convolution of Jacobsthal and P-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} J_k p_{n-k} = 2(P_{n+1} - J_{n+1}).$$

Corollary 2.21. *Using Theorem 2.9 the convolution of Mersenne and Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} M_k L_{n-k} = 3m_{n+1} - L_{n+4} - 2.$$

Corollary 2.22. *Using Theorem 2.9 the convolution of Mersenne and P-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} M_k p_{n-k} = \frac{3p_{n+1} + p_n - M_{n+3} - 1}{2}.$$

In the following theorem, we deal with the convolution of two different R-Lucas sequences.

Theorem 2.23. *The convolution of $G_n(2, A_1, A_1, B_1)$ and $H_n(2, A_2, A_2, B_2)$ is*

$$\begin{aligned} c(n) &= \sum_{k=0}^{n} G_k H_{n-k} = \\ &= \frac{\frac{\alpha^{n+1}(2\alpha-A_1)(2\alpha-A_2)}{a} - \frac{\beta^{n+1}(2\beta-A_1)(2\beta-A_2)}{b}}{\alpha-\beta} \\ &\quad + \frac{\frac{\gamma^{n+1}(2\gamma-A_1)(2\gamma-A_2)}{c} - \frac{\delta^{n+1}(2\delta-A_1)(2\delta-A_2)}{d}}{\gamma-\delta}. \end{aligned}$$

For the well-known sequences, listed in Table 1, we can get special convolution formulas:

Corollary 2.24. *Using Theorem 2.23 the convolution of Lucas and P-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} L_k p_{n-k} = 2F_{n+1} - 6F_n + 2P_{n+1} + 6P_n.$$

Corollary 2.25. *Using Theorem 2.23 the convolution of Lucas and J-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} L_k j_{n-k} = 9J_{n+1} - 5F_{n+1}.$$

Corollary 2.26. *Using Theorem 2.23 the convolution of Lucas and M-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} L_k m_{n-k} = 3M_{n+1} - L_{n+1} + 2.$$

Corollary 2.27. *Using Theorem 2.23 the convolution of P-Lucas and J-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} p_k j_{n-k} = 2P_{n+2} + p_{n+1} - 2j_{n+1}.$$

Corollary 2.28. *Using Theorem 2.23 the convolution of P-Lucas and M-Lucas numbers is:*

$$c(n) = \sum_{k=0}^{n} p_k m_{n-k} = 2P_{n+2} + 4P_{n+1} - M_{n+2} - 1.$$

3. Proofs

In the following proofs, we use the method of partial-fraction decomposition, the generating functions of second order linear recursive sequences and the idea used by Griffiths and Bramham in [1], that is $c(n)$ is the coefficient of x^n in

$$g(x)h(x) = \sum_{n=0}^{\infty} G_n x^n \cdot \sum_{n=0}^{\infty} H_n x^n = \sum_{n=0}^{\infty} c(n) x^n,$$

where $g(x)$, $h(x)$ are the generating functions of sequences $\{G_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=0}^{\infty}$, respectively.

Proof of Theorem 2.1. Using (1.4), the generating functions of the sequences $G_n(0, 1, A_1, B_1)$ and $H_n(0, 1, A_2, B_2)$ are

$$g(x) = \frac{x}{1 - A_1 x - B_1 x^2} = \frac{x}{(1 - \alpha x)(1 - \beta x)}$$

and

$$h(x) = \frac{x}{1 - A_2 x - B_2 x^2} = \frac{x}{(1 - \gamma x)(1 - \delta x)},$$

where α, β and γ, δ are the roots of the characteristic polynomial of $\{G_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=0}^{\infty}$, respectively. The generating functions can be written as (by the method of partial-fraction decomposition)

$$g(x) = \frac{1}{\alpha - \beta}\left(\frac{1}{1 - \alpha x} - \frac{1}{1 - \beta x}\right)$$

and

$$h(x) = \frac{1}{\gamma - \delta}\left(\frac{1}{1 - \gamma x} - \frac{1}{1 - \delta x}\right).$$

From this it follows that

$$g(x)h(x)(\alpha - \beta)(\gamma - \delta)$$

$$
\begin{aligned}
&= \left(\frac{1}{1-\alpha x} - \frac{1}{1-\beta x}\right)\left(\frac{1}{1-\gamma x} - \frac{1}{1-\delta x}\right) \\
&= \frac{1}{(1-\alpha x)(1-\gamma x)} - \frac{1}{(1-\alpha x)(1-\delta x)} - \frac{1}{(1-\beta x)(1-\gamma x)} + \frac{1}{(1-\beta x)(1-\delta x)} \\
&= \frac{\frac{\alpha}{\alpha-\gamma}}{1-\alpha x} - \frac{\frac{\gamma}{\alpha-\gamma}}{1-\gamma x} - \frac{\frac{\alpha}{\alpha-\delta}}{1-\alpha x} + \frac{\frac{\delta}{\alpha-\delta}}{1-\delta x} - \frac{\frac{\beta}{\beta-\gamma}}{1-\beta x} + \frac{\frac{\gamma}{\beta-\gamma}}{1-\gamma x} + \frac{\frac{\beta}{\beta-\delta}}{1-\beta x} - \frac{\frac{\delta}{\beta-\delta}}{1-\delta x} \\
&= \frac{\frac{\alpha(\gamma-\delta)}{(A_1-A_2)\alpha+B_1-B_2}}{1-\alpha x} - \frac{\frac{\beta(\gamma-\delta)}{(A_1-A_2)\beta+B_1-B_2}}{1-\beta x} + \frac{\frac{\gamma(\alpha-\beta)}{(A_2-A_1)\gamma+B_2-B_1}}{1-\gamma x} - \frac{\frac{\delta(\alpha-\beta)}{(A_2-A_1)\delta+B_2-B_1}}{1-\delta x}.
\end{aligned}
$$

Now using that $c(n)$ is the coefficient of x^n in $g(x)h(x)$ and e.g.,

$$\frac{1}{1-\alpha x} = \sum_{n=0}^{\infty} (\alpha x)^n \quad (0 < |\alpha x| < 1),$$

we get

$$
\begin{aligned}
c(n) = \frac{1}{\alpha-\beta}&\left(\frac{\alpha^{n+1}}{(A_1-A_2)\alpha+B_1-B_2} - \frac{\beta^{n+1}}{(A_1-A_2)\beta+B_1-B_2}\right) \\
&+ \frac{1}{\gamma-\delta}\left(\frac{\gamma^{n+1}}{(A_2-A_1)\gamma+B_2-B_1} - \frac{\delta^{n+1}}{(A_2-A_1)\delta+B_2-B_1}\right). \qquad \square
\end{aligned}
$$

We remark that the corollaries can be obtained from Table 1 if we use the values of A_1, B_1, A_2, B_2 and the Binet formula (1.2), e.g., the proof of Corollary 2.2:

Proof of Corollary 2.2. Now $G_n = F_n(0,1,1,1)$ and $H_n = P_n(0,1,2,1)$.

$$\alpha, \beta = \frac{1 \pm \sqrt{5}}{2}, \quad \gamma, \delta = 1 \pm \sqrt{2}.$$

By (2.1), we get that

$$
\begin{aligned}
a &= -\alpha, \\
b &= -\beta, \\
c &= \gamma, \\
d &= \delta.
\end{aligned}
$$

Applying Theorem 2.1 and (1.2), we get the result

$$c(n) = \frac{\frac{\alpha^{n+1}}{a} - \frac{\beta^{n+1}}{b}}{\alpha-\beta} + \frac{\frac{\gamma^{n+1}}{c} - \frac{\delta^{n+1}}{d}}{\gamma-\delta} = \frac{-\alpha^n + \beta^n}{\alpha-\beta} + \frac{\gamma^n - \delta^n}{\gamma-\delta} = P_n - F_n. \qquad \square$$

Proof of Theorem 2.9. Using (1.4), the generating functions of the sequences $G_n(0,1,A_1,B_1)$ and $H_n(2,A_2,A_2,B_2)$ are

$$g(x) = \frac{x}{1-A_1x-B_1x^2} = \frac{x}{(1-\alpha x)(1-\beta x)}$$

and

$$h(x) = \frac{2 - A_2x}{1 - A_2x - B_2x^2} = \frac{2 - A_2x}{(1 - \gamma x)(1 - \delta x)},$$

where α, β and γ, δ are the roots of the characteristic polynomial of $\{G_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=0}^{\infty}$, respectively. The generating functions could be written as (by the method of partial-fraction decomposition)

$$g(x) = \frac{1}{\alpha - \beta}\left(\frac{1}{1 - \alpha x} - \frac{1}{1 - \beta x}\right)$$

and

$$h(x) = \frac{1}{\gamma - \delta}\left(\frac{2\gamma - A_2}{1 - \gamma x} - \frac{2\delta - A_2}{1 - \delta x}\right).$$

From this it follows that

$$\begin{aligned}
&g(x)h(x)(\alpha-\beta)(\gamma-\delta)\\
&= \left(\frac{1}{1-\alpha x} - \frac{1}{1-\beta x}\right)\left(\frac{2\gamma - A_2}{1-\gamma x} - \frac{2\delta - A_2}{1-\delta x}\right)\\
&= \frac{2\gamma - A_2}{(1-\alpha x)(1-\gamma x)} - \frac{2\delta - A_2}{(1-\alpha x)(1-\delta x)} - \frac{2\gamma - A_2}{(1-\beta x)(1-\gamma x)} + \frac{2\delta - A_2}{(1-\beta x)(1-\delta x)}\\
&= \frac{\frac{\alpha(2\delta-A_2)}{\alpha-\gamma}}{1-\alpha x} - \frac{\frac{\gamma(2\delta-A_2)}{\alpha-\gamma}}{1-\gamma x} - \frac{\frac{\alpha(2\delta-A_2)}{\alpha-\delta}}{1-\alpha x} + \frac{\frac{\delta(2\delta-A_2)}{\alpha-\delta}}{1-\delta x}\\
&\quad - \frac{\frac{\beta(2\delta-A_2)}{\beta-\gamma}}{1-\beta x} + \frac{\frac{\gamma(2\delta-A_2)}{\beta-\gamma}}{1-\gamma x} + \frac{\frac{\beta(2\delta-A_2)}{\beta-\delta}}{1-\beta x} - \frac{\frac{\delta(2\delta-A_2)}{\beta-\delta}}{1-\delta x}\\
&= \frac{\frac{\alpha(\gamma-\delta)(2\alpha-A_2)}{(A_1-A_2)\alpha+B_1-B_2}}{1-\alpha x} - \frac{\frac{\beta(\gamma-\delta)(2\beta-A_2)}{(A_1-A_2)\beta+B_1-B_2}}{1-\beta x} + \frac{\frac{\gamma(\alpha-\beta)(2\gamma-A_2)}{(A_2-A_1)\gamma+B_2-B_1}}{1-\gamma x} - \frac{\frac{\delta(\alpha-\beta)(2\delta-A_2)}{(A_2-A_1)\delta+B_2-B_1}}{1-\delta x}.
\end{aligned}$$

Now using that $c(n)$ is the coefficient of x^n in $g(x)h(x)$ and e.g.,

$$\frac{1}{1-\alpha x} = \sum_{n=0}^{\infty} (\alpha x)^n \quad (0 < |\alpha x| < 1),$$

we get

$$\begin{aligned}
c(n) = &\frac{1}{\alpha - \beta}\left(\frac{\alpha^{n+1}(2\alpha - A_2)}{(A_1 - A_2)\alpha + B_1 - B_2} - \frac{\beta^{n+1}(2\beta - A_2)}{(A_1 - A_2)\beta + B_1 - B_2}\right)\\
&+ \frac{1}{\gamma - \delta}\left(\frac{\gamma^{n+1}(2\gamma - A_2)}{(A_2 - A_1)\gamma + B_2 - B_1} - \frac{\delta^{n+1}(2\delta - A_2)}{(A_2 - A_1)\delta + B_2 - B_1}\right). \qquad \square
\end{aligned}$$

We remark that the corollaries can be obtained from Table 1 if we use the values of A_1, B_1, A_2, B_2 and the Binet formulas ((1.2) or (1.3)), e.g., the proof of Corollary 2.10:

Proof of Corollary 2.10. Now $G_n = F_n(0,1,1,1)$ and $H_n = p_n(2,2,2,1)$.

$$\alpha, \beta = \frac{1 \pm \sqrt{5}}{2}, \quad \gamma, \delta = 1 \pm \sqrt{2}.$$

By (2.1), we get that

$$\begin{aligned} a &= -\alpha, \\ b &= -\beta, \\ c &= \gamma, \\ d &= \delta. \end{aligned}$$

Applying Theorem 2.9, (1.2) and (1.3), we get the result

$$\begin{aligned} c(n) &= \frac{\frac{\alpha^{n+1}(2\alpha - A_2)}{a} - \frac{\beta^{n+1}(2\beta - A_2)}{b}}{\alpha - \beta} + \frac{\frac{\gamma^{n+1}(2\gamma - A_2)}{c} - \frac{\delta^{n+1}(2\delta - A_2)}{d}}{\gamma - \delta} \\ &= \frac{\alpha^n(1-\sqrt{5}) - \beta^n(1+\sqrt{5})}{\alpha - \beta} + \frac{\gamma^n 2\sqrt{2} + \delta^n 2\sqrt{2}}{\gamma - \delta} \\ &= \frac{\alpha^{n-1}(-2) - \beta^{n-1}(-2)}{\alpha - \beta} + \gamma^n + \delta^n = p_n - 2F_{n-1}. \end{aligned}$$ □

Proof of Theorem 2.23. Using (1.4), the generating functions of the sequences $G_n(2, A_1, A_1, B_1)$ and $H_n(2, A_2, A_2, B_2)$ are

$$g(x) = \frac{2 - A_1 x}{1 - A_1 x - B_1 x^2} = \frac{2 - A_1 x}{(1 - \alpha x)(1 - \beta x)}$$

and

$$h(x) = \frac{2 - A_2 x}{1 - A_2 x - B_2 x^2} = \frac{2 - A_2 x}{(1 - \gamma x)(1 - \delta x)},$$

where α, β and γ, δ are the roots of the characteristic polynomial of $\{G_n\}_{n=0}^{\infty}$ and $\{H_n\}_{n=0}^{\infty}$, respectively. The generating functions could be written as (by the method of partial-fraction decomposition)

$$g(x) = \frac{1}{\alpha - \beta}\left(\frac{2\alpha - A_1}{1 - \alpha x} - \frac{2\beta - A_1}{1 - \beta x}\right)$$

and

$$h(x) = \frac{1}{\gamma - \delta}\left(\frac{2\gamma - A_2}{1 - \gamma x} - \frac{2\delta - A_2}{1 - \delta x}\right).$$

From this it follows that

$$\begin{aligned} &g(x)h(x)(\alpha - \beta)(\gamma - \delta) \\ &= \left(\frac{2\alpha - A_1}{1 - \alpha x} - \frac{2\beta - A_1}{1 - \beta x}\right)\left(\frac{2\gamma - A_2}{1 - \gamma x} - \frac{2\delta - A_2}{1 - \delta x}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{(2\alpha - A_1)(2\gamma - A_2)}{(1-\alpha x)(1-\gamma x)} - \frac{(2\alpha - A_1)(2\delta - A_2)}{(1-\alpha x)(1-\delta x)} \\
&\quad - \frac{(2\beta - A_1)(2\gamma - A_2)}{(1-\beta x)(1-\gamma x)} + \frac{(2\beta - A_1)(2\delta - A_2)}{(1-\beta x)(1-\delta x)} \\
&= \frac{\frac{\alpha(2\alpha-A_1)(2\gamma-A_2)}{\alpha-\gamma}}{1-\alpha x} - \frac{\frac{\gamma(2\alpha-A_1)(2\gamma-A_2)}{\alpha-\gamma}}{1-\gamma x} - \frac{\frac{\alpha(2\alpha-A_1)(2\delta-A_2)}{\alpha-\delta}}{1-\alpha x} + \frac{\frac{\delta(2\alpha-A_1)(2\delta-A_2)}{\alpha-\delta}}{1-\delta x} \\
&\quad - \frac{\frac{\beta(2\beta-A_1)(2\gamma-A_2)}{\beta-\gamma}}{1-\beta x} + \frac{\frac{\gamma(2\beta-A_1)(2\gamma-A_2)}{\beta-\gamma}}{1-\gamma x} + \frac{\frac{\beta(2\beta-A_1)(2\delta-A_2)}{\beta-\delta}}{1-\beta x} - \frac{\frac{\delta(2\beta-A_1)(2\delta-A_2)}{\beta-\delta}}{1-\delta x} \\
&= \frac{\frac{\alpha(\gamma-\delta)(2\alpha-A_1)(2\alpha-A_2)}{(A_1-A_2)\alpha+B_1-B_2}}{1-\alpha x} - \frac{\frac{\beta(\gamma-\delta)(2\beta-A_1)(2\beta-A_2)}{(A_1-A_2)\beta+B_1-B_2}}{1-\beta x} \\
&\quad + \frac{\frac{\gamma(\alpha-\beta)(2\gamma-A_1)(2\gamma-A_2)}{(A_2-A_1)\gamma+B_2-B_1}}{1-\gamma x} - \frac{\frac{\delta(\alpha-\beta)(2\delta-A_1)(2\delta-A_2)}{(A_2-A_1)\delta+B_2-B_1}}{1-\delta x}.
\end{aligned}
$$

Now using that $c(n)$ is the coefficient of x^n in $g(x)h(x)$ and e.g.,

$$\frac{1}{1-\alpha x} = \sum_{n=0}^{\infty} (\alpha x)^n \quad (0 < |\alpha x| < 1),$$

we get

$$
\begin{aligned}
c(n) &= \frac{1}{\alpha-\beta}\left(\frac{\alpha^{n+1}(2\alpha - A_1)(2\alpha - A_2)}{(A_1 - A_2)\alpha + B_1 - B_2} - \frac{\beta^{n+1}(2\beta - A_1)(2\beta - A_2)}{(A_1 - A_2)\beta + B_1 - B_2}\right) \\
&\quad + \frac{1}{\gamma-\delta}\left(\frac{\gamma^{n+1}(2\gamma - A_1)(2\gamma - A_2)}{(A_2 - A_1)\gamma + B_2 - B_1} - \frac{\delta^{n+1}(2\delta - A_1)(2\delta - A_2)}{(A_2 - A_1)\delta + B_2 - B_1}\right). \qquad \square
\end{aligned}
$$

We remark that the corollaries can be obtained from Table 1 if we use the values of A_1, B_1, A_2, B_2 and the Binet formula (1.2), e.g., the proof of Corollary 2.24:

Proof of Corollary 2.24. Now $G_n = L_n(2, 1, 1, 1)$ and $H_n = p_n(2, 2, 2, 1)$.

$$\alpha, \beta = \frac{1 \pm \sqrt{5}}{2}, \quad \gamma, \delta = 1 \pm \sqrt{2}.$$

By (2.1), we get that

$$
\begin{aligned}
a &= -\alpha, \\
b &= -\beta, \\
c &= \gamma, \\
d &= \delta.
\end{aligned}
$$

Applying Theorem 2.1, (1.1) and (1.2), we get the result

$$c(n) = \frac{\frac{\alpha^{n+1}(2\alpha-A_1)(2\alpha-A_2)}{a} - \frac{\beta^{n+1}(2\beta-A_1)(2\beta-A_2)}{b}}{\alpha - \beta}$$

$$
\begin{aligned}
&+\frac{\frac{\gamma^{n+1}(2\gamma-A_1)(2\gamma-A_2)}{c}-\frac{\delta^{n+1}(2\delta-A_1)(2\delta-A_2)}{d}}{\gamma-\delta}\\
&=\frac{-\alpha^n(4\alpha^2-6\alpha+2)+\beta^n(4\beta^2-6\beta+2)}{\alpha-\beta}\\
&+\frac{\gamma^n(4\gamma^2-6\gamma+2)-\delta^n(4\delta^2-6\delta+2)}{\gamma-\delta}\\
&=\frac{-\alpha^n(-2\alpha+6)+\beta^n(-2\beta+6)}{\alpha-\beta}\\
&+\frac{\gamma^n(2\gamma+6)-\delta^n(2\delta+6)}{\gamma-\delta}=2F_{n+1}-6F_n+2P_{n+1}+6P_n. \qquad \square
\end{aligned}
$$

4. Concluding remarks

In this paper, we have dealt the case, when there are no common roots of the characteristic polynomials and we have shown formulas for the convolution of R-sequences and R-Lucas sequences. In the future, we would like to continue working on the cases, when there are one or two common roots.

References

[1] Griffiths, M., Bramham A., The Jacobsthal numbers: Two results and two questions, *The Fibonacci Quarterly* Vol. 53.2 (2015), 147–151.

[2] OEIS Foundation Inc. (2011), The On-Line Encyclopedia of Integer Sequences, `http://oeis.org`.

[3] Mező, I., Several Generating Functions for Second-Order Recurrence Sequences, *Journal of Integer Sequences*, Vol. 12 (2009), Article 09.3.7

[4] Zhang, W., Some Identities Involving the Fibonacci Numbers, *The Fibonacci Quarterly*, Vol. 35.3 (1997), 225–229.

[5] Zhang, Z., Feng, H., Computational Formulas for Convoluted Generalized Fibonacci and Lucas Numbers, *The Fibonacci Quarterly*, Vol. 41.2 (2003), 144–151.

Annales Mathematicae et Informaticae
46 (2016) pp. 217–241
http://ami.ektf.hu

Evolutionary computing based QoS oriented energy efficient VM consolidation scheme for large scale cloud data centers using random work load bench

Perla Ravi Theja*, SK. Khadar Babu

[a]VIT University, Vellore Campus
Vellore - 632 014, Tamilnadu, India
ravithejaperla9048@gmail.com
khadar.babu36@gmail.com

Submitted December 22, 2015 — Accepted June 7, 2016

Abstract

In order to assess the performance of an approach, it is unavoidable to inspect the performance with distinct datasets with diverse characteristics. In this paper we had assessed the system performance with random workbench datasets. A-GA (Adaptive Genetic Algorithm) based consolidation technique has been compared with other consolidation techniques including dynamic CPU utilization techniques, VM (Virtual Machine) selection and placement policies. The proposed consolidation system had exhibited better results in terms of energy conservation, minimal Service Level Agreement (SLA) violation and Quality of Service (QoS) assurance.

MSC: 68U20; 68U99; 65K10; 65K05

1. Introduction

As computing requirements is changing in a huge rate in all-most all the bossiness's, there requires optimal utilization of resources which yields in saving costs, which motivated everyone to make use of Cloud computing. The key issues with

*Corresponding author

Cloud computing are energy consumption, SLA violation, Cloud Network Security and Privacy. As per report from Natural Resources Defence Council (NRDC), in 2013 data centers in USA used around 91 billion kilowatt-hours of electricity that is matching to the annual output of 34 large (500-megawatt) coal-fired power plants. By 2020 this demand may increase to roughly 140 billion kilowatt-hours that is matching to annual output of 50 power plants, costing American businesses 13 billion USD annually in electricity bills and emitting nearly 100 million metric tons of carbon pollution per year [21]. In this paper we are concentrating on energy conservation, reducing SLA violation and providing QoS assurance by implementing dynamic thresholding based resource utilization detection approach for real time cloud systems and by developing an evolutionary computing paradigm based A-GA for VM Placement Policy in VM Consolidation application. In this paper we exhibited performance analysis of the proposed system with other approaches having diverse CPU utilization approaches i.e. (Static Threshold (THR), Local Regression (LR), Robust Local Regression (LRR), Median Absolute Deviation (MAD), and Inter Quartile Range (IQR) etc.), VM selection policies i.e. (Minimum Migration Time (MMT) Maximum Correlation (MC) and Random Selection (RS)) and VM placement policies i.e. (Power Aware Best Fit Decreasing (PABFD) and Ant Colony Optimization (ACO)).

2. Related work

Lot of researches have made research on VM consolidation approaches [11, 16, 17, 23] for QoS and energy optimization. Coming to overload detection in data centers, researchers primarily made use of the static threshold approach [3], where they used overall CPU utilization and defined static thresholds. As dynamic prediction is better than Static prediction researches directed towards dynamic cloud pattern and workloads [4]. On the basis of the historical pattern of CPU utilization, Buyya et al [4] came up with a dynamic threshold scheme for upper and lower bound estimation. Regression based CPU utilization was suggested in [9, 10], where the CPU utilization was estimated at host nodes. Various researches used linear regression and the K-nearest neighbour regression schemes for approximating the data retrieved throughout the VMs lifetime. They gave importance for SLA optimization. Later, bin packing problem was considered in [1, 24, 25] for efficient VM consolidation. All most all existing approaches have used PABFD based consolidation [23]. Lot of algorithms tried to maximize number of host's shutdown by providing server consolidation. Multiple researches tried to provide dynamic schemes for optimal consolidation [6, 7, 13].

Various researches tends towards green cloud computing using bio-inspired technique [2]. Resource provision through ant colony model was inspected in [14]. For the purpose of VM placement, ACO and GA were suggested in [5] and [8] respectively. Disintegrated ACO placement scheme was put forwarded in [12]. Resource placement module in virtualization using GA was inspected in [15, 22, 26]. Great number of current researchers had concentrated either on VM selection or VM

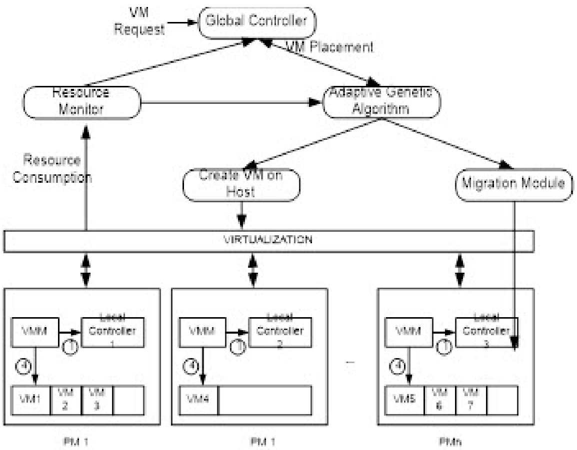


Figure 1: Proposed system architecture

placement. Real work load based A-GA VM placement schemes are used in various researches [18, 19, 20].

3. Our contribution

In our earlier work we validated the optimal efficiency of our proposed A-GA based consolidation with MMT selection policy and combined IQR and LRR CPU prediction schemes by simulation using real world work load from more than a thousand planet-lab virtual machines [18, 19, 20]. In this paper, we performed comparative analysis between different combinations of host overload detection algorithms, VM selection policies and VM placement policies for random workload traces.

3.1. Cloud model

As a part of large scale cloud infrastructure we considered varied physical machines (PMs) or hosts, which are attributed in terms of its CPU consumption and Millions Instructions per Second (MIPS). Storage will be allotted using Storage area network which supports VM migration. Here VMs seek MIPS which will be allotted to carry work.

The VM consolidation minimizing number of active hosts by allotting VM's to run on each single PM so as to conserve energy and to facilitate optimal resource consumption. This process involves SLA Violation and downtime which need to be minimized by applying optimal algorithms. Proposed VM Consolidation scheme as shown in Figure 1 involves two controllers i.e. Global and Local Controllers. Local Controller (LC) which periodically check the status of PM whether it is Overloaded/Underloaded. If it is Overloaded it will report to Virtual Machine

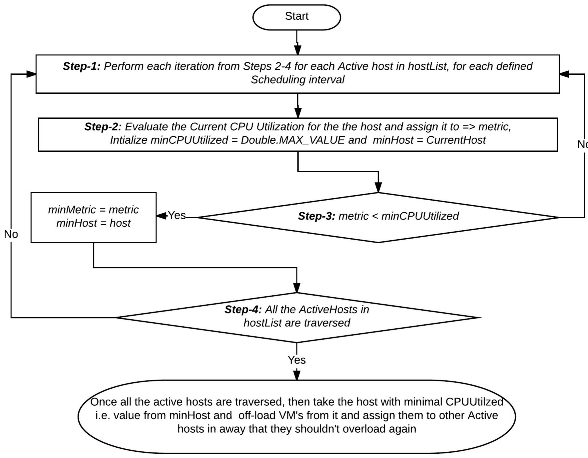


Figure 2: Data Flow diagram for Host under load detection algorithm

Manager (VMM), which in turn triggers VM selection policy and decides which VMs that need to be offloaded to avoid SLA performance degradation and reports to Global Controller (GC) which places these VMs onto other active hosts in a way that they shouldn't overload again using VM placement polices. If it is underloaded then in same way VMM triggers VM selection policy and GC uses VM placement policy to place these VM's.

3.1.1. Host under load and overload detection

The following data flow diagram i.e. Figure 2, illustrates, how the Host under detection algorithm is carried-out. The outcome of this algorithm is minimum CPU utilized host, which will be slept (turned-off) resulting in energy conservation. To investigate whether the host is overloaded or not, we make use of host overload detection algorithms. Here we considered various algorithms i.e. (Median Absolute Deviation (MAD), Static Threshold (THR), Inter Quartile Range (IQR), Local Regression (LR) and Local Robust Regression (LRR)) for this purpose. The process-flow of each algorithm is illustrated completely in terms of data flow diagrams in Figures 3 to 7 respectively as mentioned below:

To detect the overloaded host node, each host node initializes an overload detection scheme intermittently to perform VM de-consolidation, thus enabling SLA violation avoidance. CPU utilization of the host node has been used for identifying overloaded hosts. Unlike conventional schemes based on static threshold, in this paper we have developed a dynamic threshold based adaptive CPU utilization and overload detection scheme. It enables the proposed system to behave in real time scenario where there is highly fluctuating resource utilization. It adjusts resource

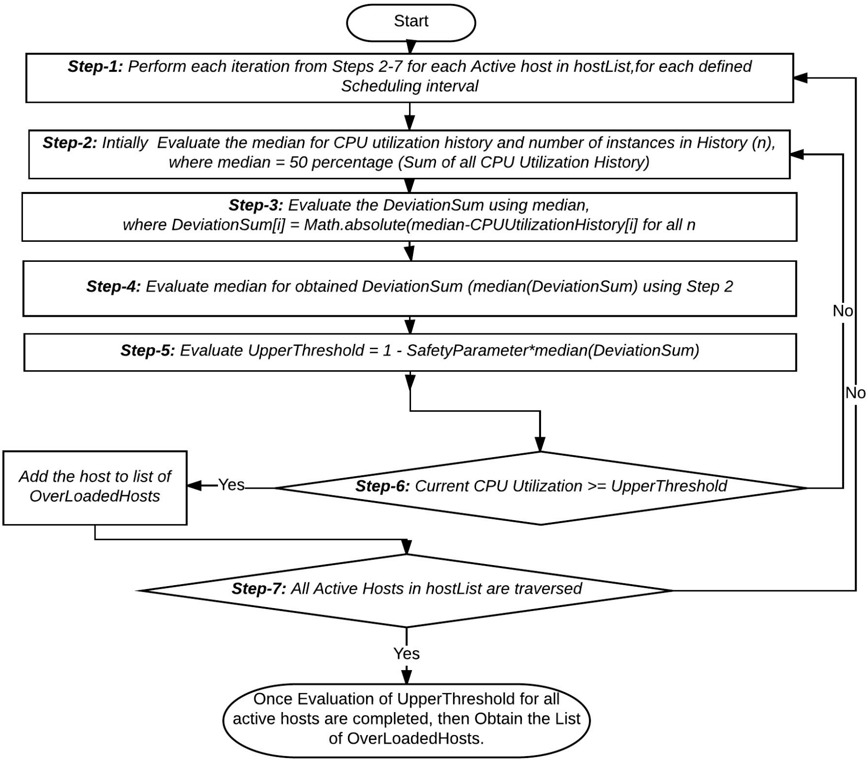


Figure 3: Data Flow diagram for Median Absolute Deviation (MAD) algorithm

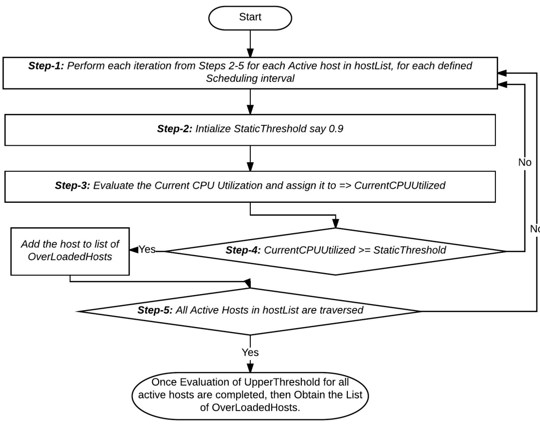


Figure 4: Data Flow diagram for Static Threshold (THR) algorithm

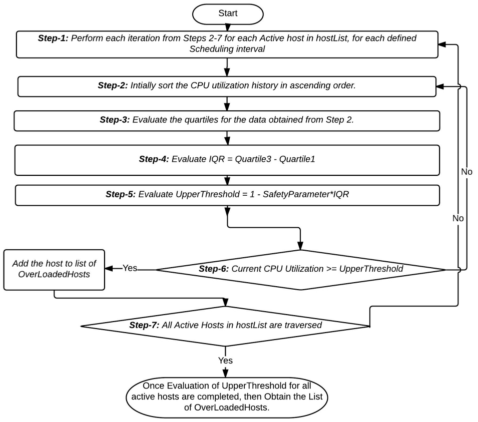


Figure 5: Data Flow diagram for Inter Quartile Range (IQR) algorithm

utilization threshold based on the variation in CPU utilization and utility map. It can be observed that higher deviation might even result into 100% CPU utilization that signifies higher overloading probability. To enable dynamic threshold detection scheme, we have used Interquartile Range (IQR) and robust local regression (LRR) algorithm. In our proposed model, the resource utilization has been examined at the interval of 5 minute and for each odd iteration, IQR algorithm has been used, while LRR has been scheduled for even iterations. Such novelty intends to employ major advantages of both approaches. A brief of IQR is given as follows:
IQR is a statistical dispersion technique which is equivalent to the differences between the third and first quartile. To estimate dynamic CPU utilization threshold, the following equation has been used.

$$IQR = Q_3 - Q_1, T_m = 1 - s.IQR \tag{3.1}$$

Where s represents the safety parameter that states the maximum extent of the tolerability of a host node in Cloud environment and its lower value signifies the higher tolerance to the fluctuation in the CPU utilization. Here we have used $s = 1.2$, and it can be changed to examine the optimal performance. In addition to IQR, we have used Loess concept [18] to derive the LRR algorithm, which has been employed for fitting a trend polynomial to the earlier k observations for the CPU utilization called utilization map. For initial recent observations, it is retrieved as

$$\hat{g}(x) = (\hat{a}) + \hat{b}x \tag{3.2}$$

It is further used for calculating the next observation $\hat{g}(x_{k+1})$. To offload some VMs from an overloaded host node, the following conditions have been fulfilled:

$$s.\hat{g}(x_{k+1}) \geq 1, \quad x_{k+1} - x_k \leq t_m \tag{3.3}$$

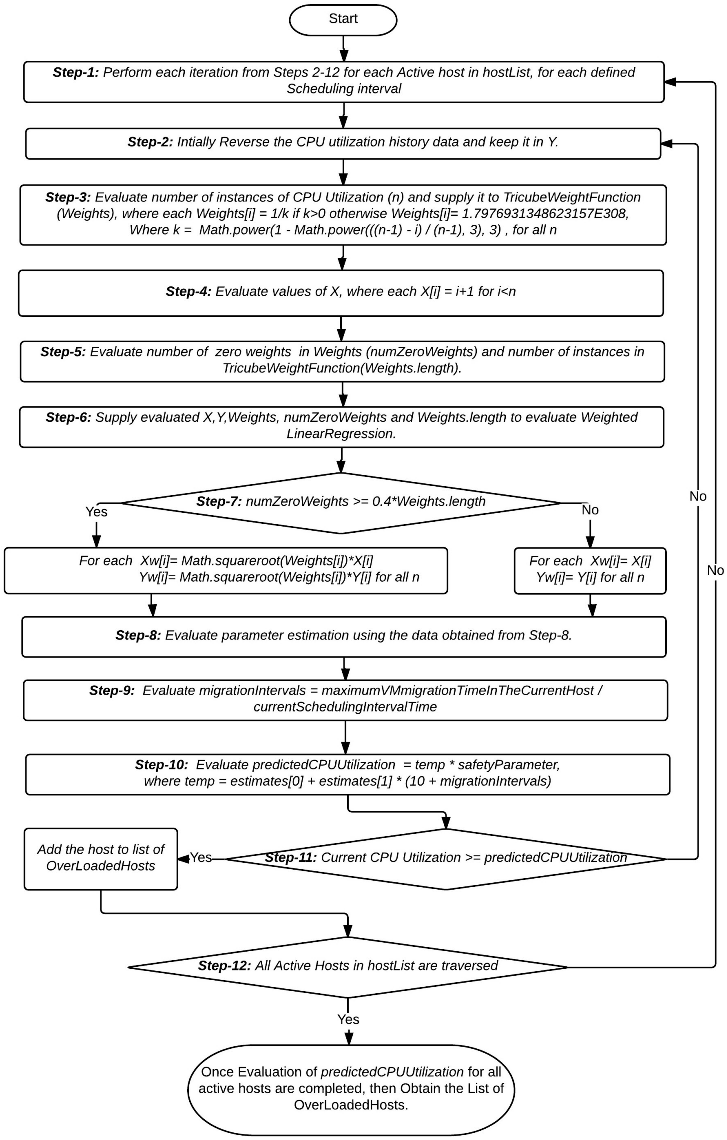


Figure 6: Data Flow diagram for Local Regression (LR) algorithm

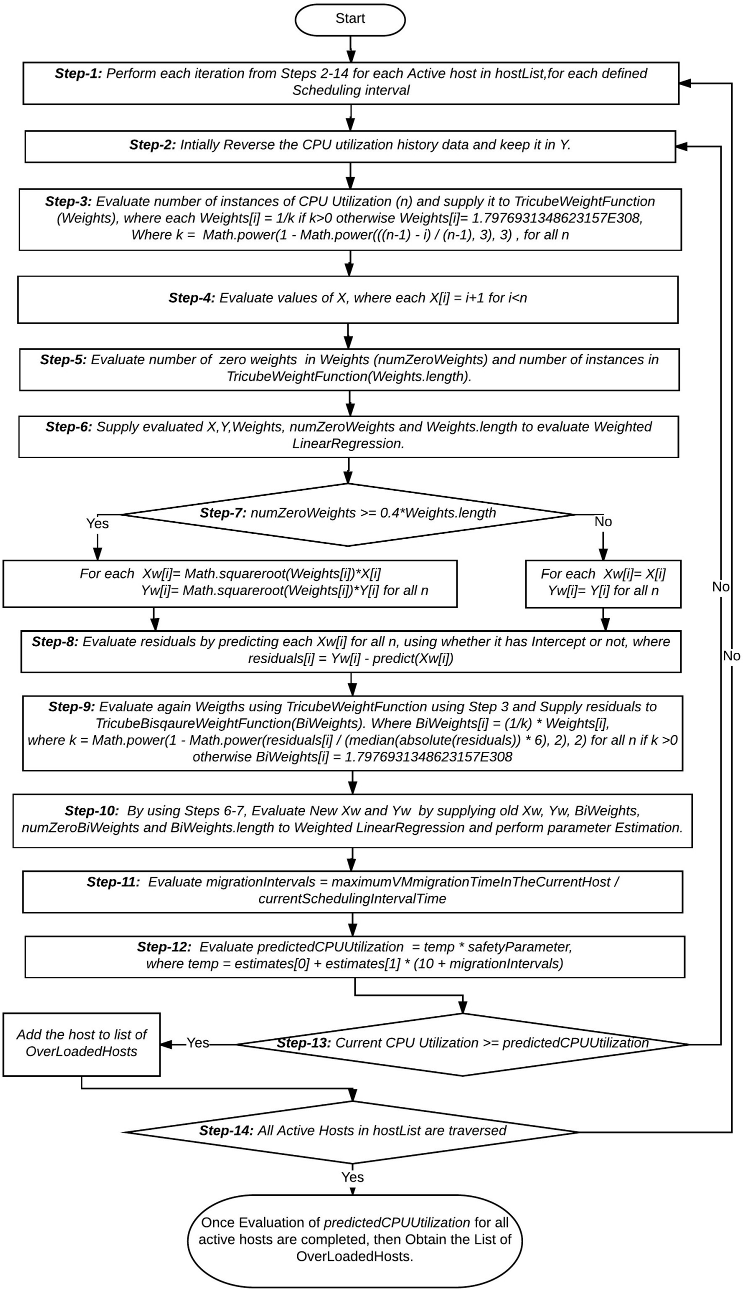


Figure 7: Data Flow diagram for Local Robust Regression (LRR) algorithm

Where $s \in \mathbb{R}^+$ is the safety parameter that signifies the maximum capability or tolerability of a host node and t_m represents the maximum time required for migrating a VM from the overloaded host. The traditional Loess concept [18] is found vulnerable to the outliers caused due to leptokurtic or heavy-tailed distributions. To alleviate this problem, we have modified the Loess concept [18] to bisquare from the conventional least-squares (LR) approach. Thus we have LRR approach iteratively for estimating the initial fitting and tricube weight function has been used for dynamic weight estimation. The fitting parameter has been retrieved at x_i to get optimal value by means of $\hat{y_i}$. Thus, the final residual value is $\epsilon_i = y_i - \hat{y_i}$. Furthermore, the final retrieved value (x_i, y_i) has been assigned a robustness factor $\mathbb{R}_i$ that primarily depends on the magnitude of ϵ_i. Mathematically, R_i is obtained as

$$\mathbb{R}_i = \mathbb{B}\left(\frac{\hat{\epsilon}_i}{6L_{MAD}}\right) \tag{3.4}$$

Where $\mathbb{B}(.)$ gives the bisquare weight function and L_{MAD} represents the Median Absolute Deviation (MAD) for the least square fit. Mathematically:

$$B(.) = \begin{cases} (1-u^2)^2 & \text{if} |u| < 1, \\ 0 & \text{otherwise.} \end{cases}$$

In this paper, we assigned $\mathbb{R}_i$ for each observation (5 minute), where L_{MAD} has been obtained as:

$$L_{MAD} = median|\hat{\epsilon}_i| \tag{3.5}$$

Using equation (3.3), the next observation has been obtained for the estimated trend line, where observing any inequalities, the host can be identified as overloaded.

3.1.2. VM selection policy

In this phase, the VM selection takes place where it is intended to select the VM which should be migrated to minimize overhead from overloaded host. Estimating the dynamic CPU utilization threshold, VM selection has been performed that offloads host for avoiding SLA violation and unwanted energy consumption caused due to overload. In this paper, we have examined three different selection policies; the Minimum Migration Time (MMT), Maximum Correlation (MC) and Random Selection (RS) policy. A brief discussion of the implemented VM selection policies with respective data flow diagrams are reprsented in Figures 8 to 10 as mentioned below:

1. **Minimum Migration Time(MMT) Policy**
   Once assessing the host's CPU utilization levels and identifying any probable or overloaded host, VMs selection algorithm performs offloading of that host node to avoid any probability of SLA violation. The developed MMT selection policy performs migration of only those VMs (v) which requires minimal migration time than the other. In this paper, the migration time has been

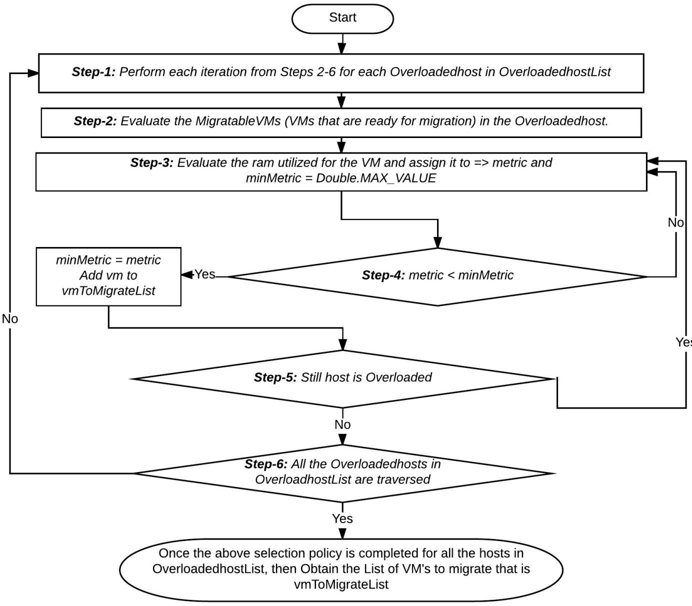


Figure 8: Data Flow diagram for Minimum Migration Time (MMT) policy

estimated in terms of the resource, RAM being used by VM divided by the supplementary network bandwidth available for hostj. Let $V_(j)$ be a set of VMs connected with the hostj. Therefore, the VM to be migrated is selected based on the following condition.

$$v \in V_j | \forall_a \in V_j, \frac{RAM_u(v)}{NET_j} \leq \frac{RAM_u(a)}{NET_j} \tag{3.6}$$

Where $RAM_u(a)$ depicts the amount of RAM currently being used by VMa; and NET_j refers the bandwidth available for migration from the host$j$.

2. **Maximum Correlation (MC) Policy**
In case of maximum correlation [19][20] policy, it is assumed that the higher correlation between the CPU utilization by VMs connected to certain host signifies higher overloading probability. In MC policy, VMs having highest correlation for the CPU utilization are needed to be migrated from the current host to assist energy conservation and SLA violation avoidance. We have used the concept of multiple correlation coefficients (MCC) for estimating the intra-VM correlation and respective CPU utilization. The used MCC coefficients are in relation to the squared correlation between the real and the predicted values of the dependent variable. In fact, it can be interpreted as the fraction of variance of the dependent variable elucidated by the associated

independent variables.
Let $X_1, X_2, \cdots, X_n$ be the CPU utilization of n VMs connected to a host and Y be the specific VM to be migrated, where $(n-1)$ are the random independent variables. Here Y being the VM to be migrated, be the dependent variable. In our proposed MC policy, we have calculated the correlation between Y and $(n-1)$. The obtained augmented matrix $(n-1)\times n$ encompasses the observed instances or the values of $(n-1)$, indicated as X. Similarly, the observation vector or mapped vector $(n-1)\times 1$ of the dependent variable Y be y. Thus, the overall observation vector or the mapping vectors for the VM's CPU utilization can be given as follows:

$$x = \begin{bmatrix} 1 & x_{1,1} & \cdots & x_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n-1,1} & \cdots & x_{n-1,n-1} \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} \tag{3.7}$$

Observing equation (3.7), it can be found that the first column of X is all 1 (for all instances), therefore it can be considered as an augmented matrix. The predicted values of the VM's CPU utilization or Y can be presented as $\hat{Y}$, which can be obtained by $\hat{Y} = Xb$, where $b = (X^TX)^{-1}X^TY$. Obtaining the predicted values, MCC coefficients, the final correlation $R^2_{Y,1,\cdots,n-1}$ has been obtained as

$$R^2_{Y,X_1,\cdots,X_{n-1}} = \frac{\sum_{i=1}^{n}(Y_i - m_Y)^2(\hat{Y}_i - m_{\hat{Y}})^2}{\sum_{i=1}^{n}(Y_i - m_Y)^2 \sum_{i=1}^{n}(\hat{Y}_i - m_{\hat{Y}})^{2'}} \tag{3.8}$$

Where m_Y and m_Y give the observation means of Y and $\hat{Y}$ respectively. In MC based VM selection, the MCCs for all mapped instances X_i have been obtained as $R^2_{X_i,X_1,\cdots,X_{i-1},X_{i+1},...,X_n}$. Finally, based on correlation value, Eq. (3.9) has been used to select a specific VM to be migrated:

$$\begin{aligned} v \in VM_j|\forall_a & \\ \in V_j, R^2_{X_{vm},X_1,\cdots,X_{vm-1},X_{v+1}X_n} & \qquad (3.9)\\ \geq R^2_{X_v,X_1,\cdots,X_{a-1},X_{a+1},\cdots,X_n}. & \qquad (3.10) \end{aligned}$$

In addition to the MMT and MC selection policy, we have also examined an algorithm called random selection (RS) policy.

3. **Random Selection Policy**
In Random Selection policy, a VM is randomly selected for migration from the host node as per a uniformly distributed discrete random variable $A = U(0|V_j)$, whose values signify a set of VMs, V_j placed at jth host. Since VM consolidation is a bin packing problem and therefore, an optimal approach for placement is of great significance to ensure minimal downtime, energy consumption and probable SLA violation. The following section discusses the proposed evolutionary computing based VM placement approach.

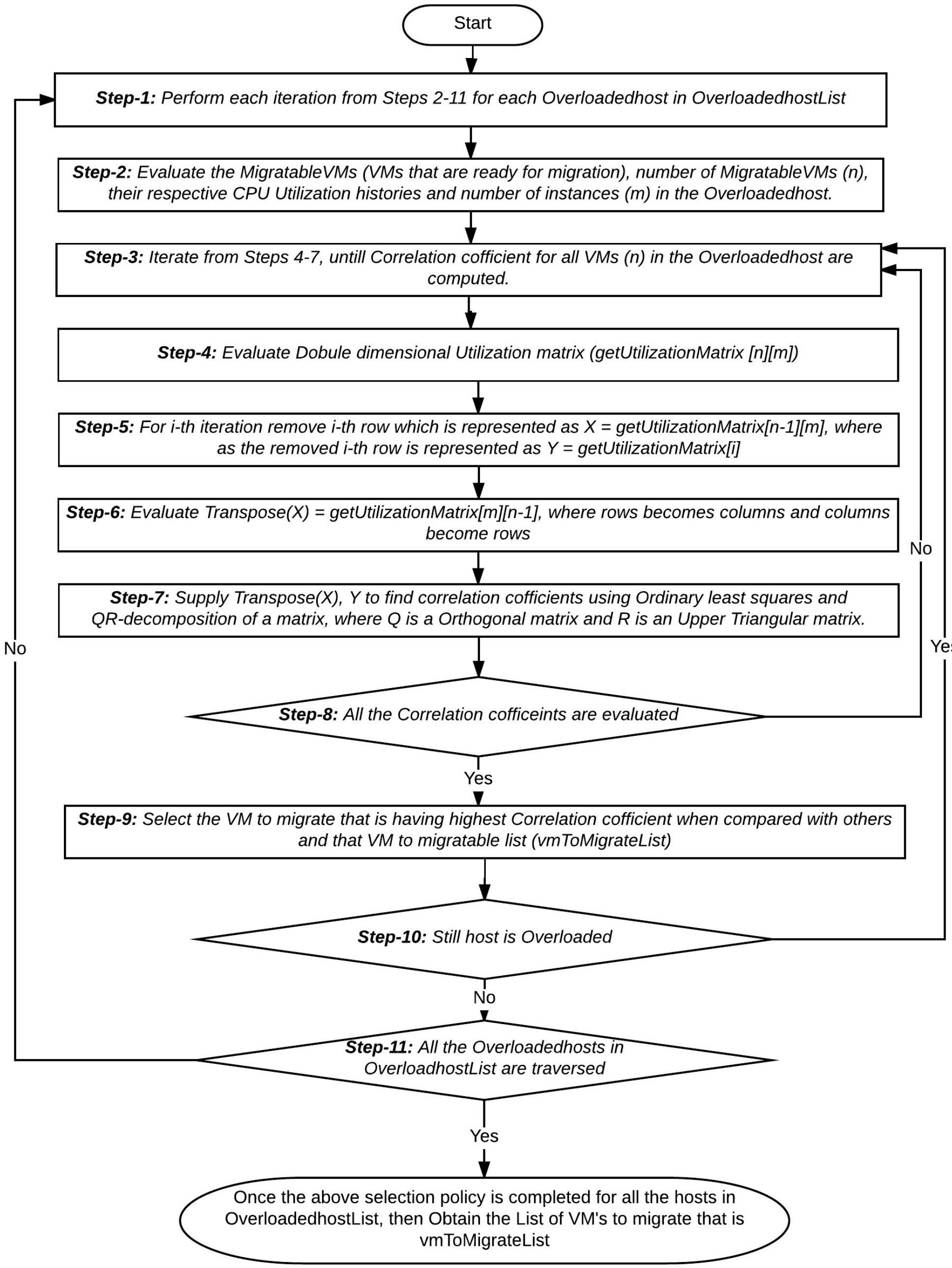


Figure 9: Data Flow diagram for Maximum Correlation (MC) policy

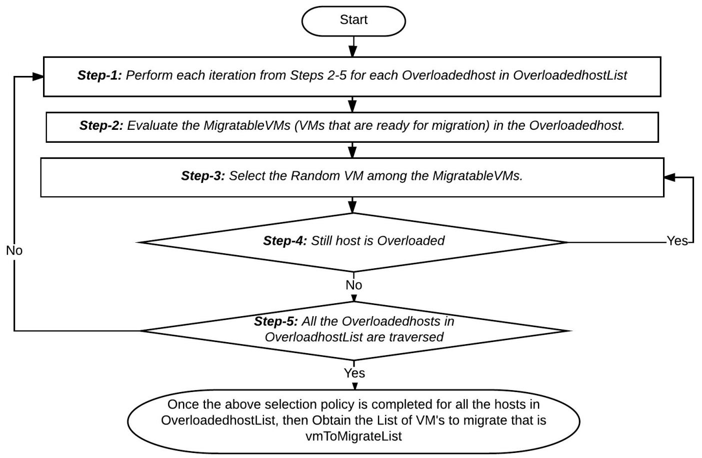


Figure 10: Data Flow diagram for Random Selection (RS) policy

3.1.3. VM placement policy

VM placement can be stated to be a problem of bin packing that encompasses bins, items and prices as the three parameters, where bins represent the host nodes, VMs represent the items to be allocated, bin size refers the available resource on the host node; and the resource or CPU power consumed by a host is stated in terms of price. In bin packing, it is intended to accommodate as much as VMs that makes the overall scenario NP-hard. In order to deal with such non-convexity problem, certain heuristic approach or evolutionary computing scheme can be the potential solution. In this paper, we have proposed Adaptive Genetic Algorithm (A-GA) as VM placement policy. We have interfaced A-GA with CloudSim simulator comprising multiple host nodes, and VMs in the data center. In our proposed model, each host is equipped with one or multiple processing elements (PE). The executing VMs on the hosts have one or multiple running Cloudlets. In simulation model, the user requests have been stated in terms of Cloudlets, where the needed processing power for each Cloudlet has been defined in terms of Million Instructions Per Second (MIPS). In the proposed placement policy, the scheduler considers all hosts, VMs and VM maps as input and generates mapping for nodes, where it divides overall MIPS into different components like hosts and VMs running in parallel. The functional discussion of the proposed A-GA scheme is given in (Figure 13).

1. **A-GA Based VM Placement**
   As depicted in (Figure 1), the proposed A-GA processes VM scheduling based on the resource utilization information provided by local controller (LC) and the upper threshold value estimated by dynamic threshold estimation scheme. We have considered upper threshold so as to satisfy transient variations of resource demand by different VMs on a host node. The CPU utilization pattern

or history of VMs has been considered for VMs placement onto destination host node. Considering a large scale cloud infrastructure or data center, the time efficient and effective placement scheduling is of great significance. In this research we have intended to reduce the number of VM migration and migration time, so as to enable energy efficient and QoS oriented consolidation mechanism. In addition, we have scheduled the system to enable maximum host shut down so as to conserve energy. In this paper, placing or allocating VM on certain host, our algorithm estimates the energy of the data center and accordingly performs further scheduling to minimize energy consumption.

At first, the proposed A-GA algorithm initializes a definite set of population where individual host is a tree comprising global controller as its root, hosts are the next level nodes and VMs are the child nodes (Figure 1). It calculates the total energy consumption and CPU utilization for each mapping in the deployed cloud center. Here, the VM mapping history also known as utilization pattern, allocated VMs and their resource utilization mapping, future mapping for VMs based dynamic utilization, hosts and its available resource availability etc. have been used as population. The precise discussion of the proposed system is given in (Figure 13). From these chromosomes, our proposed A-GA algorithm initially selects two VM mappings with minimal energy values on which the initial genetic operators (crossover p_c and mutation probability p_m) are applied. Thus, the mapping obtained for VMs onto the host nodes is added to the overall population based on the fitness values. In our proposed A-GA based VM placement policy, p_c selects the host with the best CPU utilization based on the previous VMs mapping. Here p_c and p_m try to reduce host nodes by means of SWITCHING OFF or turning it into SLEEP MODE. Here, it should be noted that unlike conventional genetic approach (i.e., A-GA), we have applied adaptive genetic parameter selection, where these variables are updated dynamically after every iteration, till stopping criteria (100 generations) is obtained.

Consider, the host nodes in data centers be $PM = pm_1, pm_2, pm_3, \cdots, p_{m_m}$ and pm_i be ith host node, where $(1 \leq i \leq m)$. Similarly, VMs in the network be $VM_i = vm_1, vm_2, vm_3, \cdots, vm_{n,i}$, which are connected to the ith host. Consider $vm_{(j,i)}$ be the ith VM on jth PM. The variable $x_{(i,j)}$ signifies whether ith VM is placed on host j or not. Let $P_{(r,i)}$ be the resource capacity r (CPU utilization) on jth host node. The resource needed by ith VM is $v_(r,j)$. Thus, overall load on j th host node would be the sum of all resource needed by all VMs running over it. Consider, T be the duration of past observations, thus the sub-intervals can be obtained by dividing T into $(q-1)$ sub intervals such that $T = [(t_2 - t_1)(t_3 - t_2)(t_q - t_(q-1))]$. The slot $(t_k - t_(k-1))$ is the time periodk. In such manner, for periodk, we have estimated the CPU utilization at a host $(CPU_(i, Util)(k))$ using following equation:

$$CPU_{(i,Util)}(k) = \sum_{j=1}^{n} \frac{vm_{CPU,j}}{pm_(CPU, j)} \tag{3.11}$$

Where, k represents the duration for which the CPU utilization has to be retrieved. Finally, the average CPU utilization at a host node has been obtained as:

$$pm_(i, AvgUtil) = \sum_{t-t_k}^{t_k-n} \frac{pm_(i, Util)(t)}{(q-1)} \tag{3.12}$$

Where $(q-1)$ represents the total number of sub intervals in T time. Consider pm_i represents the power of jth host node during t_k. Thus, the power utilization can be obtained in terms of CPU utilization at the host node. Consider, $pm_iE^{(k)}$ be the power or energy consumption of the jth host node in between the last time interval and the current time, then it can be obtained as

$$pm_iE^{(k)} = pm_{iw}(k-1) + (pm_{iw}(k-1) + (pm_{iw}(k))(t_k - t_k - 1) \tag{3.13}$$

The energy consumption for the jth host, $E(pm_j)$ can be estimated at certain host pm_j having CPU usage as $CPU_{i,Util}(k)$

$$E(pm_j) = K_j.e_j^{max} + (1-k_j).e_j^{max}.CPU_{i,Util}(k) \tag{3.14}$$

Where K_j states the part of energy consumed when the host pm_j is in idle state; e_j^{max} states for the energy consumption of host pm_j when it being used 100%. The variable $CPU_{(i,Util)}(k)$ represents the CPU utilization by host pm_j. We have used this approach to estimate the energy consumption at certain host so as to perform placement scheduling. Similarly, the energy consumption for all host nodes, $D_E(k)$ can be obtained for a period using following equation:

$$D_E(k) = \sum_{i=1}^{m} pm_iE^{(k)} \tag{3.15}$$

In this paper, the prime objective of the proposed A-GA scheme for VM placement is to retrieve the set of mapping from VM set to the host set PM while ensuring minimal energy consumption $D_E(k)$, provided:

$$\forall_i \sum_{j=1}^{m} x_{ij-1} \tag{3.16}$$

$$\forall_j \sum_{i=1}^{n} vm_{CPU,i}X_{ij} \leq pm_{CPU,j}. \tag{3.17}$$

In this paper A-GA is compared with ACO and PABFD VM placement algorithms. The implementation of our proposed A-GA based consolidation scheme using LRR and IQR for host overload detection and MMT for VM Selection [20] is shown in Figure 13. The whole process flow for ACO and PABFD is shown in Figure 11 and Figure 12 respectively.

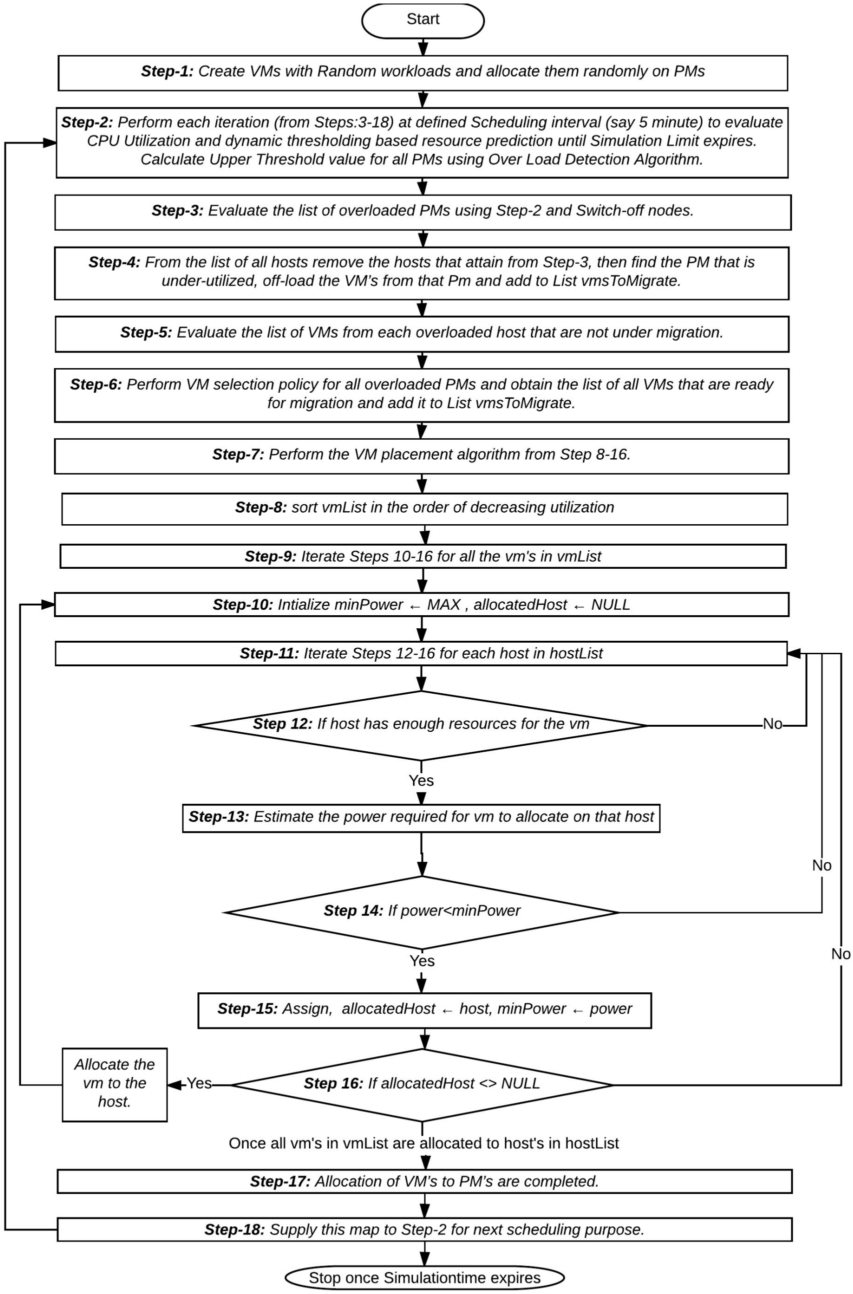


Figure 11: Data Flow diagram for Power-Aware Best Fit Decreasing (PABFD) based VM Consolidation Scheme

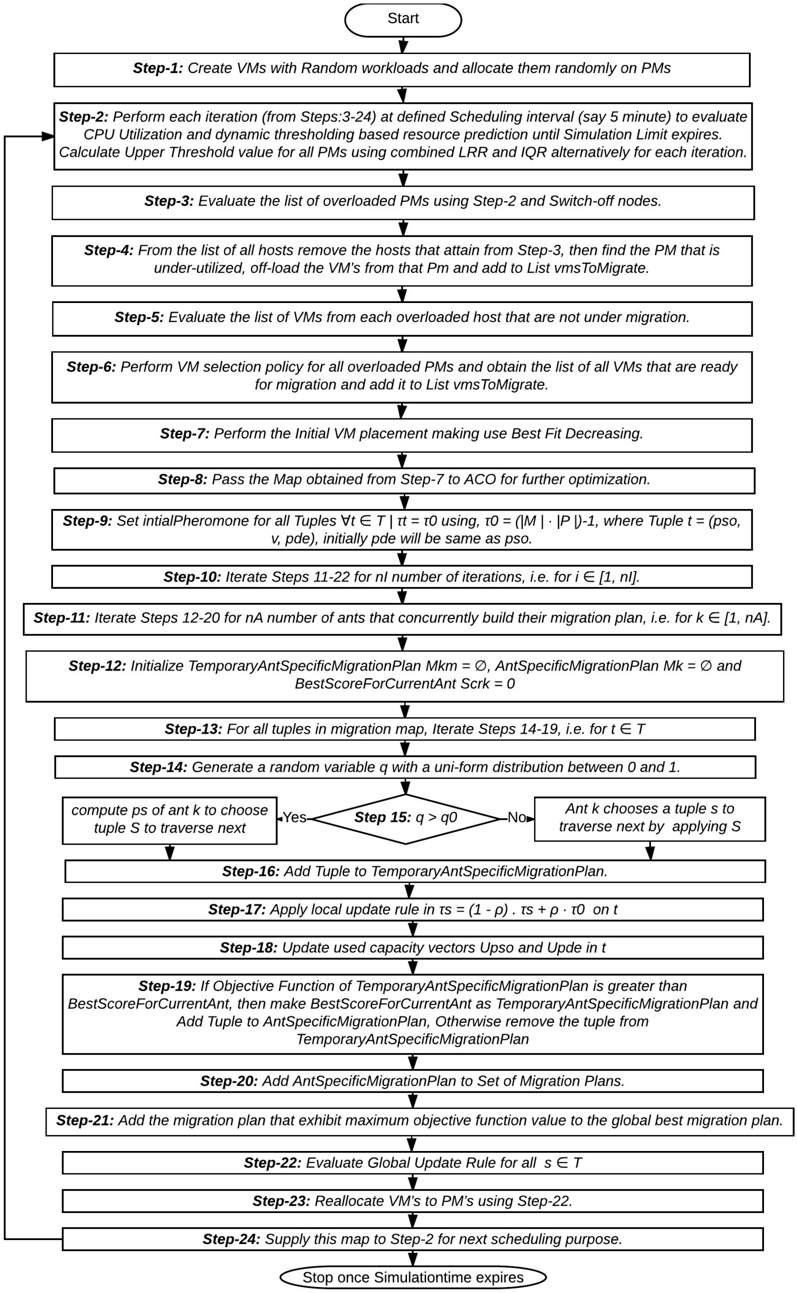


Figure 12: Data Flow diagram for Ant Colony Optimization (ACO) based VM Consolidation Scheme

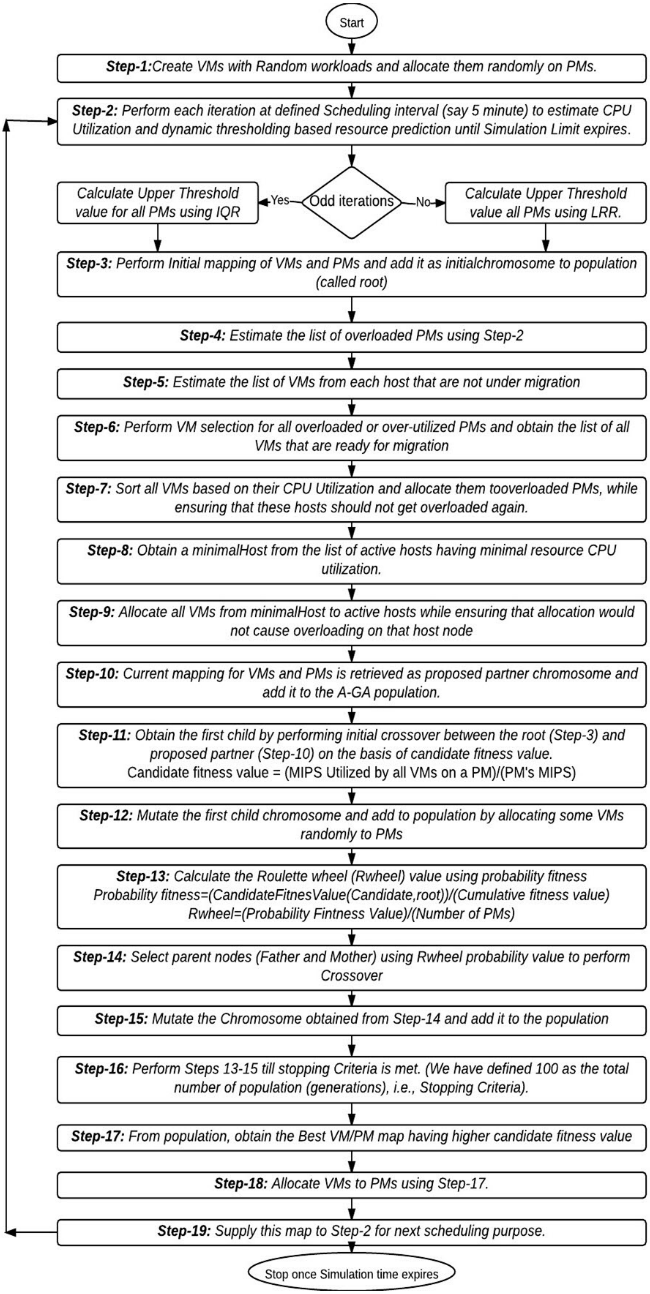


Figure 13: Data flow diagram for proposed Adaptive Genetic Algorithm (A-GA) based VM Consolidation Scheme

Workbench Data	**Number of host Nodes**	**Number of VM's**
Random Workload	300	400, 500, 600
Random Workload	600	800,1000,1200
Random Workload	800	1000,1200,1600

Table 1: Workbench data with different number of host nodes and VM's

4. Experimental setup

In this paper, we have examined the performance of the proposed system using random cloud workload traces. To evaluate the robustness of the proposed consolidation scheme a large scale cloud infrastructure with different network size and configurations has been considered as per Table 1. For simulation we utilized CloudSim in Eclipse IDE. Intially, servers frequency is mapped to MIPS ratings with 1860 and 2660 MIPS in HP ProLiant ML110 G4 and HP ProLiant ML110 G5, respectively. Each server has been assigned 1 GB/s network bandwidth.

5. Results and discussion

The different simulation scenarios used for evaluation are mentioned in Table 2. With different network and process algorithm combination, the simulations have been done and respective performance has been analyzed in terms of the following performance parameters:

1. Number of VM Migrations (in Nos.)
2. Service Level Agreement Violation (SLAV) (%)
3. SLA time per Active Host (%)
4. SLA Performance Degradation (%)
5. Number of Host Shutdown (Nos)
6. Energy Consumption (KWh)

In this paper, we plotted graphs only for random workload with 800 hosts and VMs with 1000, 1200, 1600 (as similar results are observed for 300 and 600 hosts). As shown in Figure 14, the proposed A-GA based VM consolation yields minimal VM migration thus providing minimal downtime. Figure 15 illustrates SLAV, where the proposed A-GA provides minimal SLA violation with MMT VM selection policy.

Figure 16 represents SLA performance degradation and it clearly states proposed evolutionary A-GA with IQR and LRR as overload prediction schemes and MMT

CPU Utilization Threshold	**VM Selection**			**VM Placement**
IQR	MMT	MC	RS	Power Aware Best Fit Decreasing
LR	MMT	MC	RS	Power Aware Best Fit Decreasing
MAD	MMT	MC	RS	Power Aware Best Fit Decreasing
THR	MMT	MC	RS	Power Aware Best Fit Decreasing
LRR	MMT	MC	RS	Power Aware Best Fit Decreasing
Combined IQR and LRR	MMT	MC	RS	Ant Colony Optimization
Combined IQR and LRR	MMT	MC	RS	Adaptive Genetic Algorithm

Table 2: Implementation and simulation scenarios

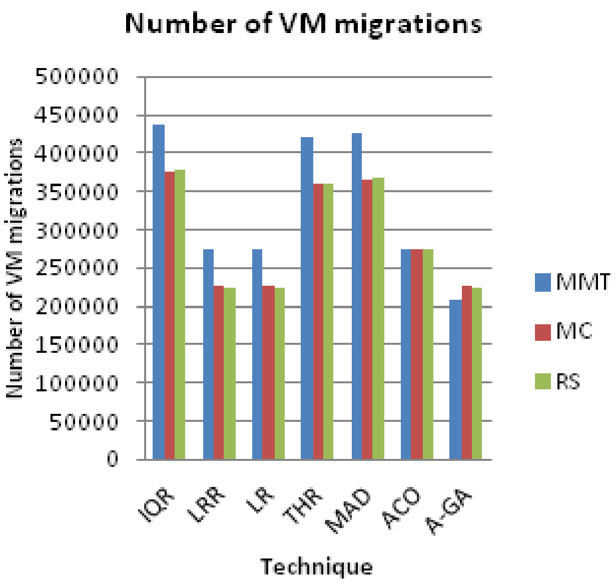


Figure 14: Number of VM migrations for 800 Hosts

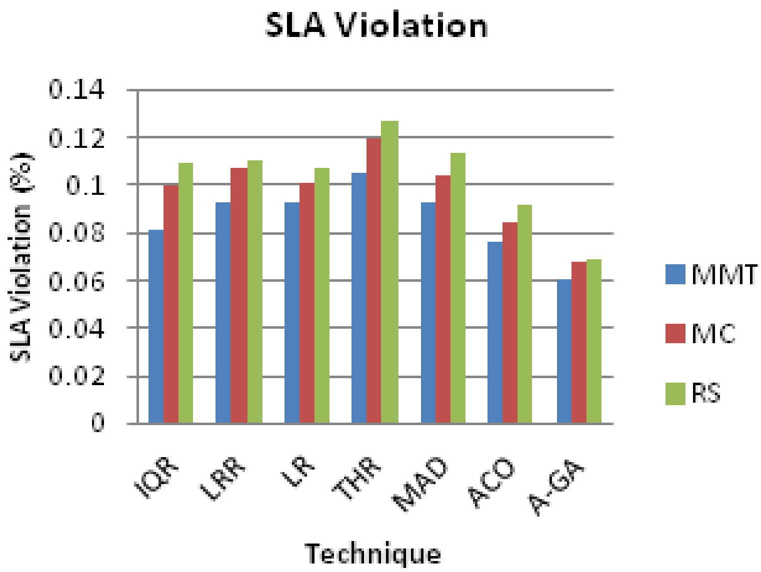


Figure 15: SLA violation for 800 Hosts

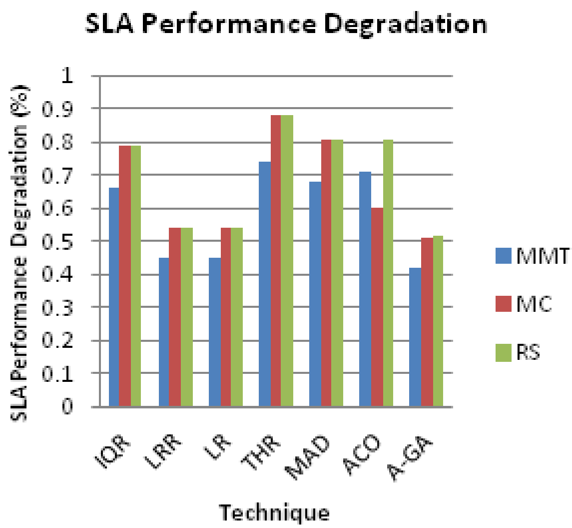


Figure 16: SLA performance degradation for 800 Hosts

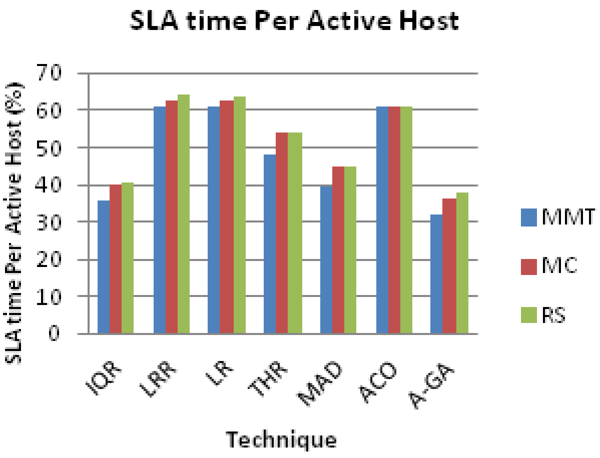


Figure 17: SLA time per active host for 800 Hosts

as VM selection policy exhibits minimal performance degradation. Figure 17, also reflects that the proposed system performed better in terms of SLA time per active host. Figure 18 affirms that maximum hosts shutdown happened with proposed system. The complete analysis states that proposed system provides complete QoS assurance with reliability.

Figure 19 clearly shows the proposed system with IQR and LRR + MMT + A-GA ensure minimal energy is consumed thus drive towards green cloud computing.From the plotted graphs it is clearly found that the proposed A-GA based scheme can provide better results as compared to the other heuristic techniques such PABFD and ACO. Taking into account among selection policies, it has been observed that MMT policy outperforms other approaches such as MC and RS.

6. Conclusion

In this paper we presented highly optimal random workbench based performance evaluation using adaptive genetic algorithm, where different complexities are considered. The execution of merging both IQR and LRR for dynamic CPU prediction schemes yielded optimal results compared with other schemes. In terms of all other paramaters MMT in combing with A-GA outperforms other approches such as MC and RS. A-GA outperforms ACO and PABFD based VM consolidation. As a reult we can quote that the proposed system can be utilized in large scale cloud data centers using random work load for green cloud computing and QoS assurance.

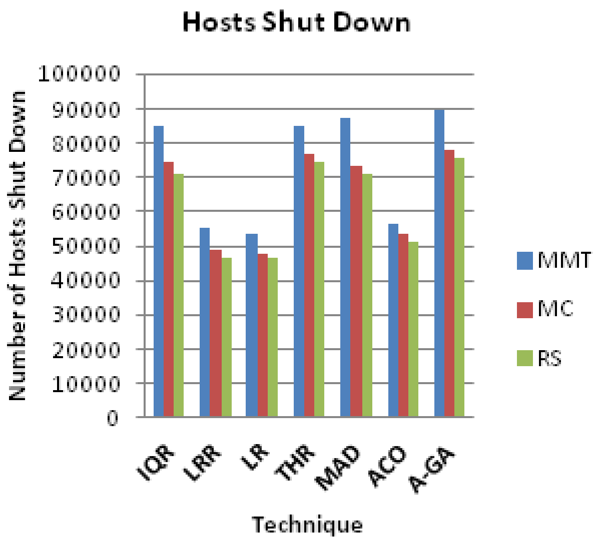


Figure 18: Number of hosts shutdown for 800 Hosts

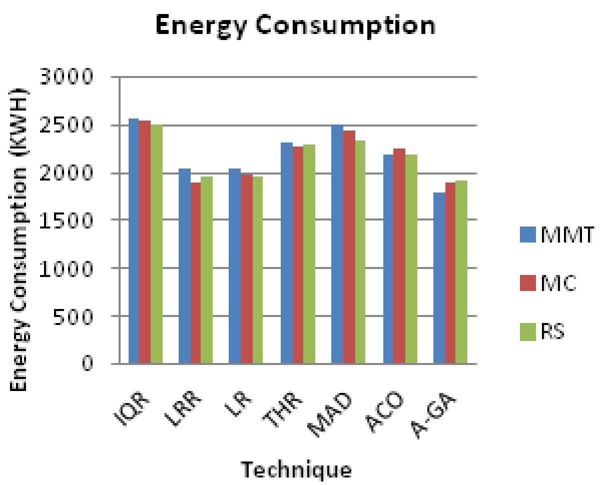


Figure 19: Energy consumption for 800 Hosts

References

[1] Y. Ajiro, A. Tanaka. Improving packing algorithms for server consolidation. - Proceedings of the International Conference for the Computer Measurement Group, San Diego, California, USA, 2007, pp. 399–407.

[2] D. Barbagallo, E. Di Nitto et al. A bio-inspired algorithm for energy optimization in a self-organizing data center. - Self-Organizing Architectures, Springer, 2010, pp. 127–151.

[3] Beloglazov, J. Abawajy et al. Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing. - Grid Computing and eScience, Future Generation Computer Systems, Vol 28, 2012, pp. 755–768.

[4] Beloglazov, R. Buyya. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in Cloud data centers. - Concurrency and Computation: Practice and Experience, Vol 24, 2012, No 13, pp. 1397–1420.

[5] Y.S. Dong,G.C.Xu et al. A distributed parallel genetic algorithm of placement strategy for virtual machines deployment on Cloud platform. -The Scientific World Journal, 2014, pp. 1–12.

[6] M. Dorigo, L. Gambardella. Ant colony system: a cooperative learning approach to the traveling salesman problem. - IEEE Transactions on Evolutionary Computation, Vol 1, 1997, No 1, pp. 53–66.

[7] M. Dorigo, G. Di Caro et al. Ant algorithms for discrete optimization. - Artificial Life, Vol 5, 1999, No 2, pp. 137–172.

[8] A. Esnault, E. Feller et al. Energy-aware distributed ant colony based virtual machine consolidation in IaaS Clouds bibliographic study. - Informatics Mathematics (INRIA), 2012, pp. 1–13.

[9] F. Farahnakian, P. Liljeberg et al. Linear regression based CPU usage prediction algorithm for live migration of virtual machines in data centers. - Software Engineering and Advanced Applications, 39th Euromicro Conference, Santander, Spain, 2013, pp. 357–364.

[10] F. Farahnakian, T. Pahikkala et al. Energy aware consolidation algorithm based on K-nearest neighbor regression for Cloud data centers. - 6th IEEE/ACM International Conference on Utility and Cloud Computing, Dresden, Germany, 2013.

[11] E. Feller, C. Morin et al. A case for fully decentralized dynamic VM consolidation in Clouds. -Cloud Computing Technology and Science, 4th IEEE International Conference, Taipei, Taiwan, 2012, pp. 26–33.

[12] M.H. Ferdaus, M. Murshed et al. Virtual machine consolidation in Cloud data centers using ACO metaheuristic. - Euro-Par 2014 Parallel Processing, 20th International conference, Porto, Portugal, 2014, pp. 306–317.

[13] M. Harman, K. Lakhotia et al. Cloud engineering is search based software engineering too. - Journal of Systems and Software, Vol 86, 2013, No 9, pp. 2225–2241.

[14] Hongwei Chen, Lei Xiong et al. Cloud task scheduling simulation via improved ant colony optimization algorithm. - Journal of Convergence Information Technology, 2013.

[15] B. Madhusudhan, K.C. Sekaran. A genetic algorithm approach for virtual machine placement in Cloud. - Proceedings of International conference on Emerging research in computing, information, communication and applications, 2013, pp. 115–122.

[16] M. Marzolla, O. Babaoglu et al. Server consolidation in Clouds through gossiping. -World of Wireless, Mobile and Multimedia Networks, 12th IEEE International Symposium, Lucca, Italy, 2011, pp. 1–6.

[17] Murtazaev, S. Oh.Sercon: Server consolidation algorithm using live migration of virtual machines for green computing. - IETE Technical Review, Vol 28, 2011, No 3, pp. 212–231.

[18] Perla Ravi Theja, SK. Khadar Babu. An Evolutionary Computing based Energy Efficient VM Consolidation Scheme for Optimal Resource Utilization and QoS Assurance. - Indian Journal of Science and Technology, 77179, Vol 8, 2015, No 26, pp. 1–11.

[19] Perla Ravi Theja, Dr. SK. Khadar Babu. An Adaptive Genetic Algorithm based Robust QoS Oriented Green Computing Scheme for VM Consolidation in Large Scale Cloud Infrastructures. - Indian Journal of Science and Technology, 79175, Vol 8, 2015, No 27, pp. 1–13.

[20] Perla Ravi Theja, SK. Khadar Babu. Evolutionary computing based QoS oriented energy efficient VM consolidation scheme for large scale cloud data centers. - Cybernetics and Information Technologies, Bulgarian Academy of sciences, Vol. 16, No. 2, pp. 97–112.

[21] Pierre Delforge. America's Data Centers Consuming and Wasting Growing Amounts of Energy. -Available at : https://www.nrdc.org/resources/americas-data-centers-consuming-and-wasting-growing-amounts-energy , February,06, 2015.

[22] M. Tang, S. Pan, A hybrid genetic algorithm for the energy-efficient virtual machine placement problem in data centers. - Neural Processing Letters, Vol 41, 2015, No 2, pp.211–221.

[23] W. Vogels.Beyond server consolidation. - ACM Queue, Vol , 2008, No 1, pp. 20–26.

[24] M. Wang, X. Meng et al. Consolidating virtual machines with dynamic bandwidth demand in data centers. - 30th IEEE International conference on computer communications, Shanghai, China, 2011, pp. 71–75.

[25] T. Wood, P. Shenoyet al. Sandpiper: Black-box and gray-box resource management for virtual machines. - Computer Networks, Vol 53, 2009, pp. 2923–2938.

[26] H. Zhong, K. Tao et al. An approach to optimized resource scheduling algorithm for open-source Cloud systems. - China Grid Conference (China Grid), 2010 Fifth Annual, Guangzhou, China, 2010 pp. 124–129.

Methodological papers

Annales Mathematicae et Informaticae
46 *(2016) pp. 245–250*
http://ami.ektf.hu

On a variant of the Lucas' square pyramid problem

Salima Kebli[a], **Omar Kihel**[b]

[a]Départment de Mathématiques, Université d'Oran 1 Ahmed Benbella
Bp 1524, Algeria
kabli.salima@univ-oran.dz

[b]Department of Mathematics, Brock University, Ontario, Canada L2S 3A1
okihel@brocku.ca

Submitted October 31, 2015 — Accepted February 24, 2016

Abstract

In this paper we consider the problem of finding integers k such that the sum of k consecutive cubes starting at n^3 is a perfect square. We give an upper bound of k in terms of n and then, list all possible k when $1 < n \leq 300$.

Keywords: Diophantine equation, Lucas' square pyramid problem, sum of squares, sum of cubes

MSC: 11A99, 11D09, 11D25

1. Introduction

The problem of finding integers k such that the sum of k consecutive squares is a square has been initiated by Lucas [3], who formulated the problem as follows: when does a square pyramid of cannonballs contain a number of cannonballs which is a perfect square? This is equivalent to solving the diophantine equation

$$1^2 + 2^2 + 3^2 + 4^2 + \cdots + k^2 = y^2. \tag{1.1}$$

It was not until 1918 that a complete solution to Lucas' problem was given by Watson [5]. He showed that the diophantine equation (1.1) has only two solutions, namely $(k, y) = (1, 1)$ and $(24, 70)$. It is natural to ask whether this phenomenon keeps occurring when the initial square is shifted. This is in fact equivalent to solving the following diophantine equation

$$n^2 + (n+1)^2 + \cdots + (n+k-1)^2 = y^2. \tag{1.2}$$

This problem has been considered by many authors from different points of view. For instance, Beeckmans [1] determined all values $1 \leq k \leq 1000$ for which equation (1.2) has solutions (n, y). Using the theory of elliptic curves Bremner, Stroeker and Tzanakis [2] found all solutions k and y to equation (1.2) when $1 \leq n \leq 100$. Stroeker [4] considered the question of when does a sum of k consecutive cubes starting at n^3 equal a perfect square. He [4], considered the case where k is a fixed integer. In this paper we take $n > 1$ a fixed integer and consider the question of when does a sum k consecutive cubes starting at n^3 equal a perfect square. We will give in theorem 1 an upper bound of k in term of n, and then use this upper bound to do some computations to list all possible k when $1 \leq n \leq 300$. Our method uses only elementary techniques.

2. The sum of k consecutive cubes being a square

Stroeker [4] considered the question of when does a sum of k consecutive cubes starting at n^3 equal a perfect square. He [4] considered the case where k is a fixed integer. This is equivalent to solving the following diophantine equation:

$$n^3 + (n+1)^3 + \cdots + (n+k-1)^3 = y^2. \tag{2.1}$$

The problem is interesting only when $n > 1$. In fact, when $n = 1$, because of the well known equality $1^3 + 2^3 + \cdots + k^3 = \left(\frac{k(k+1)}{2}\right)^2$ equation (2.1) is always true for any value of the integer k. Stroeker [4] solved equation (2.1) for $2 \leq k \leq 50$ and $k = 98$. We prove the following.

Theorem 2.1. *If $n > 1$ is a fixed integer, there are only finitely many k such that the sum of k consecutive cubes starting at n^3 is a perfect square. Moreover, $k \leq \lfloor \frac{n^2}{\sqrt{2}} - n + 1 \rfloor$.*

Proof. The equality

$$1^3 + 2^3 + 3^3 + \cdots + (n-1)^3 = \left(\frac{(n-1)n}{2}\right)^2$$

leads

$$n^3 + (n+1)^3 + \cdots + (n+k-1)^3 = \left(\frac{(n+k)(n+k-1)}{2}\right)^2 - \left(\frac{n(n-1)}{2}\right)^2.$$

Hence equation (2.1) gives

$$\left(\frac{(n+k)(n+k-1)}{2}\right)^2 - \left(\frac{n(n-1)}{2}\right)^2 = y^2.$$

It is well known that the positive solutions of the last equation are given by

$$\begin{cases} \frac{(n+k)(n+k-1)}{2} = \alpha(a^2+b^2), \\ \frac{n(n-1)}{2} = \alpha(a^2-b^2) \qquad\qquad \alpha \in \mathcal{N} \\ y = \alpha(2ab) \end{cases} \tag{2.2}$$

or

$$\begin{cases} \frac{(n+k)(n+k-1)}{2} = \alpha(a^2+b^2) \\ \frac{n(n-1)}{2} = \alpha(2ab) \qquad\qquad \alpha \in \mathcal{N} \\ y = \alpha(a^2-b^2) \end{cases} \tag{2.3}$$

where $a, b \in \mathbb{N}$, $\gcd(a,b) = 1$, $a > b$, $a \neq b \pmod 2$. The first equation in system (2.2) gives that

$$(n+k-1)^2 < 2\alpha(a^2+b^2). \tag{2.4}$$

The second equation in system (2.2) gives

$$\frac{n^2}{2} > \frac{n(n-1)}{2} = \alpha(a^2-b^2) \geq \alpha(a+b).$$

Hence

$$\left(\frac{n^2}{2}\right)^2 > (\alpha(a+b))^2 \geq \alpha(a^2+b^2). \tag{2.5}$$

Inequality (2.4) combined with inequality (2.5) yield

$$(n+k-1)^2 < 2\alpha(a^2+b^2) \leq 2\left(\frac{n^2}{2}\right)^2.$$

Whereupon

$$n+k-1 < \frac{n^2}{\sqrt{2}},$$

hence,

$$k \leq \frac{n^2}{\sqrt{2}} - n + 1.$$

The second equation in system (2.3) implies that

$$\frac{n(n-1)}{2} = 2\alpha(ab),$$

hence

$$\frac{n^2}{4} > \alpha ab.$$

This last inequality combined with the first equation in system (2.3) yield

$$2\left(\frac{n^2}{4}\right)^2 > 2\alpha^2a^2b^2 > \alpha(a^2+b^2) > \left(\frac{n+k-1}{2}\right)^2.$$

Whereupon

$$k \leq \frac{n^2}{\sqrt{2}} - n + 1. \qquad \square$$

3. Some computations

Based upon Theorem 2.1, we wrote a program in MAPLE and found the solutions to equation (2.1) for $1 < n \leq 300$. The solutions are listed in the following table.

n	k	y^2
4	1	64
9	1	729
	17	104329
14	12	97344
	21	345744
16	1	4096
21	128	121528576
23	3	41616
25	1	15625
	5	99225
	15	518400
	98	56205009
28	8	254016
33	33	4322241
36	1	46656
49	1	117649
	291	3319833924
64	1	262144
	42	26904969
	48	34574400
69	32	19998784
78	105	268304400
81	1	531441
	28	24147396
	69	114383025
	644	68869504900
88	203	1765764441
96	5	4708900
97	98	336098889
100	1	1000000
105	64	171714816

Remark 3.1. Let $C_n = |\{(k, y) \text{ solution to equation (2.1) }\}|$. We see from theorem 1, that for every n, C_n is finite, and from the table above, that for $1 \leq n \leq 300$, $C_n \leq 7$. It would be interesting to see if there exists a constant C such that $C_n \leq C$ for every n.

111	39	87609600
118	5	8643600
	60	200505600
120	17	35808256
	722	125308212121
121	1	1771561
	1205	771665618025
133	32	106007616
144	1	2985984
	13	43956900
	21	77053284
	77	484968484
	82	540423009
	175	2466612225
	246	5647973409
153	18	76055841
	305	10817040025
165	287	10205848576
168	243	6902120241
169	1	4826809
	2022	5755695204609
176	45	353816100
	195	4473603225
189	423	34640654400
196	1	7529536
216	98	1875669481
	784	248961081600
217	63	976437504
	242	10499076225
	434	44214734529
221	936	446630236416
225	1	11390625
	35	498628900
	280	15560067600
	3143	32148582480784
232	87	1854594225
	175	6108204025
256	1	16777216
	169	7052640400
	336	29537234496
	1190	1090405850625
265	54	1349019441
	2209	9356875327801

289	1	24137569
	4616	144648440352144
295	76	2830240000
298	560	133210400400

Acknowledgements. The authors express their gratitude to the anonymous referees for constructive suggestions which improved the quality of the paper. The second author was supported in part by NSERC.

References

[1] L. Beeckmans, Squares expressible as sum of consecutive squares, *Amer. Math. Monthly* 101 (1994), no. 5, 437–442.

[2] A. Bremner, R. J. Stroeker, N. Tzanakis, n sums of consecutive squares, *J. Number Theory* 62 (1997), no. 1, 39–70.

[3] E. Lucas, Question 1180, *Nouvelles Annales de Mathématiques*, ser. 2, 14 (1875), 336.

[4] R. J. Stroeker, On the sum of consecutive cubes being a perfect square. *Special issue in honour of Frans Oort. Compositio Math.* 97 (1995), no. 1–2, 295–307.

[5] G. N. Watson, The Problem of the Square Pyramid, *Messenger of Mathematics* 48 (1918–19), 1–22.

Annales Mathematicae et Informaticae
46 *(2016) pp. 251–264*
http://ami.ektf.hu

Equivalence relation as a tool to create new structures
How could they be prepared and taught in schools?

Z. N. Lehocká[a], Ö. Vancsó[b]

[a]Constantine the Philosopher University in Nitra
znlehocka@ukf.sk

[b]Eötvös University, Budapest
vancso.odon@gmail.com

Submitted November 12, 2015 — Accepted April 18, 2016

In this paper, we look for responses to ideas of introducing the teaching of number and vectors using equivalence relations with children 10-13 years old. We have to express that our goal with this paper is to wake up the interest of teachers and researchers only and we have not elaborated a concrete teaching material or textbook part for this level. We analysed the problems with vectors and rational numbers from the *theoretical point of view* including correct mathematical thoughts. Therefore our work could be seen as *a starting point* for developing new materials and experiments with pupils 10-13 years old. Our interest for this topic was awakened by personal experiences of the authors from the teacher training at university level. We experienced, that students are very confused about vectors and directed segments even as rational numbers and fractions. The relation between these notions is not clear.

In the textbooks used in Hungary and Slovakia, equivalence relations are not *explicitly* mentioned when introducing pupils to such ideas as rational numbers (fractions) and vectors, where the equivalence relation plays a crucial role. The reason for this, in our opinion, is that the authors of those textbooks think (otherwise well) that the notion of equivalence relation is too complicated for pupils at that stage. It is right but followed the ideas T. Varga the complex mathematical notions should be introduced as early as possible to pupils. However, given that there are problems with the introduction of the notions of rational numbers and vectors in our countries, in this article we try to argue against this trend, showing how we could teach those ideas to students and using equivalence relations from an early age and going step by step, developing the idea leading to correct notions later on.

In the first part of the article we show, using textbooks from Hungary and Slovakia, how the notions of rational numbers and vectors are currently introduced in school mathematics. In the second part, we discuss supporting ideas from psychology and cognitive science related to the building of a concept. The third part consists of exercises for simple structures we can rely on in giving tasks for practice, by the help of which the conceptual progress related to the rational numbers might be promoted. In the fourth part, we summarize the ideas of a framework we have used when discussing these ideas with prospective teachers of mathematics.

1. Textbooks

In Hungary, there are many different textbook families from which, normally, the teacher chooses after a common decision within the department of mathematics. In Slovakia, there is a centralized educational system for textbooks, where, usually, only one textbook exists for all students of the same grade. Therefore, our examples are more varied in Hungary.

1.1. Hungarian examples

Firstly,

> A rational number is such a *number* which can be written in the form $\frac{a}{b}$ where both a and b are whole numbers (and $b \neq 0$). See e.g. in Hajdu, S. [7, p. 18], Csatár, K. [5, p. 28] [translation Vancsó, Ö.]

It is not clear what the connection is between the terms fraction and rational number. Another weakness of this definition is that it assumes we know what *number* means. But this is the main problem. We want to extend the notion of number (from whole to rational) and in such a case we can only use the earlier notion of number and not the new one. This is a typical mistake repeated later in the case of real numbers. An exception is where the "value" of a fraction is introduced as a rational number (Vancsó, Ö. [10, p. 69]). This is correct but it uses a non-mathematical notion, which needs to be mathematized later.

Secondly,

> A vector is a segment with a given direction. See Csatár, K. [6, p. 199.], Koller, M. [8, p. 224], Korányi, E. [9, pp. 279–281].

In this sentence, two different notions are identified, and, with the exception of a single book (see [11, p. 253]), there is a similar introduction in all textbooks. It is interesting that both [8] and [9] introduce the idea of a vector connecting it to the idea of translation but mixing the notions of directed segment and vector. The definition says a vector is an equivalence class of segments that have the same direction and the same length. There is not elaborated enough clearly the difference between the translation as a "global" transformation of a plane (space) and the result of this translation for each one points.

1.2. Slovakian examples

The main focus in the current mathematical books in the field of teaching rational numbers is dual. On the one hand the emphasis is on the display when two different notations (their forms) mean the same rational number. On the other hand the emphasis is on the possible ways of comparison by size of rational numbers having different forms and not the least to make the students able to determine the image of rational numbers on the number line.

The tasks and examples of the given chapter also focus on this. They mention the reduced form of rational numbers, but the fact that the reduced form can be seen as the class representative of the given rational number is not showed. The formulation that one point of the number line is the image of infinitely many but equivalent fractions is also included. The various fractions express the same rational number. The fact, that a rational number can be expressed with the help of different fractions, is also included, but the proof of that is not included according to the teachers. It is said, that the result of the operations does not depend on which form of the fractions are used, by which the rational number is expressed.

Firstly, in summary, we can say that the set of rational numbers is introduced in a way that makes it unclear what kinds of numbers constitute the set (see Sedivy a. coll. [12, pp. 3335], and Sedivy a. coll. [13, pp. 28-35]).

Fractions and rational numbers are dealt with in the 6th-7th grades (11-12 year olds) without mentioning that rational numbers come from classifying fractions. This might be the kernel of the problem. Even at university level, students are still not quite aware of what fractions or rational numbers are. On the other hand, the textbook speaks about the value of the fraction, which is vividly described and illustrated, though mathematically not a well-established notion. Basically the value of a fraction should be made mathematical, which could then lead to the notion of rational numbers. These two notions follow each other, thus certain topic headings speak about rational numbers, while the subheadings are about operations with fractions. Serious misunderstandings may be created because the relationship between these notions is not clarified. The notion of rational numbers is not explained in any form, thus, in place of the rules referring to the operations with rational numbers, there are always fractional operations.

Secondly,

> "It is a rarity that a notion is surrounded by so many mysteries, and

> carry so many meanings as the notion of the vector. The substance of the hardships might come from the fact that we can manipulate with single numbers but the world surrounding us cannot be described by single numbers. If a single number is not sufficient, then we try to describe the situation by more numbers and we can speak about vectors right away... For us there will always be only one meaning for the vector and this is a shift." (Discussion of vectors in the 3rd class (16-17 year olds) of secondary schools Hecht T. [15, pp. 25] (translation Lehocka, Z.)

It is fairly confusing not to explain why numbers make a vector; on the other hand the notion is identified clearly with a geometrical transformation which might be hard to understand at first hearing (see Hecht T. [14, pp. 24–25]). One positive aspect of this introduction is that later on this notion can be developed and its relation to directed segments becomes clear. This differs from the Hungarian textbooks, where a vector is identified falsely with a transformation which is undoubtedly not identical to a directed segment. Of course it might be represented that way (see later).

2. Building concepts

The understanding of the equality of fractions, i.e. the fact that different fractions can express the same quantity, is crucial in the linking of quantities and fractions, and also in the perspective of adding and subtracting fractions. Researches show that fraction equivalence is not easy to understand for all students.

The wish to become acquainted with the surrounding world develops in children at an early age. As teachers, we have to build on children's natural curiosity and interest. An interesting task or a problem situation tailored to the child's level of development may challenge them and mobilize their inventiveness. Tamás Varga [4] said the following: "if the child comes to know the geometrical phenomena through aesthetic experience, if we let them learn mathematics through their toys, this world will not be a strange land to them. The connections between thinking and observations, activity-based experience, clearly indicate that children learn how to see, observe, solve tasks and think via actions, that is, independent activities. Teachers who are open, accepting and affirmative towards children's exhibitions of initiative and inventiveness contribute to the development of thinking. We can strengthen in children their feeling of self-trust, initiative and a need for searching out, new ways as well as finding suitable solutions among various possibilities."

According to Bruner, teaching must be built on the structures of mathematics. The advantage of this method is a more easily comprehensible syllabus if students understand the basics; single things and details are soon forgotten if they are not treated in structured forms; understanding basic phenomena contributes to transferring effect; and the difficulties of transmission to higher grades of schools are lessened. Processing the syllabus in an intuitive form in an early phase as well as a later re-discussion of the given topic on a level that suits the development of a certain age group is necessary for a more complex understanding of that syllabus by students.

Concepts are created by putting certain objects in one group according to their existing common characteristics (abstraction). For example, the family's car and the neighbour's car are objects which are alike (since they have a car-body, four seats, can be driven, and can get us to faraway places). A common word „car" has been made for the naming (classifying) of these objects. Naming (classifying) is as important as the concept creating process itself. The primary concepts abstracted from the objects might also be classified according to their existing common characteristics creating secondary concepts (cars, buses and bicycles are all vehicles). Third-level concepts may be abstracted from the second-level concepts and from those even higher level concepts. During the abstraction process it is more expedient to illustrate the concept by examples instead of definitions. Text in facts is the only viable way.

For example, when we say to a child that the colour red is a sensation which is created by the rays of light of about 0.6 micron wavelength reaching the eyes, they cannot create the concept red and cannot say whether a given car is red or not. On the other hand, if we show a lot of examples of red, after some time they will create the concept of red themselves.

In the article we undertake a theoretical concept to the introduction of the rational number and vector terms with the help of equivalence classes. The basic idea of our concept is that, by gradual building, the design of these terms is done intuitively. Meanwhile it is lucky if the students understand the concepts in context.

Our goal is not to create new definitions or theses. We rather aim to outline a possible way in which, by suitable tasks, certain mathematical concepts can be introduced and substantiated.

According to Richard Skemp [3, p. 20-45], in mathematics, the only way of teaching concepts which students have not acquired yet is to help children organize a suitable set of examples into one group in their mind.

If the new concept is a kind of primary concept such as for example, red, we can do this without using any symbols, simply by pointing. If the concept is a secondary one such as all mathematical concepts, then the only method of helping the student collect the set of suitable examples into one group in their mind is collecting the suitable words.

Definitions are not completely useless because they close scientific arguments, ensuring the possibility of unambiguous classification as well as ensuring the proper positioning of the concept in the structures of the concepts of the individuals. However their function is only secondary in the original creation of the concepts.

Sciences work with many concepts and the basic concepts have to be learnt at primary and secondary schools. Without understanding the first concepts, anything built on it, will not be comprehended. If we do not understand which phenomenon/situation is covered in the first concept we will not understand the secondary- and third-level concepts built on the first one. For example, if the concept of the circle is not clear (or if a student does not understand what a set of points equally distant from a given point means), then the teacher will not be able to teach children the triangle's circumcircle either.

Creating clear concepts, seeing the essence and setting up unambiguous and simple models are important functions of a teacher. Clear concepts are needed for logically proper statements. Existence of clear concepts means that students find adequate concepts for various situations. It is also important that students can differentiate two close concepts, for instance parallelogram and trapezium, and as will be discussed later directed segment and vector.

3. Constructing structures by using equivalence relations

Mathematics textbooks were examined from the viewpoint of dealing with the relation between fractions and rational numbers as well as vectors and directed segments. The idea of equivalence behind them and the building of structures are usually not stated explicitly, thus missing the clearing up of the notions and their relations. How could we avoid these mistakes? This is an important goal of this article. Let us give an example to clarify our situation.

The first step can be made in primary schools where all notions are introduced by classification. The essence is always to follow this classification, which means the pupils have to understand how well defined "classes"could be constructed from difference things, objects. We have to know which things belong to a certain group and we can categorize all things in such a way.

3.1. The introduction of equivalence relations and classes

Equivalence classes should be introduced in early ages, grades 1-3 (6-8 years olds).

Example: take a set of words (more or fewer words, depending on the age), for example {Johnny, more cunning, horse, wonderful, pen, intimate, Adam}. Students are asked to group the words according to the numbers of vowels. So, the words with equal numbers of vowels will belong to the same group. As a support, we should use a set of letters or a magnetic ABC set so that students can find the solution by manipulation. The set is freely expandable and we can play with the idea that this is a real classification (i.e. as relations, they are reflexive, symmetrical and transitive, since this is the condition of becoming a real classification).

Let H be a set of a plane's lines. Let us examine the features of the parallel relation on this set H. We can use a match-stick, a drinking-straw and a bigger cardboard sheet. Firstly students examine mutual relations of one, two then three lines. They will experience intuitively that this relation is reflexive, symmetrical and transitive, thus a classification on the set of the plane's lines. The classification might be identified with the direction; therefore there are as many classes as many different directions. This example is more complex, two steps further in abstraction than the earlier one was. The number of classes is infinite, and the pupils can draw them but they cannot manipulated them.

To illustrate this structure we give three different examples, which help to understand how new structures among equivalence classes can be built up using the old one among the original elements of the classes.

3.2. Finite constructions

3.2.1. Residue classes by the division with a given positive whole number

In a more simple case (compared to rational numbers or vectors) we consider a finite number of classes.

It is useful to make pupils do many exercises with this case.

Example 3.1. Divisibility by 2. There are only two classes: the even and the odd numbers (for example: 4 or 13). We can define the "sum" of two classes since the sum of two even and two odd numbers is always even and the sum of one odd and one even number is odd. So, regardless of the representative elements of the classes always the same result occurs. If we note these classes for example by E(ven) and O(dd) then our addition rules can be expressed by: $\mathrm{E} + \mathrm{E} = \mathrm{E}$, $\mathrm{O} + \mathrm{E} = \mathrm{O}$ and $\mathrm{O} + \mathrm{O} = \mathrm{E}$. Now we have to analyse the traditional rules of addition: associativity, commutativity, and the existence of neutral element and inverse. The first two are proved very easily, the neutral element is E, because: $\mathrm{E} + \mathrm{E} = \mathrm{E}$, $\mathrm{O} + \mathrm{E} = \mathrm{O}$. The inverse element is the same as the element itself: $\mathrm{E} + \mathrm{E} = \mathrm{E}$, consequently $\mathrm{E}^{-1} = \mathrm{E}$ similar way $\mathrm{O} + \mathrm{O} = \mathrm{E}$ so $\mathrm{O}^{-1} = \mathrm{O}$. (We call such a structure a group.)

The other operation is multiplication. We remark that if two odd numbers are multiplied, always an odd number comes out. In the other three cases always even numbers come out. It means the following rules: $\mathrm{O} \cdot \mathrm{O} = \mathrm{O}$, $\mathrm{E} \cdot \mathrm{O} = \mathrm{O} \cdot \mathrm{E} = \mathrm{E} \cdot \mathrm{E} = \mathrm{E}$.

It's useful to give children tasks which highlight the rules of associativity and commutativity and help them recognize these rules. There is a unit element (O), because $\mathrm{O} \cdot \mathrm{O} = \mathrm{O}$ and $\mathrm{E} \cdot \mathrm{O} = \mathrm{E}$. It is interesting that among the whole numbers, which are elements of these two classes, there is not an inverse element but in our new structure there is. It means that this is a field.

Such examples and tasks can be posed for primary school children as well.

Example 3.2. Division by 5.

There are five classes here, which means much more complicated operation tables (see below). These tables can be constructed by pupils of grades 6-7 (12-13 years olds). They can check that the class operations are independent of choosing the representative elements from the classes. The two tables are Table 1 and Table 2.

In the general case the division goes by n (this case is $n = 5$) but this is too abstract for primary-school pupils, it is for secondary school students of grades 11 or 12 (17-18 years olds).

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

Table 1: Addition

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Table 2: Multiplication

These ideas are intended for further elaboration for teachers and textbook writers and for structures by which students might be more activate, and by which we could prepare the operations among the desired equivalence classes, hereby approximating our aims.

3.2.2. Iterated sums of the digits of natural numbers

This is an interesting example which appears to have an even more complicated structure than examples 3.1 and 3.2.

We will represent the natural numbers by 1, 2, 3, ..., 9 the following way. In the first step we sum all the digits of the given number. If the result is less than 10, the process is finished. If not, it will be repeated. After the final step we reach a number between 1 and 9. This construction classifies all positive integers into 9 different classes, and induces an equivalence relation on them as it shown below.

An example: 198564218 is equivalent to 3752, since $1+9+8+5+6+4+2+1+8 = 44$ and after the second step $4+4$ we get 8, similarly $3+7+5+2 = 17$, and after the second step $1 + 7 = 8$. It is easy to see that this is indeed an equivalence relation (simply because it is based on a partition, which always induces an equivalence relation). Furthermore, we may construct a "number type" structure on the set of these nine classes according to the original sum and product operations given on the set of natural numbers. We offer this example for secondary students of grades 10-11 (16-17 years of age).

The operations introduced according to this equivalence relation are well-defined, because the classes are actually the remainder classes of division by 9. The reason of this is the fact that a positive integer has same remainder after division by 9 as the sum of its digits. Therefore the result of an operation does not depend on choosing the representative elements from the classes. For example, the sum of classes "1" and "2" is the class "3". The question is whether sum always belongs to class "3", if we choose other representative numbers from both classes "1" and "2". Example: $46 \to 10 \to 1$; $89471 \to 29 \to 11 \to 2$; $46 + 89471 = 89517 \to 30 \to 3$, it demonstrates that (at least in this case) the operation works well. The case of multiplication is similar. We believe this example is suitable for creative students and provides a possibility to exercise building mathematical structures, in this case by using number theoretic methods.

3.3. Rational numbers

Young children are already able to organise things in different classes. For example, vehicles or games. These experiences can be extended in the case of fractions. A fraction is a relation between two natural numbers (a ratio). It is possible to create classes among fractions. Two fractions $\frac{a}{b}$ and $\frac{c}{d}$ belong to one class if and only if the following two products are the same: $ad = cb$. (That means the "value" of the two fractions are the same but this is not a mathematical notion.) An important step is to show that this relation is really a classification, meaning mathematically to fulfill three assumptions: reflexivity, symmetry and transitivity. The first is very formal for pupils it is enough to deal with symmetry and transitivity. Symmetry can be proven very easily (because multiplication is symmetrical), the second is a bit more complex. If $\frac{a}{b}$ and $\frac{c}{d}$ and also $\frac{c}{d}$ and $\frac{e}{f}$ belong to the same class, then it is true that $\frac{a}{b}$ and $\frac{e}{f}$ also belong to the same class. To prove this we have to write the assumptions in the form of equations: $ad = cb$ and $cf = ed$. We have to derive from them that $af = eb$. In order to derive this let the two given equations be multiplied: $(ad)(cf) = (cb)(ed)$. Using two rules of multiplication, associativity and commutativity we then get $afcd = ebcd$. Dividing through by cd we have finished the proof.

Of course, in school such tasks have to be posed where the connection with concrete numbers can be seen. There are links again to number theory and divisibility.

The next and most complicated step is the introduction of addition and multiplication between classes. Our goal is *to regard one class* as *a "new number"*. To do this we have to formulate the operations. We wanted to use the operations of fractions. This is the correct way if we can prove that these operations are well defined because the result is independent of which element of a class was chosen.

a) Addition

Let A and B be two classes of fractions. We define C as a "sum" of these classes in such a way that if $\frac{a_1}{a_2} \in A$ and $\frac{b_1}{b_2} \in B$, then C will be the class where the fractions $\frac{a_1 b_2 + b_1 a_2}{a_2 b_2}$ belong. It is a good definition only if the result does not depend on the representatives of A and B. To prove it we have to show, if $\frac{a_1^*}{a_2^*} \in A$ and $\frac{b_1^*}{b_2^*} \in B$, then $\frac{a_1^* b_2^* + b_1^* a_2^*}{a_2^* b_2^*} \in C$ is true.

In order to prove it we have to point out that

$$(a_1^* b_2^* + b_1^* a_2^*)(a_2 b_2) = (a_1 b_2 + b_1 a_2)(a_2^* b_2^*)$$

supposing that $a_1a_2^* = a_1^*a_2$ and $b_1b_2^* = b_1^*b_2$. The following equation chain shows why $\frac{a_1^*b_2^*+b_1^*a_2^*}{a_2^*b_2^*} \in C$ is true. $(a_1^*b_2^* + b_1^*a_2^*)(a_2b_2) = a_1^*b_2^*a_2b_2 + b_1^*a_2^*a_2b_2 = a_1a_2^*b_2^*b_2 + b_1b_2^*a_2^*a_2 = (a_1b_2 + b_1a_2)(a_2^*b_2^*)$. Therefore the addition is well defined.

b) Multiplication

Let A and B are two classes of fractions. We define C as a "product" of these classes in the following way: if $\frac{a_1}{a_2} \in A$ and $\frac{b_1}{b_2} \in B$ then C will be the class where the fractions $\frac{a_1b_2}{a_2b_2}$ belong. Our goal is again to show, that this is a good definition meaning if $\frac{a_1^*}{a_2^*} \in A$ and $\frac{b_1^*}{b_2^*} \in B$ then $\frac{a_1^*b_1^*}{a_2^*b_2^*} \in C$ as well. We know that $a_1a_2^* = a_1^*a_2$ and $b_1b_2^* = b_1^*b_2$. We have to prove that $a_1b_1a_2^*b_2^* = a_1^*b_1^*a_2b_2$. On the left-hand side, we can replace $a_1a_2^*$ by $a_1^*a_2$, then $a_1b_1a_2^*b_2^* = a_1^*a_2b_1b_2^* = a_1^*b_1^*a_2b_2$ since $b_1b_2^* = b_1^*b_2$. This means that the multiplication is independent of the representative elements, consequently it is correctly defined.

In school we do not think these abstract proofs should be derived, only shown by concrete number examples. We have just proved the operations are well defined. The next step is, if we would like to regard these classes as numbers, to check the usual rules of these operations. This means that for multiplication it is a commutative group similarly to addition. Furthermore, multiplication is the distributive with respect to addition. Most of these rules can be derived by using similar rules of whole numbers. The only exception is the inverse element for multiplication (this does not exist for whole numbers). We can prove easily that $A^{-1} = B$ where B is defined by the following way: if $\frac{a}{b} \in A$ then $\frac{b}{a} \in B$. $\left(\frac{a}{b}\right)^{-1} = \frac{b}{a}$, since $\frac{a}{b} \cdot \frac{b}{a} = 1$ and the class which contains 1 is the identity element for multiplication. Of course the identity element also has an inverse which is the identity element itself, because only the fractions written in form $\frac{a}{a}$ belong to class "1".

In this way, the student has to understand the difference between fractions and rational numbers and see an important example of how a new structure can be constructed by using an older structure.

3.4. Vectors

There are at least two different ways to introduce vectors. One way is geometrically, using the translation as a transformation (as in the Slovakian textbook) and it is briefly written at the end of chapter 1.1. The first step is on this way is to distinguish between the transformation as a function whose domain is the points of a plane (or space) and the range is the same plane (space) and the concrete operation from points to points. Each point and its image define a directed segment. These are parallel, have the same lengths and direction. One of them can represent the translation as a transformation. This distinction is the most important to avoid the misconceptions.

The notion of vector, is perhaps more confusing than the notion of rational numbers. Among fractions there are operations which are the basis of the operation of their classes. Among directed sections which are the elements of the classes

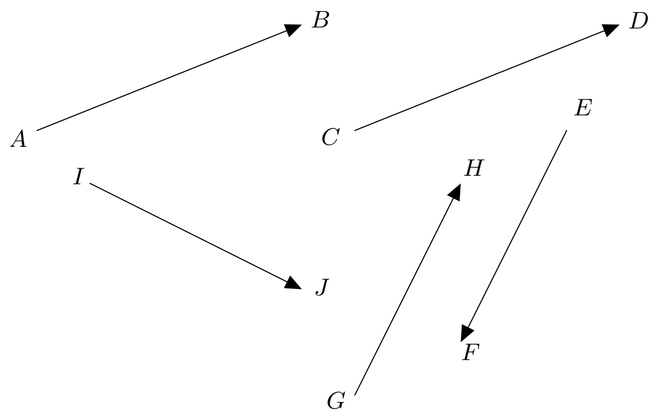


Figure 1

(named later vector) there is only addition in restricted situations, the other operation, named scalar multiplication, is defined without any restriction.

We would like to follow another way, using directed sections and equivalence classes as before.

Our starting point is the set of all directed sections and two operations between them: addition and scalar multiplication (which means a directed segment can be multiplied by a real (or earlier by a rational) number. Our way is: first define a vector as a class of directed segment. Two directed segments are equal (equivalent) if they are parallel, have the same direction and the same lengths (to put it briefly, using a translation they can be rapped into each other). The next step in our case is to define addition and multiplication between classes and to show they are well-defined operations. First step: introduce an equivalence relation $\approx$ the above mentioned way.

In Figure 1 $\overrightarrow{AB} \approx \overrightarrow{CD}$, but only these two sections are equivalent.

It is easy to prove that this relation is really equivalence. We then introduce addition among classes. We have to choose such representatives of the classes which either have common starting points or the ending point of one is the same as the second's starting point, see Figure 2.

To prove this addition is well-defined we have to use only translations, if we have chosen other representatives, the whole picture will be the "same" only translated (see figure 3). This shows that addition is a well-defined operation on the equivalence classes of directed sections. One class can be regarded as a vector. The above operation has important characteristics which are those of a commutative group. After this operation another can be defined as well, the so called multiplication by a scalar (see figure 4). Here it is again easy to prove the well-defined operation. The translation plays the crucial role again.

An easily provable rule of this operation is: $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$. The connection of

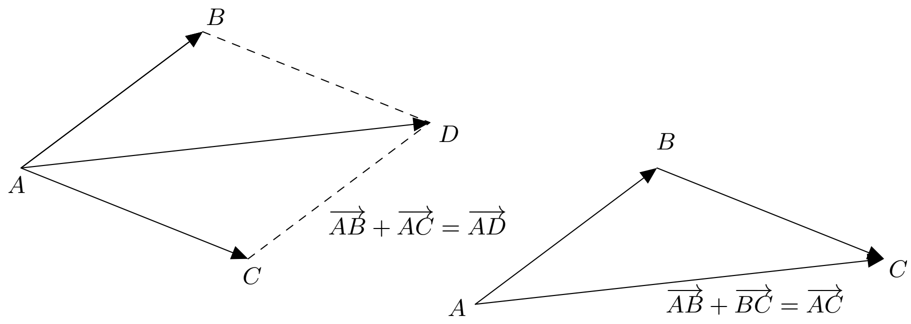


Figure 2

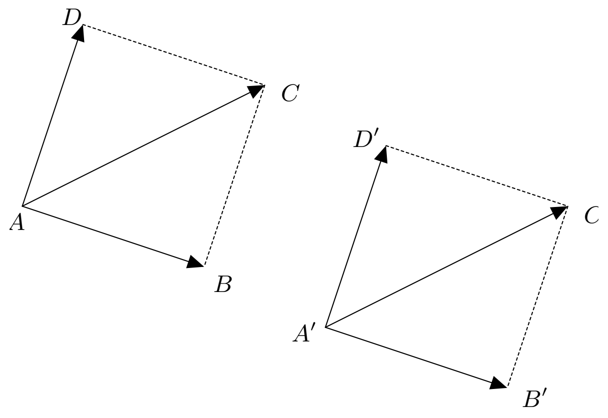


Figure 3

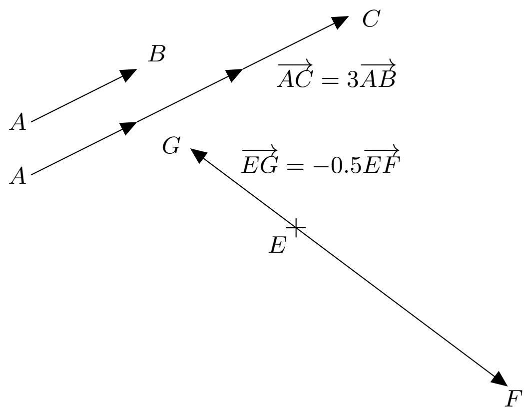


Figure 4

the two operations is the following two forms of the distributive law:

$$(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$$
$$\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$$

Both of them can be proved using directed segments as representatives of vectors to check these rules which belong to the required axioms of vector-space.

The most important thing is to see the difference between directed sections and vectors. They are on different levels of the building of mathematics. Helping later to understand the abstraction of vector space which is algebraically and not geometrically constructed. It is very useful to distinguish and abstract the notion of vector from geometry. This is impossible if we follow the classical false way to introduce the notion of the vector.

4. Summary and new results

Our aim was to show the similar mathematical connection between fractions and rational number respectively through directed segments and vectors. Behind both case the equivalence relation stands. We collected some examples which could help to deal earlier with this relation in school mathematics.

Our aim was to show how the notions of rational number and vector are taught traditionally. We found some typical misconceptions dealing with these notions in the school-mathematics that were illustrated by different textbooks from our countries. We think these textbook-examples are everywhere typical not only in Hungary and Slovakia. Finally we sketched some ideas about teaching these notions, first of all, *for teachers* and *text book authors* in the future. It would be necessary to plan experiments for students 10-13 years old to verify our thoughts. We hope, we will be able to report such experiments in the next future.

References

[1] Bruner, J. S. (1974): Új utak az oktatás elméletéhez (New ways to the theory of teaching) Gondolat Kiadó, Budapest.

[2] Clark, J. M., Paivio, A. (1991). Dual coding theory and education. Educational Psychology Review, 3(3), 149–170.

[3] Skemp, R. (1975): A matematikatanulás pszichológiája (Psychology Of Learning Mathematics), Gondolat, Budapest.

[4] Varga, T. (1976): Játszunk matematikát! (Let's play mathematics!) 1.-2., Móra Könyvkiadó, Budapest.

Hungarian Textbooks:

[5] Csatár, K. (ed.): Matematika 7. Apáczai Publisher, 2003.

[6] Csatár, K. (ed.): Matematika 9. I. Apáczai Publisher 2009.

[7] Hajdú, S. (ed.) Matematika 6. Tankönyvkiadó 1988.

[8] Koller, M.: Matematika 9. Műszaki Kiadó, 1998.

[9] Korányi, E.: Matematika 9. Konsept-H Publisher 1999.

[10] Vancsó, Ö. (ed.): Matematika 7-8. Műszaki Kiadó 2001.

[11] Vancsó, Ö. (ed.) Matematika 9. Műszaki Kiadó 2003.

Slovakian Textbooks:

[12] Šedivý a kol.: Matematika az alapiskola 6.osztálya számára 1.rész, SPN, Bratislava, 2004.

[13] Šedivý a kol.: Matematika az alapiskola 7.osztálya számára 1.rész, SPN, Bratislava, 2006.

[14] Hecht T.: Matematika a gimnáziumok és szakközépiskolák 2. osztálya számára, Orbis Pictus Istropolitana, Bratislava, 2001.

[15] Hecht T.: Matematika a gimnáziumok és szakközépiskolák 3. osztálya számára, Orbis Pictus Istropolitana, Bratislava, 2003.

Annales Mathematicae et Informaticae
46 *(2016) pp. 265–276*
http://ami.ektf.hu

Gender differences in spatial visualization skills of engineering students

Rita Nagy-Kondor

University of Debrecen, Faculty of Engineering, Hungary
rita@eng.unideb.hu

Submitted May 30, 2016 — Accepted December 14, 2016

Abstract

Spatial visualization skills are very essential in engineering studies. This study investigated the effects of gender in important component of spatial skills among engineering students. A total of 126 first-year engineering students (68 female, 58 male) completed different spatial visualization tests. In order to analyze the obtained data, means, standard deviations and independent samples t-test were used.

Keywords: Spatial ability, mathematics education, engineering education

MSC: G20, G30, G40

1. Introduction

Spatial visualization skills are essential to success in several disciplines. Spatial thinking has an important role in the teaching and learning of mathematics process. Previous studies showed that this ability has positive correlations with geometry and mathematics education. (see [2, 40]) According to previous studies students with high scores on a mental rotation test systematically score higher on anatomy examinations (see [44]) and spatial visualisation ability is a predictor for success in technical education and it is very important in engineering training. (see [1, 16, 22, 23, 26, 27]) The skills needed to develop a person's spatial ability are acquired through programs or activities that teach engineering or drafting. (see [31]) Shea, Lubinski and Benbow (see [35]) state that the intellectually talented adolescents, who has better spatial than verbal abilities are more likely to be found in the field of engineering, computer sciences and mathematics.

According to Gardner (see [9, p. 9]) the "spatial intelligence is the ability of forming a mental model of the spatial world and maneuvering and working with this model". Spatial visualization can be defined as the abilities of imagine the visualization of an object from different viewpoints, rotation of it and blend or integrate of the parts of the given object (see [12, 21, 31]). McGee (see [21]) defines spatial ability as "the ability to mentally manipulate, rotate, twist or invert pictorially presented stimuli", McGee (see [21]) and Maier (see [18]) classify five components of spatial skills as

- Spatial perception: the vertical and horizontal fixation of direction regardless of troublesome information;
- Spatial visualization: it is the ability of depicting of situations when the components are moving compared to each other;
- Mental rotation: rotation of three dimensional solids mentally;
- Spatial relations: the ability of recognizing the relations between the parts of a solid;
- Spatial orientation: the ability of entering into a given spatial situation.

Several studies have indicated that there are various factors effecting spatial ability; one of these is gender. According to Tsutsumi et al. (see [37]) females are much less likely to get high scores in Mental Cutting Test (see [29, 30]). In general, boys have a higher spatial ability than girls which may be caused by biological and/or environmental factors (see [42]). And the related literature shows that there is a significant male advantage on mental rotation tasks at every age (see [32, 45]). Yet, difference between males and females on those tests are much smaller than those found in mental rotation and spatial perception (see [42, p. 90]). There are conflicting results in the reviewed literature. Previous studies (see [19, 39, 41]) found that there is not a significant difference between male and female groups' scores in spatial visualization of prospective elementary mathematics teachers. Due to these conflicting results the educators are still interested in gender difference in case of spatial ability.

In the light of the existing literature, the aim of this paper to evaluate MCT, HSVT, PSVT (developments, rotations (without coordinate axes), views) and PSVT-R (with coordinate axes) test results of first-year engineering students in University of Debrecen, with special emphasis on gender differences, in comparison with the international results. Such a comprehensive survey has never been made in our University.

2. Measurement of spatial ability

The measurement of spatial abilities is standardized by international tests, among which the Mental Cutting Test (MCT) is of greatest importance. MCT presents a 3D object with an imaginary cutting plane and five possible solutions for the cross-section shape. Heinrich Spatial Visualization Test (HSVT), Purdue Spatial Visualization Test (PSVT) and Purdue Spatial Visualization Test – Visualization of Rotation (PSVT-R) are widely used for testing the spatial ability.

Much work has been reported an analysis of MCT (see [1, 16, 37, 38]), HSVT (see [13, 39]) and PSVT-R (see [1, 7, 34, 36]) results of engineering students or prospective mathematics teachers. Most US researchers have used the PSVT-R to measure visualization skills; MCT is widely used in Europe and Japan.

The standard MCT (a sub-set of CEEB Special Aptitude Test in Spatial Relations, (see [5])) consists of 25 items. In each item presents a three dimensional object which is to be cut with an assumed cutting plane in perspective projection, and five alternative views of the resulting section (one correct alternative and four incorrect alternatives), resulting from the cut. The students is required to choose the single view that represents the correct section. It was shown by Saito et al. (see [33]) and Tsutsumi et al. (see [37]) that in order to solve the MCT problems, subjects go through three phases of information processing, which are:

- Recognizing the solid from the perspective drawing,
- Cutting the solid by the assumed cutting plane,
- Judging the characteristic quantity of the section, if necessary.

One of paper-and-pencil test was selected to measure spatial visualization ability of prospective elementary mathematics teachers: a reduced version of HSVT (planar geometrical transformations). This test was developed by Heinrich (see [13]) to examine the spatial abilities of engineering graphics students. The original HSVT includes two major expert skills in spatial visualization: synthesis and decomposition. For each two basic skills she hypothesized that when mental rotation was added to these tasks at three hierarchical levels of complexity, this would render the spatial problem solving progressively more difficult (see [6, p. 2]). The original HSVT consists of 48 items divided into 6 scales:

- synthesis without rotation;
- decomposition without rotation;
- synthesis with one-step rotation;
- decomposition with one-step rotation;
- synthesis with two-step rotation;
- decomposition with two-step rotation.

Example items of the test are given in the following figures.[1]

[1] Similar items used in the reduced test are given in the light of the examples of Chen (see [6]).

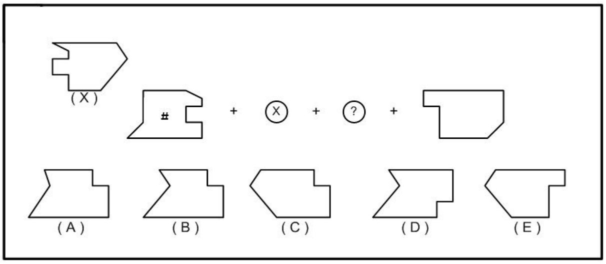


Figure 1: Example item for Synthesis section

Figure 1 expresses an example about the part of "Synthesis". Synthesize four pieces, adjusting Probe X to fit piece # and selecting one of 5 options A, B, C, D, E to replace the question mark. (see [6, p. 3]) Decompose given pattern three pieces, X + ? + Y, where probes X, Y may need to be adjusted, and after selecting one of 5 options, A, B, C, D, E to replace the question mark. The reduced test includes 15 items for the part of "synthesis", and 10 items for the part of "decomposition".

Guay (see [11]) developed the PSVT in 1976 to determine student's ability to visualize, recognize orthograpic drawings. The PSVT includes three sections: developments (identification of the figure), object rotations and views. Most researchers use only the object rotations portion (PSVT-R). The rotations section shows an object in two different positions. The first object is rotated on the X, Y or Z-axis, to show the rotation pattern. A second object is presented with five alternative views, one represents the second object subjected to the same rotation as the example. In our study coordinate axes were added to the first and second stimulus objects, but they were not added to the five solution choices (see [3]). The first stimulus object was shown in its new, rotated position. Figure 2 expresses an example about the part of PSVT-R (with coordinate axes) (see [3]).

3. Background

On the basis of our experience at the University of Debrecen, the high number of engineering students who fail to meet the requirements of the different foundation and special subjects can be explained by their insufficient secondary school knowledge in geometry, mathematics and physics. There are not sufficient opportunities for high school students to learn geometry, number of geometry class hours per secondary school year and especially number of three dimensional type class hours per secondary school year are very few.

Our aim of the introduced of "Basic Knowledge of Natural Sciences" course is to summarise and repeat those parts of elementary and secondary school mathematics, geometry and physics which are necessary as basic knowledge at our Faculty (see [28]). In the syllabus of the subject "Basic Knowledge of Natural Sciences" there are only secondary school topics, and from these topics only the ones which

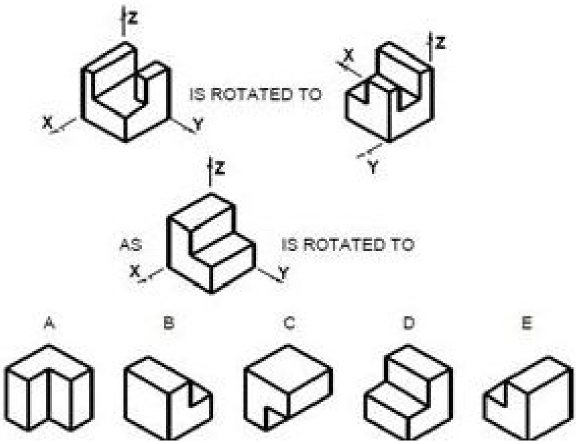


Figure 2: Example item for PSVT-R with coordinate axes added (see [3])

are highly important for engineering students. For example, in order to support teaching Descriptive Geometry, there are some topics of secondary school geometry: Basics of representation; Plane and space geometry; Coordinate geometry. Before they begin their studies, all students participate in this course. On the other hand, at our faculty the teaching of Mathematics is based on course book entitled "Mathematical Tools in Engineering Applications" (see [25]) to develop the mathematical knowledge. The book follows a new "engineer friendly" approach by demonstrating the application of mathematical tools through problems which typically occurring in the different engineering and economic fields.

4. Methodology

At the University of Debrecen 126 (68 female, 58 male) first-year engineering students took the tests. Subjects of the study are volunteered to participate and confidential feedbacks were given to those participants who are interested in. The students were recruited after "Basic Knowledge of Natural Sciences" course.

All data were collected by the author during the autumn semester. The test was administered in groups to explain instructions, efficiently. We used a personal information form for gathering data on gender. Standard instructions were given to two samples.

5. Results

Data were analysed using the SPSS statistical analysis program. The performance of the students is presented in the Figure 3.

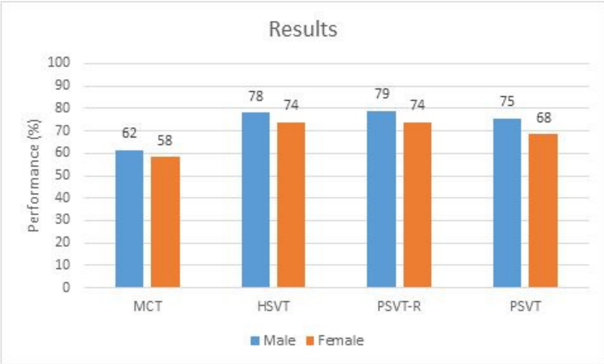


Figure 3: Results of Spatial Tests

Male students were 4% in MCT, 4% in HSVT, 5% in PSVT-R and 7% in PSVT better than female students on the tasks of spatial tests. The first-year engineering students scored 59.89% in MCT, 75.84% in HSVT, 76.3% in PSVT-R and 71.22% in PSVT.

According to these results, it can be said that Hungarian engineering students have adequate spatial performance. Means, standard deviations of spatial visualization ability and statistical differences of each group are analyzed in terms of descriptive statistics (Figure 4).

Test		Female	Male	Sex Difference
	n	68	58	
MCT	M	14.59	15.38	p>.05
	SD	4.85	4.91	
	n	52	54	
HSVT	M	18.63	19.46	p>.05
	SD	3.81	4.83	
	n	52	54	
PSVT-R	M	22.19	23.63	p>.05
	SD	4.90	4.97	
	n	68	58	
PSVT	M	25.62	27.14	p>.05
	SD	6.32	8.68	

Figure 4: Mean scores of each sample and statistical differences

Investigating of each sample's and all tests' means and standard deviations, we find that female/male engineering students mean score of MCT is 14.59/15.35 (SD = 4.85/4.91). Female/male engineering students mean score of HSVT is 18.63/19.46 (SD = 3.81/4.83). Female/male engineering students mean score of PSVT-R is 22.19/23.63 (SD = 4.90/4.97) and mean score of PSVT is 25.62/27.14 (SD = 6.32/8.68) (Figure 4). The results of female students did not reach the results of male students in all spatial ability tests. We found that there is not a significant difference between male and female groups' scores in spatial visualization in all spatial ability tests.

The performance of the students in PSVT is presented in the Figure 5.

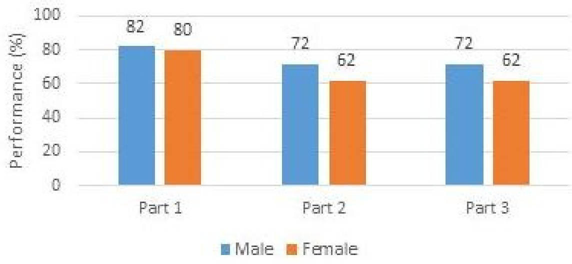


Figure 5: Results of PSVT

Male students were 2% (Part 1), 10% (Part 2) and 10% (Part 3) better than female students on the tasks of PSVT.

Means, standard deviations of part of PSVT and statistical differences of each group are analyzed in terms of descriptive statistics. The results appear in the Figure 6.

Investigating of each sample's and all part of tests' means and standard deviations, we find that female/male engineering students mean score of PSVT developments is 9.65/9.90 (SD = 1.91/2.83), mean score of PSVT object rotations is 7.50/8.62 (SD = 2.86/2.92) and mean score of PSVT views is 7.47/8.62 (SD = 2.83/3.82) The results of female students did not reach the results of male students in parts of PSVT. We found significant difference ($p < 0.05$) between male and female groups' scores in object rotations without axes.

Tasks of object rotation of PSVT focus on the imaginary manipulation of the solid. First task is the identification of the figure; second task is manipulation of mental representations. Many female students have bigger problems with imaginary manipulation without axes.

In general, the average scores of the PSVT-R are consistently reported to be around 75% across US four-year engineering universities; many other studies report similar data among engineering students. Results of the MCT test are comparable with an average around 60% for universities in the US, Australia and Europe (see [1, 7, 8, 10, 16, 36]). Spatial visualization skills of first year engineering students at the University of Debrecen are similar to the most engineering students' results (PSVT-

Test: PSVT		Female	Male	Sex Difference
Part 1: developments	n	68	58	p>.05
	M	9.65	9.90	
	SD	1.91	2.83	
Part 2: object rotations	n	68	58	p<.05**
	M	7.50	8.62	
	SD	2.86	2.92	
Part 3: views	n	68	58	p>.05
	M	7.47	8.62	
	SD	2.83	3.82	

Figure 6: Mean scores of PSVT of each sample and statistical differences

R is 76.3% and MCT is 59.84%) of international studies. Spatial visualization skills of entering first year engineering students at the Polytechnic of Namibia are significantly lower than those of most students in industrialized countries, but comparable to cohorts of minority engineering students in the US. Ault and John (see [1]) hypothesize that these differences are due to factors of prior experience and educational background.

6. Conclusion

Spatial ability of engineering students is greatest importance in terms of their professional achievement. Yet the results of the survey verify that many students have problems with imagining a spatial figure and therefore to solve the spatial geometry exercises. There are not sufficient opportunities for high school students to learn geometry. It would be very useful in the high schools and in the university training as well, if we devote more time for spatial ability, for summarizing the spatial geometry knowledge, for solving spatial geometry tasks. In general we can conclude that the basic geometry, mathematical and physical knowledge of students, who start their studies at Faculty of Engineering, is insufficient for their further engineering studies. It is very important to detect early the students with lower levels of spatial ability, because later will be negative impact by engineering studies.

The existing literature reports, that males perform better than females on spatial tests, especially on mental rotation tasks (see [1, 17, 29, 32, 37]). The results of female students did not reach the results of male students in all spatial ability tests, but we found that there is not a significant difference between male and female groups' scores in spatial visualization in all tests, excepting part 2 of PSVT, in object rotations without axes. Many female students have problems with imaginary manipulation without axes. We hypothesize that these differences in object rotation are due to use of axes in other mental rotation test (PSVT-R). The use of axes helps our female students for example in PSVT-R. It is possible that instead of manipulation of mental representations, female students watched shape and position of two-dimensional figure of solid in the coordinate plane, and thus deduce the solution. So it is useful to use several types of spatial test for measurement of spatial abilities.

In the future this difference can also be investigated in other areas of study. Based on studies of visualization skills we hypothesize that these differences are due to factors of students' prior experience and educational background.

Studies suggest that interactive animation and virtual solids are promising tools for training spatial thinking and we can achieve better results in mathematics classes with the use of dynamic geometry systems (see [4, 20, 23, 24, 36, 43]).

GeoGebra was designed to combine features of dynamic geometry software (e.g. Cabri Geometry, Cinderella, Geometer's Sketchpad) and computer algebra systems (e.g. Derive, Maple) in a single, integrated system for teaching and learning mathematics (see [14, 15]). Budai (see [4]) combined with plane space analogies for development of spatial ability and problem-solving skills in three dimensions of solid geometry. He presented the plane and spatial analogues of coordinate geometry with the help of GeoGebra, since part of the concept of this software is to show the algebraic and geometric view of the objects in parallel.

How we can help female students to develop their spatial ability, particularly in object rotations and views (and results of PSVT)? In the future it would be useful to focus on task based student interviews to reveal our female students' spatial problems, to find the reasons and typical mistakes of these female-male difference. Based on the results, as a future study, it is being planned to make special interactive worksheets to develop of spatial ability in collaboration with other researchers and to research the relationships academic success and spatial ability.

References

[1] Ault, H. K., John, S., Assessing and Enhancing Visualization Skills of Engineering Students in Africa: A Comprehensive Study, *Engineering Design Graphics Journal* Vol. 74(2) (2010), 12–20.

[2] Bosnyák, Á., Nagy-Kondor, R., The spatial ability and spatial geometrical knowledge of university students majored in mathematics, *Acta Didactica Universitatis Comenianae*, Vol. 8 (2008), 1–25.

[3] Branoff, T., Connolly, P., The Addition of Coordinate Axes to the Purdue Spatial Visualization Test – Visualization of Rotations: A Study at Two Universities, *Proceedings of the American Society for Engineering Education Annual Conference.*

[4] Budai, L., Improving Problem-Solving Skills with the Help of Plane-Space Analogies, *Center for Educational Policy Studies Journal*, Vol. 3(4) (2013), 79–98.

[5] CEEB Special Aptitude Test in Spatial Relations, *Developed by the College Entrance Examination Board, USA*, (1939).

[6] Chen, K. H., Validity studies of the Heinrich spatial visualization test, *Doctoral Dissertation, Ohio State University, Ohio, USA*, (1995).

[7] Ferguson, C., Ball, A., McDaniel, W., Anderson, R., A Comparison of Instructional Methods for Improving the Spatial Visualization Ability of Freshman Technology Seminar Students, *Proceedings, IAJC-IJME International Conference*, (2008).

[8] Field, B., A Course in Spatial Visualization, *Journal for Geometry and Graphics*, Vol. 3(2) (1999), 201–209.

[9] Gardner, H., Frames of mind: the theory of multiple intelligences, *Basic Books, New York*, (1983).

[10] Gorska, R., Sorby, S., Testing Instruments For The Assessment Of 3-D Spatial Skills, *Proceedings of the American Society for Engineering Education Annual Conference*, (2008).

[11] Guay, R. B., Purdue Spatial Visualisation Test: Rotations, *West Lafayette: Purdue Research Foundation*, (1977).

[12] Haanstra, F. H., Effects of art education on visual-spatial and aesthetic perception: two meta-analysis, *Rijksuniversiteit Groningen, Groningen*, (1994).

[13] Heinrich, V. L. S., The development and validation of a spatial perception test for selection purpopes, *Master Science Dissertation, Ohio State University, Columbus, Ohio, USA*, (1989).

[14] Hohenwarter, M., Lavicza, Z., Gaining momentum: GeoGebra inspires educators and students around the world, *GeoGebra The New Language For The Third Millennium, Zigotto Publishing House, Braila*, Vol. 1(1) (2010).

[15] Hohenwarter, M., Preiner, J., Dynamic mathematics with GeoGebra, *Journal of Online Mathematics and its Applications*, Vol. 7 (2007), ID 1448.

[16] Leopold, C., Gorska, R. A., Sorby, S. A., International Experiences in Developing the Spatial Visualization Abilities of Engineering Students, *Journal for Geometry and Graphics*, Vol. 5(1) (2001), 81—91.

[17] Linn, M. C., Petersen, A. C., Emergence and characterization of sex differences in spatial ability: a-meta analysis, *Child Development*, Vol. 56 (1985), 1479—1498.

[18] Maier, P. H., Spatial geometry and spatial ability – How to make solid geometry solid? *In: Elmar Cohors-Fresenborg, K. Reiss, G. Toener, H.-G. Weigand, editors, Selected Papers from the Annual Conference of Didactics of Mathematics 1996, Osnabrueck*, Vol. 56 (1998), 63—75.

[19] Manger, T., Eikeland, O.J., The Effects of Spatial Visualization and students' Sex on Mathematical Achivement, *British Journal of Psychology*, Vol. 89 (1998), 17–25.

[20] Martín-Gutiérrez, J., Gil, F. A., Contero, M., Saorín, J. L., Dynamic three-dimensional illustrator for teaching descriptive geometry and training visualisation skills, *Computer Applications in Engineering Education*, Vol. 21(1) (2013), 8–25.

[21] McGee, M. G., Human Spatial Abilities: Psychometric studies and environmental, genetic, hormonal and neurological influences, *Psychological Bulletin*, Vol. 86 (1979), 899—918.

[22] Nagy-Kondor, R., Spatial ability of engineering students, *Annales Mathematicae et Informaticae*, Vol. 34 (2007), 113–122.

[23] Nagy-Kondor, R., The results of a delayed test in Descriptive Geometry, *International Journal for Technology in Mathematics Education*, Vol. 15(3) (2008), 119–128.

[24] Nagy-Kondor, R.,Using dynamic geometry software at technical college, *Mathematics and Computer Education, Fall*, Vol. 42(3) (2008), 249–257.

[25] Nagy-Kondor, R., Technical Mathematics in the University of Debrecen, *Annales Mathematicae et Informaticae*, Vol. 38 (2011), 157–167.

[26] Nagy-Kondor, R., Importance of spatial visualization skills in Hungary and Turkey: Comparative Studies, *Annales Mathematicae et Informaticae*, Vol. 43 (2014), 171–181.

[27] Nagy-Kondor, R., Sörös, C., Engineering students' Spatial Abilities in Budapest and Debrecen, *Annales Mathematicae et Informaticae*, Vol. 40 (2012), 187–201.

[28] Nagy-Kondor, R., Szíki, G. Á., "Basic Knowledge of Natural Sciences": a new foundation subject at the Faculty of Engineering, University of Debrecen, *Horizons of mathematics, physics and computer sciences*, Vol. 41(2) (2012), ISSN 1335-4981, 9–17.

[29] Németh, B., Hoffmann, M., Gender differences in spatial visualization among engineering students, *Annales Mathematicae et Informaticae*, Vol. 33 (2006), 169—174.

[30] Németh, B., Sörös, C., Hoffmann, M., Typical mistakes in Mental Cutting Test and their consequences in gender differences, *Teaching Mathematics and Computer Science*, (2007), 385–392.

[31] Olkun, S., Making Connections: Improving Spatial Abilities with Engineering Drawing Activities, *International Journal of Mathematics Teaching and Learning, http://www.ex.uk/cimt/ijmt1/ijabout.htm.*

[32] Pietsch, S., Jansen, P., Different Mental Rotation Performance in Students of Music, Sport and Education, *Learning and Individual Differences*, Vol. 22 (2012), 159–163.

[33] Saito, T., Shiina, K., Makino, K., Suzuki, K., Jingu, T., Analysis of Problem Solving Process and Causes of Error in a Mental Cutting Test, *Proc. 2nd China-Japan Joint Conference on Graphics Education, Chengdu, China*, Vol. 22 (1995), 259–264.

[34] Scribner, S. A., Novice drafters' spatial visualization development: influence of instructional methods and individual learning styles, *Dissertation, Southern Illinois University, Carbondale*, (2004).

[35] Shea, D. L., Lubinski, D., Benbow, C. P., Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study, *Journal of Educational Psychology*, Vol. 93 (2001), 604-–614.

[36] Sorby, S., A New and Improved Course for Developing Spatial Visualization Skills, *Proceedings, ASEE Annual Conference*, (2001).

[37] Tsutsumi, E., Shiina, K., Suzaki, A., Yamanouchi, K., Takaaki, S., Suzuki, K., A Mental Cutting Test on female students using a stereographic system, *Journal for Geometry and Graphics*, Vol. 3 (1999), 111–119.

[38] Turgut, M., Nagy-Kondor, R., Comparison of Hungarian and Turkish prospective mathematics teachers' Mental Cutting performances, *Acta Didactica Universitatis Comenianae*, Vol. 13 (2013), 47–58.

[39] Turgut, M., Nagy-Kondor, R., Spatial Visualisation Skills of Hungarian and Turkish prospective mathematics teachers, *International Journal for Studies in Mathematics Education*, Vol. 6(1) (2013), 168–183.

[40] Turgut, M., Uygan, C., Spatial ability training via 3D modelling software, *In E. Faggiano, A. Montone (Eds) Proceedings of the 11th International Conference on Technology in Mathematics Teaching–ICTMT11, University of Bari, 9-12 July 2013, Italy: Università degli Studi di Bari Aldo Moro*, (2013), 292-–297.

[41] Turgut, M., Yenilmez, K., Spatial visualization abilities of preservice mathematics teachers, *Journal of Research in Education and Teaching*, Vol. 1(2) (2012), 243–252.

[42] Yilmaz, H. B., On the Development and Measurement of Spatial Ability, *International Electronic Journal of Elementary Education*, Vol. 1(2) (2009), 83–96.

[43] Yue, J., Spatial visualization by realistic 3D views, *Engineering Design Graphics Journal*, Vol. 72(1) (2009).

[44] Vorstenbosch, M. A., Klaassen, T. P., Donders, A. R. T., Kooloos, J. G., Bolhuis, S. M., Laan, R. F., Learning anatomy enhances spatial ability, *Anatomical sciences education*, Vol. 6(4) (2013), 257–262.

[45] Voyer, D., Voyer, S., Bryden, M. P., Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables, *Psychological Bulletin*, Vol. 117 (1995), 250–270.

Annales Mathematicae et Informaticae
46 *(2016) pp. 277–291*
http://ami.ektf.hu

The impacts of the introduction of the function concept on students' skills

Gyöngyi Szanyi

University of Debrecen
szanyi.gyongyi@science.unideb.hu

Submitted December 7, 2015 — Accepted September 7, 2016

Abstract

The concept of function is of fundamental importance to the learning of mathematics [7]. In the function concept development process the rule-followig and rule-recognition skills have an important role, that are necessary in the construction of value tables, which help children to figure out the relationship between quantities [4]. In this study, the skills mentioned above were tested among seventh grade students from Ukrainian and Hungarian schools, then consequences have been compared to results of previous studies [15]. We attempted to find out whether the introduction of this concept has an effect on the aforementioned skills.

Keywords: function concept, rule following and rule recognition skills, Hungarian and Ukrainian secondary education, comparison.

MSC: I23

1. Introduction

Over the past at least twenty years the concept of function has emerged as a unifying theme in mathematics curricula internationally. The notion of functions evolved from dependence relationships of real life phenomena to an abstract correspondence that is usually best describe in symbolic terms [13, 14].

In her study, Sierpinska [14] sets out the conditions of understanding the notion of function. These conditions illustrate that it takes time to reach a thorough understanding of the function concept. There is a long journey from the beginning to develop an understanding of the relationships between the elements of sets to the robust function concept. According to Kwari [10] the rule is an element of the function concept. Regarding the skills that can be linked to the function concept [10], possessing rule-recognition and rule-following skills (hereafter referred to as RR and RF) is exceptionally important in the period before providing the definition of function (preparation period) in order to be able to recognize function-like relations. However, studying this period in Hungarian [9] and Ukrainian mathematics curricula frameworks, we can state that the latter focuses less on developing skills discussed above [15].

Thus, the base of this paper is a previously conducted study among a group of 6^{th} grade students of a Hungarian school and a group of 6^{th} grade students of a Ukrainian school, where we tested their RF and RR skills before introducing them to the function concept. Taking the results of this study into consideration, the aim of this paper is to answer the following questions: (a) do the same students from Ukraine develop without targeted development in the aforementioned skills year by year; (b) what are the differences in the skills between the students of the two countries by the end of the 7^{th} form, after the introduction of the concept of function, does it have an impact on these skills.

2. Theoretical background

Mathematical definitions play an important role in the study of practically every area of mathematics ([5, 17, 18]).

The modern definition of function that frames this study is the Dirichlet-Bourbaki definition, which is "a correspondence between two nonempty sets that assigns to every element in the first set (the domain) exactly one element in the second set (the codomain)" (as cited in [6, p. 357]).

Vinner [20] drew attention to the distinction between the concept definition that mathematicians use to define a mathematical concept and the concept image which people generate in their mind. He also showed that most students use their own personal concept definition for the notion of function. Vinner and Dreyfus [6] categorized the students' definitions of a function into six categories, a refinement of the categorization by Vinner [20]: (A) correspondence (the Dirichlet-Bourbaki definition); (B) dependence relation (dependence between two variables); (C) rule (a function is a rule; a rule is expected to have some regularity, whereas a correspondence may be "arbitrary"); (D) operation (a function is an operation or manipulation); (E) formula (a function is a formula, an algebric expression, or an equation); and, (F) representation (graphical or symbolic representation) (as cited in [6, p. 360]).

Taking into account these categories, it can be highlighted that the function can be defined in various ways. Sierpinska [14] described the "worlds" that the study of functions should focus on: the world of changes or changing objects; the world of relationships or processes; and, the world of rules, patterns, and laws. According to Sierpinska [14] the difference between the rule and relationship is subtle because the rules, patterns and laws are simply well defined relationships. Relationships can be expressed verbally or using diagrams, tables, graphs or in symbols. A rule can be a verbal statement or a formula. It is possible for one to detect a relationship but fail to explicity state the rule. Finding rules, patterns and laws can be used as an entry point to the development of the function concept.

Among the skills that could be linked to the above listed "worlds", the RR and RF skills play a highly important role in recognising function-like relations, as well as in representing them in various ways during the concept-forming process. Because the concept of functions can be represented in a variety of ways (verbal, set diagram, function box, set of ordered pairs, table of values, graph, formula) [1] an important aspect of its understanding is the ability to use these multiple representations [11]. Thompson [16] claims that if students do not see something remain the same as they move among different representations then they see each representation as a topic to be learnt in isolation.

To summarize, the arguments above we can state that in order to form the function concept, it is crucial for the students to be able to: (1) recognize relations between cohesive elements if these relations are represented in different ways; (2) to express relations in various ways; (3) follow a rule, based on which elements are ordered to one another.

The aim of this study is to investigate whether the introduction of the function concept to 7^{th} grade students, after providing them with its definition and expanding its ways of different representations (e.g. arrow diagrams and graphs of functions are considered new ways of representations since the students of both groups are already aware of tables of values and formulas) have an impact on students' RF and RR skills.

3. Methodology

3.1. Sample

Participants were 24 seventh grade students (13-14 years old), with moderate abilities, in a school with Hungarian as the language of instruction in Ukraine and 20 students from the education system of Hungary (13-14 years old). When choosing our sample, we tried to balance between the two groups in a way that none of them are specialised classes in Mathematics, they study the subject in 4 hours per week and by the end of the sixth grade they acquire the same material. Based on their grades the students are on the same level of knowledge. In both schools they use the textbook supported by the Ministry of Education of the given country.

As the research was carried out in April, during the second semester of the seventh grade, students of both countries were already familiar with the natural numbers, fractions (common fractions and decimals), and had learned arithmetic operations with rational numbers, direct proportionality, equations. Also, the concept of function had been introduced to both groups in a more or less similar way and the level of acquisition had been tested before this research was conducted.

3.2. Background

In the Ukrainian and Hungarian (hereafter referred to as UA and HU) education system, function as a mathematical concept is defined at the seventh grade of the secondary school. In the lower classes, students are prepared with the use of different materials for the introduction of the function concept. Analysing the HU and UA curriculum and the textbooks for the fifth and sixth grade we can highlight that the major differences are in the requirements for developing RF and RR skills (in the lower classes in the UA education it does not exist at all) [15]. Prior researches (cf., studies cited above), however, suggest that they are necessary for the development of the function concept. Nonetheless, findings of previous studies prove that UA students participating in this research, at this stage of intellectual development (according to Piaget this is the transition from the concrete operations phase to the formal one) (as cited in [2, p. 48]), are able to recognise and follow a rule in order to assign cohesive elements without any targeted development by the end of 6^{th} grade, using only their previously acquired mathematical knowledge. However, argue in favour of a well-recognized rule either in a written form or with formulas remained a difficulty for them.

3.3. The questionnaire

A written test was used in order to investigate the RF and RR skills of students. Students worked independently and had 40 minutes to complete the test. The test contained six tasks that were based on the recognition and application of the relationship between the cohesive elements, as well as on the expression of the recognised rule, including as a formula. In other words, the base of these tasks was recognition or expression function-like connections between cohesive elements, represented by words, formulas, tables of values and graphs; without mentioning the word "function" in the instructions.

The cohesive element pairs did not clearly define a function, so more rules might be possible. In the instructions of the test, however, we tried to make it clear that we wanted students to find only one adequate rule. Some of the tasks were created by ourselves while the others were chosen from the literature. There were tasks that were included in the test from the previous year.

Task 1: Find a relationship between the x and y values of the columns and based on it, complete the table with the missing elements.

x	*16*	*12*	*4*	*-8*	*20*	*4*	*0*			
y	*4*	*3*	*1*					*6*	*-3*	*-1*

Write down the relationship: a) with words; b) with expression. Plot the pairs of values on a coordinate plane.

Task 1 targeted the recognition, application, expression of a relationship between cohesive elements using words and formula, and also the graphing of these ordered pairs on a coordinate plane. The filling in of the blank places of the tables assessed the following the rule. This table-filling task targeted the recognition and the application of a simple, one-step rule (e.g. determining y values based on x values and vice versa), furthermore, it provided us with the opportunity to study whether students were able to represent a rule that define the relationship between cohesive elements in various ways (such as by words, formulas, graphs).

Task 2: Find a relationship between the x and y values of the columns and based on it, complete the table with the missing elements.

x	1	10	7	0	9	20	38
y	5	23	17				

$y = \ldots$

Write down the relationship: a) with words; b) with expression.

In Task 2 it was necessary to recognise some kind of relationship between the cohesive elements (x and y), hint to record the recognised relationship in the language of arithmetic, and express the relationship with a formula, in other words, the aim of this task was to find out, whether students were able to define a rule that determined the relationship between the elements. The "end product" (y value) should be found with the help of the given "raw material" (x value) according to the recognised rule (as cited in [3, p. 23]), while in the previous task knowing the "end product" and using the recognised rule, the "raw material" should be found, in this way, Task 1 served as a preparation for the use of inverse function. Unlike the previous task, students had to recognise a multi-step rule and represent it in different ways.

Task 3: Fill in the table according to the following rule: $y = \frac{x}{3} - 7$. Write down the rule in words.

x	9	0	42	63	15	3
y						

Task 3 was aimed at the interpretation and following of a predefined rule, given by a formula. The correct completion of the table indicated correct interpretation of the symbolic rule.

Task 4: 2 litres/second of water flows from a tap to a tank. How much water is in the tank at:

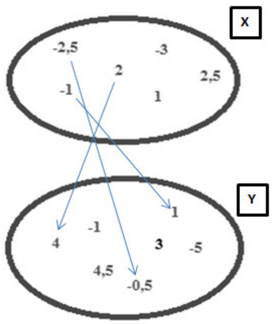
a) 1 s,
b) 2 s,
c) 5 s,
d) 10 s,
e) 16 s,
f) x s

later if the tank was empty at the beginning?

Illustrate the relationship between the amounts
I) in a table,
II) with a formula.
III) How much water is in the tank when, at the moment of opening the tap, the tank contains 3 litres of water? Use a formula to describe the connection between the values.

In Task 4, the rule was given verbally, in context. We take students' correct responses for parts (a) through (e) as an indication that the student had correctly interpreted the rule. Correct responses to exercise (f) indicated students' understanding of function (e.g. they could fill in the table with cohesive values based on a recognised rule since students were able to generalise the task, i.e. write down the rule using a formula). The aim of the third part of this task was to study if there is consistency between multi-step rulemaking and its description with a formula in cases where cohesive elements are provided verbally, in context and not in a table form (see Task 2).

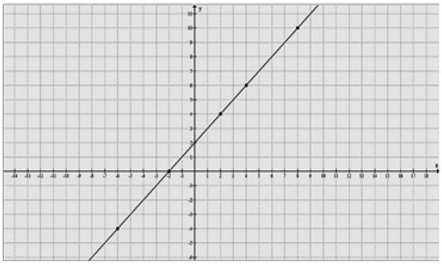
Task 5: In the figure below, several cohesive elements of two sets are connected with arrows. a) Find the assignment rule applicable. b) Connect the remaining cohesive elements using arrows.



In Task 5 cohesive elements were represented by arrows instead of a table of values. Students had to recognise the rule, based on which, the elements of set X were assigned to the elements of set Y. They were also required to argue in favour of the previously recognised rule using words. Part b) of this task was aimed at the application of the recognised rule.

The point of Task 6 was to find out whether students were able to transit from a mode of representation to another (convert data from a graph to a table) and whether they could read plotted points on a coordinate plane. The aim of part b) was to find out whether students could find a formula that describe the recognised rule (determination of assignment rules respresented by graph was not included in the HU and UA curricula frameworks before the beginning of seventh grade).

Task 6: Ordered pairs are graphed below.



a) Fill in the table with the pairs of values indicated with dots, and two additional ordered number pairs based upon your choice.

x								
y								

b) Write down the rule which recognized between the pairs of values with formula.

4. Results

4.1. Analysis of the students' answers

Task 1

18 UA students of 24 managed to fill in the table with correct values based upon the rule they recognised. 6 remaining students made the following mistake: they tried to apply different rules to each row of each column. The reason behind this could be that these students might find it problematic to represent or interpret a set of values in a table if these values are assigned to each other by a particular assignment rule. They might find it difficult even after introducing them to the concept of function. All these deficiencies were found despite the fact that they were able to read graphed points on a coordinate plane and to fill in the table with this set of ordered pairs.

10 UA students managed to argue in favour of a recognised rule using words, 9 of these students were able to use a formula for that. 14 students have found a rule that was not applicable to the table of values or just simply have skipped this part of the task. Hence, it is important to highlight that for these 14 students, the wording of thoughts and so describing a rule using words or formulas a problematic area.

17 students who found an applicable rule and followed it (e.g. filled in the table of values) could plot the points on a coordinate plane as well.

All of the 20 HU students managed to recognise an applicable rule. Following these rules, they could fill in the table of values correctly and also they could argue in favour of the recognised rule using formulas. 4 students either skipped the verbalization of the recognised rule or were inaccurate in representing the rule on a coordinate plane.

Several correct and similar answers of the UA and HU students:
"If we divide the x-values by 4, we get the y values."
"x divided by four equals y, or x is four times y."
"x is four times y."
"$x = 4y; y = \frac{x}{4}$"

Task 2

Only 7 of 24 UA students managed to complete Task 2. These students could answer every part of this question, i.e. they could fill in the table based on the rule they recognised when pairing numbers and also they could word it and represent it by formula. Therefore, it seems that those students, who despite possessing the ability of one-step RR and RF, might have difficulties in the process of two-step RR and RF.

11 of the 20 HU students managed to recognise a rule between the cohesive elements. Also they could describe it with the help of words and formula. The table of values reflected the rule as well. There was only 1 student who did not fill in the table, however managed to answer correctly part a) as well as part b) of the question and thus followed an appropriate thought-process. Results suggest that similarly to the UA participants, there are students who have difficulties with recognising and following a two-step rule. On the other hand, if they manage to recognise a rule, they are able to represent and describe it in different ways.

Task 3

Task 3 can be considered successful, since 23 of the 24 UA students could complete this task correctly: they could fill in the table of values which allowed us to assume that they could interpret the given rule correctly. However, similarly to the previous tasks, some difficulties arose when students tried to describe it, since only 9 students were able to describe the explicite given rule verbally. Presumably, these students have problems with expressing their thoughts verbally.

All of the HU students performed well in interpreting the rule then, based upon it, they could fill in the table of values. There was only 1 student who, instead of describing the assignment rule, presented the procedure of his calculation (e.g. *"we have to find the common denominator so we can get the final result").*

Task 4

18 UA students were able to make calculations with concrete values (a-e), however, they could not answer exercise f) correctly. 12 students managed to represent the values (a-f) in a table form but they could not describe the rule using a formula (Part II).

No more than 8 UA students could answer Part I and Part II of the task correctly. In addition, without knowing the elapsed time, they were able to generalise the rule, they could argue in favour of the rule using a formula. There were only 5 students who could solve Part III, i.e. write down the rule "$y = 2x + 3$" (letters used for variables may have varied). Those who could fill in the table with correct values based upon a rule they recognised and found an applicable formula as well (Task 2), were also included in this number. Similarly to the Task 2 where they had to deal with a set of related number pairs in a table, the majority of students found difficult with multi-step rulemaking in the case when they were provided with these values in an implicit way, embedded in text.

All of the 20 HU students could represent the connection between cohesive values in a table form. Similarly to the UA students trying to answer these exercises, also HU students could recognise a rule between related number pairs. In the case when the elapsed time was unknown (exercise f)), there were 9 students who either skipped this part of the question or their answers were incorrect. There were only 11 HU students, who were able to represent the ordered number pairs in a table form. However, when they had to represent the relationship (the recognised rule) by a formula, there were 19 students who could answer correctly. The various solutions of the two different, yet substantively identical questions (exercise f) and part II) suggest that it is problematic for the students to interpret the variable in a formula that indicates elapsed time in different ways. In other words, they have difficulties with interpreting the meaning of dependent and independent variables, in terms of the concept of function. There were 16 HU students who could solve Part III (wrote down the rule).

Task 5

In Task 5, a function was illustrated by arrow diagram; cohesive elements were indicated by arrows. Only 6 UA students of 24 could recognise the rule, based on which the elements of set X had been assigned to the elements of set Y. 5 students managed to phrase this assignment rule. This low ratio might be due to the "unpopularity" of teaching the function concept with the help of arrow diagrams. This fact is also supported by a research of Vinner [19] where he stated that a function, as taught in schools, is often identified with just one of its representations, either the symbolic or the graphical one. In this way, students are less capable of interpreting and understanding this type of representations of cohesive values. Most of the students attempted to solve this task but later on, they overlooked the cohesive elements given as an example and they assigned the remaining elements to each other using a different rule (e.g. they paired each non-negative value with their negative equivalents).

HU students managed to accomplish this task more effectively than UA students, although low success ratio occurred in here as well. 10 students recognised an applicable rule and managed to phrase it. There were 13 students who, in spite of the fact that they did not put the recognised assignment rule into words, the way they paired the remaining elements reflected correct judgements.

Overall, we can conclude that recognizing a connection between cohesive elements means difficulty for the students of both countries if these elements are represented by arrow diagram (which is a representation form of function).

Task 6

Reading plotted points on a coordinate plane required little effort from the students of both countries. All of the 20 HU and 22 UA students managed to fill in the table with this set of ordered number pairs and even could find two additional number pairs. It should be highlighted that most of the students could see none of these additional number pairs on the coordinate plane, since these pairs had not been graphed in the exercise. 8 UA and 16 HU students managed to describe a rule using formula; thus it can be stated that UA students can easily interpret a rule if it is described by a formula (see Task 3). On the other hand, as results of the previous tasks suggest (Task 1 and Task 2 for instance, where the function was represented by a table) these students have difficulties with describing a rule using symbols.

Analysis of the results of certain parts of the tasks

Figure 1 represents an analysis of the success ratio of certain parts of tasks that were aimed at assessing one-step RR and RF. We tried to answer the following questions: whether students will recognise and then follow a rule if cohesive elements are represented by different ways (by a table of values, by words, by graph or by using arrow diagram). We can see the results in the figure. The value of accuracy of arguing in favour of the recognised rule using words or formulas is not included (except for in the case of Task 6).

Comments:

Results of the following tasks were included when creating the diagram: Task 1 (filling in a table based on a recognised rule); Task 3 (filling in a table based on a predefined rule); Task 4, a)-e) (recognizing and following a rule embedded in text); Task 5, part b) (recognizing and following a rule illustrated by arrows, demonstrating how the function behaves by drawing missing arrows); Task 6, part b) (recognising an assignment rule that given by graph), since it reflects if students had recognised an assignment rule.

Clearly, the most problematic field for participants of both countries is recognizing and following a rule, based on which cohesive elements of sets are connected by arrows.

Figure 2 represents the success ratio of tasks that were aimed at assessing if students could recognise a multi-step rule and describe it with the use of formula (Task 2 and Task 4, part III).

Obviously, recognising and describing a multi-step rule with formulas proved to be a more difficult task than what we could experience in the case of a one-step rule. As seen in the Figure 2, recognizing and describing a multi-step rule using formula requires much more effort from UA students. Compering this to the results of one-step RR, the performance of both groups proved to be weaker in the case of multi-step RR and RF.

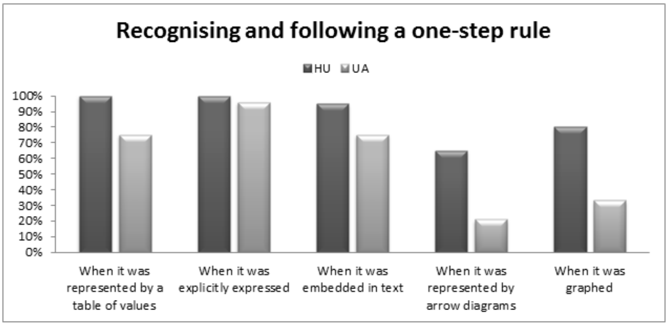


Figure 1

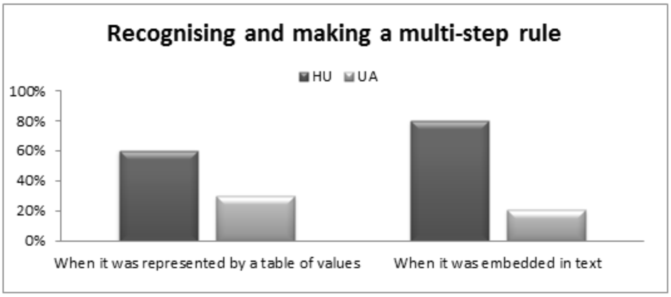


Figure 2

4.2. Do the students from Ukraine develop without targeted development in the RR and RF skills year by year?

As analysing the RF and RR skills of the same UA students who also had participated in the previous research (in 2014) it turned out that they were able to identify a rule between cohesive elements, however there were only 10 of 26 students who could desribe either a one-step or a multi-step rule using words, and there were only 2 who could argue in favour of these rules using formulas [15]. The aim of this study was to attempt to answer if the RR and RF skills of investigated UA students had developed after introducing them to the concept of function and its different ways of represantations.

The figures below represent the success ratio of UA students attempting to answer questions from the present and last year's assessment. These questions were either identical (Task 2 and Task 4, part I and II) or despite differing only in data sets and were the same types of exercises.

Figure 3 represents the success ratio of task from 2014 and 2015 that were used to assess the effectiveness of recognising and following a rule represented in different ways. The numbers in the brackets indicates the tasks which are used in current and previous investigations.

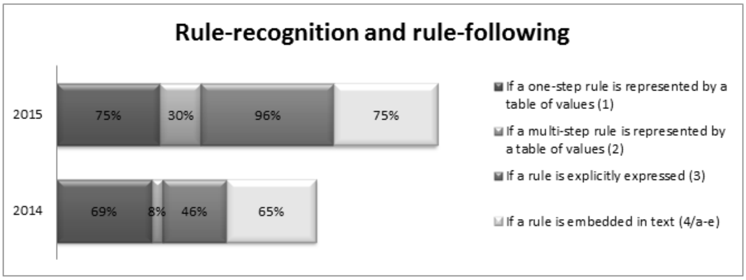


Figure 3

As it can be seen, greater development can be observed in the annual follow-up among investigated UA students, mainly in recognizing and interpreting a multi-step rule, also in following a rule that was explicitly expressed. Thus we can state that introducing the concept of function and its representations had an impact on the development of the students' RR and RF skills .

As it is represented in Figure 4, we could observe no development in arguing in favour of a rule using words; results are relatively stagnant. This allows us to come to the conclusion that the investigated UA students have difficulties with expressing their thoughts verbally. Many of the students, similarly to the previous research in 2014, skipped this part of the tasks, they did not even try to solve it.

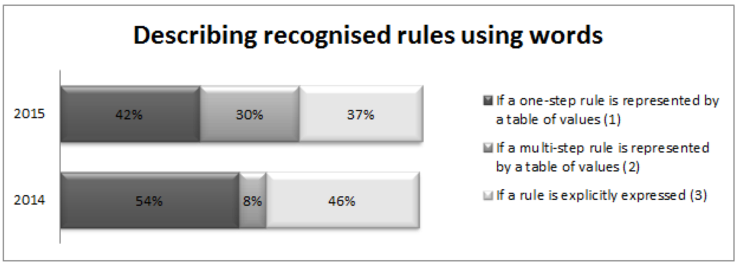


Figure 4

In describing the recognised rule using symbols some minor improvement occured as well (see Figure 5). As it can be seen, those students who were able to recognise a multi-step rule, could easily describe it using formula.

5. Conclusions

The focus of this study was to obtain information on the RR and RF skills of students in a 7^{th} grade of Hungarian and in a 7^{th} grade of Ukrainian schools,

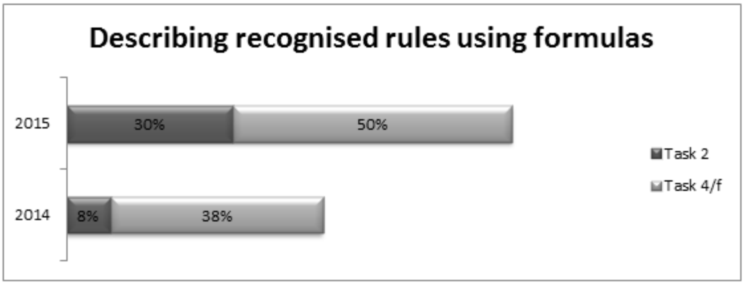


Figure 5

after introducing the concept of function, comparing the consequences to results of previous studies [15].

Considering the results of this current study, the following changes and deficiencies of the annual follow-up can be documented regarding the RF and RR skills of the students:

- Compared to previous studies, the majority of the 24 UA students could recognise and follow a one-step rule represented in different ways (in table form, by words, by graph).
- We could experience development in interpreting and following an explicitly expressed rule (see Figure 3).

Probably, that in the case of UA students, during the preparation process of the function concept, less attention was paid on developing RR and RF skills (before 7^{th} grade), the success ratio of the tasks done by UA students approached, in some cases reached (Task 1/representing, Task 3/a, Task 4/f, Task 6) the success ratio of the tasks done by HU students, whose learning/teaching process was more effectively assembled for improving RR and RF skills.

- Describing rules using words and/or symbols remained a slightly problematic area for the investigated UA students, similarly to last year.

This means, they have difficulties not only with wording their thoughts, but with the field of symbolism as well. Considering the representational complexity of the function concept [1], we can state that regarding the various representation forms of function, describing a function using either words (linguistic complexity) or formulas (mathematical complexity) requires a lot of effort. We could observe some minor development in recognising, following and describing a multi-step rule, but it is not significant (see Figure 3–5).

- Both of the investigated student groups experienced difficulties with recognising a rule in cases when it was represented by arrows (Task 5).

The recognition of functions represented in this way is problematic even for high school students who can define a function correctly [8]. Norman [12] explains this phenomenon with many mathematics teaching methods often favouring functions represented by graphs. This statement can be confirmed by what we have observed when students tried to convert different types of representations. Both UA and HU students needed little effort to convert data from a table to a graph and vice versa. Describing the function using words and/or symbols remained a difficulty for most of the UA student.

Considering the results it also should be highlighted that those UA and HU students who could describe a rule using words and symbols, demonstrated a similar way of thinking.

In conclusion, we can state that introducing the concept of function has an impact on the development of the RR and RF skills. However, results suggest that developing these skills would be useful to start even before introducing the function concept, since HU students performed better in this assessment, similarly to the previous year's.

Studying the characteristics of concept images of the investigated students in each country may be the subject of further researches.

References

[1] Akkoç, H., Tall, D.O., The simplicity, complexity and complication of the function concept, *Proceedings of the 26th International Conference on the Psychology of Mathematics Education* Vol. 2 (2002), Norwich, UK, 25–32.

[2] Ambrus, A., Bevezetés a matematikadidaktikába, ELTE, Eötvös Kiadó, Budapest (1995).

[3] Anderson, J., The mathematics curriculum: algebra, Blackie, London (1978).

[4] Blanton, M., Kaput, J., Functional thinking as a route into algebra in the elementary grades, In: Cai, J., Knuth, E. (eds.), Early algebraization (2011), Springer, Berlin, 5–23.

[5] de Villiers, M., Govender, R., Patterson, N., Defining in geometry, In: Craine, T.V., Rubenstein, R. (eds.), Understanding geometry for a changing world (2009), VA: National Council of Teachers of Mathematics, Reston, 189–203.

[6] Dreyfus, T., Vinner, S., Images and definitios for the concept of function, *Journal for Research in Mathematics Education* Vol. 20 (1989), 356–366.

[7] Eisenberg, T., On the development of a sense for functions, In: Harel, G., Dubinsky, E. (eds.), The concept of function: Aspects of epistemology and pedagogy (1992), DC: Mathematical Association of America, Washington, 153–174.

[8] Elia, I., Panaoura, A., Eracleous, A., Gagatsis, A., Relations between secondary pupils' conceptions about functions and problem solving in different representations, *International Journal of Science and Mathematics Education* Vol. 5 (2007), 533–556.

[9] Herendiné Kónya, E., A matematika tanítása alsó tagozaton, Nemzedékek Tudása Tankönyvkiadó, Budapest (2013).

[10] Kwari, R., An investigation into the development of the function concept through a problem-centered approach by form 1 pupils in Zimbabwe (Master's thesis) (2007).

[11] Lin, F., Cooney, T., Making sense of mathematics teacher education, Kluwer, The Netherlands (2001).

[12] Norman, A., Teachers' mathematical knowledge of the concept of function, In: Harel, G., Dubinsky, E. (eds.), The concept of function: Aspects of epistemology and pedagogy (1992), DC: Mathematical Association of America, Washington, 215–232.

[13] Ponte, J.P., The history of the concept of function and some educational implications, *The Mathematics Educator* Vol. 3(2) (1992), 3–7.

[14] Sierpinska, A., On understanding the notion of function, In: Harel, G., Dubinsky, E. (eds.), The concept of function: Aspects of epistemology and pedagogy (1992), DC: Mathematical Association of America, Washington, 25–58.

[15] Szanyi, Gy., The investigation of students' skills in the process of function concept creation, *Teaching of Mathematics and Computer Science* Vol. 13(2) (2015), 249–266.

[16] Thompson, P.W., Students, functions, and the undergraduate curriculum, In: Dubinsky, E., Schoenfeld, A., Kaput, J. (eds.), Research in Collegiate Mathematics Education, I, CBMS Issues in Mathematics Education, Vol. 4 (1994), 21–44.

[17] Usiskin, Z., Griffin, J., The classification of quadrilaterals: A study of definition, NC: Information Age Publishing, Charlotte (2008).

[18] Vinner, S., The role of definitions in the teaching and learning of mathematics, In: Tall, D. (ed.), Advanced mathematical thinking, Kluwer Academic Publishing (1991), Dordrecht, The Netherlands , 65–81.

[19] Vinner, S., The function concept as a prototype for problems in mathematicslearning, In: Harel, G., Dubinsky, E. (eds.), The concept of function: Aspects of epistemology and pedagogy (1992), DC: Mathematical Association of America, Washington, 195–214.

[20] Vinner, S., Concept definition, concept image and the notion of function, *International Journal of Mathematical Education in Science and Technology* Vol. 14 (1983), 293–305.