Journal of Integer Sequences, Vol. 8 (2005), Article 05.2.8

# On the Density of Languages Representing Finite Set Partitions ${ }^{1}$ 

Nelma Moreira and Rogério Reis<br>DCC-FC \& LIACC<br>Universidade do Porto<br>R. do Campo Alegre 823<br>4150 Porto<br>Portugal<br>nam@ncc.up.pt<br>rvr@ncc.up.pt


#### Abstract

We present a family of regular languages representing partitions of a set of $n$ elements in less or equal $c$ parts. The density of those languages is given by partial sums of Stirling numbers of second kind for which we obtain explicit formulas. We also determine the limit frequency of those languages. This work was motivated by computational representations of the configurations of some numerical games.


## 1 The languages $L_{c}$

Consider a game where natural numbers are to be placed, by increasing order, in a fixed number of columns, subject to some specific constraints. In these games column order is irrelevant. Numbering the columns, game configurations can be seen as sequences of column numbers where the successive integers are placed. For instance, the string

11213
stands for a configuration where $1,2,4$ were placed in the first column, 3 was placed in the second and 5 was placed in the third. Because column order is irrelevant, and to have a unique representation for each configuration, it is not allowed to place an integer in the

[^0]$k$ th column if the $(k-1)$ th is still empty, for any $k>1$. Blanchard and al. [BHR04] and Reis and al. [RMP04] used this kind of representation to study the possible configurations of sum-free games.

Given $c$ columns, let $\mathbb{N}_{c}=\{1, \ldots, c\}$. We are interested in studying the set of game configurations as strings in $\left(\mathbb{N}_{c}\right)^{\star}$, i.e., in the set of finite sequences of elements of $\mathbb{N}_{c}$. Game configurations can be characterised by the following language $L_{c} \subset\left(\mathbb{N}_{c}\right)^{\star}$ :

$$
L_{c}=\left\{a_{1} a_{2} \cdots a_{k} \in\left(\mathbb{N}_{c}\right)^{\star} \mid \forall i \in \mathbb{N}_{k}, a_{i} \leq \max \left\{a_{1}, \ldots, a_{i-1}\right\}+1\right\} .
$$

For $c=4$, there are only 15 strings in $L_{4}$ of length 4, instead of the total possible 256 in $\left(\mathbb{N}_{4}\right)^{4}$ :

| 1111 | 1112 | 1121 | 1122 | 1123 |
| :--- | :--- | :--- | :--- | :--- |
| 1211 | 1212 | 1213 | 1221 | 1222 |
| 1223 | 1231 | 1232 | 1233 | 1234 |

Given a finite set $\Sigma$, a regular expression (r.e.) $\alpha$ over $\Sigma$ represents a (regular) language $L(\alpha) \subseteq \Sigma^{\star}$ and is inductively defined by: $\emptyset$ is a r.e and $L(\emptyset)=\emptyset ; \epsilon$ (empty string) is a r.e and $L(\epsilon)=\{\epsilon\} ; a \in \Sigma$ is a r.e and $L(a)=\{a\} ;$ if $\alpha_{1}$ and $\alpha_{2}$ are r.e., $\alpha_{1}+\alpha_{2}, \alpha_{1} \alpha_{2}$ and $\alpha_{1}^{\star}$ are r.e., respectively with $L\left(\alpha_{1}+\alpha_{2}\right)=L\left(\alpha_{1}\right) \cup L\left(\alpha_{2}\right), L\left(\alpha_{1} \alpha_{2}\right)=L\left(\alpha_{1}\right) L\left(\alpha_{2}\right)$ and $L\left(\alpha_{1}{ }^{\star}\right)=L\left(\alpha_{1}\right)^{\star}$, where we assume the usual precedence of the operators (see [HMU00]). A regular expression $\alpha$ is unambiguous if for each $w \in L(\alpha)$ there is only one path through $\alpha$ that matches $w$.
Theorem 1.1. For all $c \geq 1, L_{c}$ is a regular language.
Proof. For $c=1$, we have $L_{1}=L\left(11^{\star}\right)$. We define by induction on $c$, a family of regular expressions:

$$
\begin{align*}
& \alpha_{1}=11^{\star}  \tag{1}\\
& \alpha_{c}=\alpha_{c-1}+\prod_{j=1}^{c} j(1+\cdots+j)^{\star} \tag{2}
\end{align*}
$$

It is trivial to see that

$$
\begin{equation*}
\alpha_{c}=\sum_{i=1}^{c} \prod_{j=1}^{i} j(1+\cdots+j)^{\star} . \tag{3}
\end{equation*}
$$

For instance, $\alpha_{4}$ is

$$
11^{\star}+11^{\star} 2(1+2)^{\star}+11^{\star} 2(1+2)^{\star} 3(1+2+3)^{\star}+11^{\star} 2(1+2)^{\star} 3(1+2+3)^{\star} 4(1+2+3+4)^{\star} .
$$

It is also obvious that $L\left(\alpha_{c-1}\right) \subseteq L\left(\alpha_{c}\right)$, for $c>1$. For any $c \geq 1$, we prove that

$$
L_{c}=L\left(\alpha_{c}\right) .
$$

$L_{c} \supseteq L\left(\alpha_{c}\right):$ If $x \in L\left(\alpha_{c}\right)$ it is obvious that $x \in L_{c}$.
$L_{c} \subseteq L\left(\alpha_{c}\right):$ By induction on the length of $x \in L_{c}$ : If $|x|=1$ then $x \in L\left(\alpha_{1}\right) \subseteq L\left(\alpha_{c}\right)$. Suppose that for any string $x$ of length $\leq n, x \in L\left(\alpha_{c}\right)$. Let $|x|=n+1$ and $x=y a$, where $a \in \mathbb{N}_{c}$ and $y \in L\left(\alpha_{c}\right)$. Let $c^{\prime}=\max \left\{a_{i} \mid a_{i} \in y\right\}$. If $c^{\prime}=c$, obviously $x \in L\left(\alpha_{c}\right)$. If $c^{\prime}<c$, then $y \in L\left(\alpha_{c^{\prime}}\right)$, and $x \in L\left(\alpha_{c^{\prime}+1}\right) \subseteq L\left(\alpha_{c}\right)$.

## 2 Counting the strings of $L_{c}$

The density of a language L over a finite set $\Sigma, \rho_{L}(n)$, is the number of strings of length $n$ that are in $L$, i.e.,

$$
\rho_{L}(n)=\left|L \cap \Sigma^{n}\right| .
$$

In particular, the density of $L_{c}$ is

$$
\rho_{L_{c}}(n)=\left|L_{c} \cap \mathbb{N}_{c}^{n}\right| .
$$

Using generating functions we can determine a closed form for $\rho_{L_{c}}(n)$. Recall that, a (ordinary) generating function for a sequence $\left\{a_{n}\right\}$ is a formal series (see [GKP94])

$$
G(z)=\sum_{i=0}^{\infty} a_{n} z^{n}
$$

If $A(z)$ and $B(z)$ are generating functions for the density functions of the languages represented by unambiguous regular expressions $A$ and $B$, and $A+B, A B$ and $A^{\star}$ are also unambiguous r.e., we have that $A(z)+B(z), A(z) B(z)$ and $\frac{1}{1-A(z)}$, are the generating functions for the density functions of the corresponding languages (see [SF96], page 378).

As $\alpha_{c}$ are unambiguous regular expressions, from (3), we obtain the following generating function for $\left\{\rho_{L_{c}}(n)\right\}$ :

$$
T_{c}(z)=\sum_{i=1}^{c} \prod_{j=1}^{i} \frac{z}{(1-j z)}=\sum_{i=1}^{c} \frac{z^{i}}{\prod_{j=1}^{i}(1-j z)} .
$$

Notice that

$$
S_{i}(z)=\frac{z^{i}}{\prod_{j=1}^{i}(1-j z)}
$$

are the generating functions for the Stirling numbers of second kind

$$
S(n, i)=\frac{1}{i!} \sum_{j=0}^{i-1}(-1)^{j}\binom{i}{j}(i-j)^{n}
$$

which are, for each $n$, the number of ways of partitioning a set of $n$ elements into $i$ nonempty sets (see [GKP94] and A008277).

Then, a closed form for the density of $L_{c}, \rho_{L_{c}}(n)$, is given by

$$
\begin{equation*}
\rho_{L_{c}}(n)=\sum_{i=1}^{c} S(n, i) \tag{4}
\end{equation*}
$$

i.e., a partial sum of Stirling numbers of second kind.

In Table 1 we present the values of $\rho_{L_{c}}(n)$, for $c=1 . .8$ and $n=1 . .13$. For some sequences, we also indicate the corresponding number in Sloane's On-Line Encyclopedia
of Integer Sequences [Slo03]. The closed forms were calculated using the Maple computer algebra system [Hec03].

From expression (4), it is also easy to see that

## Theorem 2.1.

$$
\lim _{c \rightarrow \infty} \rho_{L_{c}}(n)=B_{n},
$$

where $B_{n}$ are the Bell numbers, i.e., for each $n$, the number of ways a set of $n$ elements can be partitioned into nonempty subsets.
Proof. Bell numbers, $B_{n}$, can be defined by the sum

$$
B_{n}=\sum_{i=1}^{n} S(n, i)
$$

And, as $S(n, i)=0$ for $i>n$, we have

$$
\lim _{c \rightarrow \infty} \rho_{L_{c}}(n)=\sum_{i=1}^{n} S(n, i)+\lim _{c \rightarrow \infty} \sum_{i=n+1}^{c} S(n, i)=B_{n}
$$

In Table 1, for each $c \geq 1$, the subsequence for $n \leq c$ coincides with the first $c$ elements of $B_{n}$ (A000110).

Moreover, we can express $\rho_{L_{c}}(n)$, and then the partial sums of Stirling numbers of second kind, as a generic linear combination of $n$th powers of $k$, for $k \in \mathbb{N}_{c}$. Let $S^{j}(n, i)$ denote the $j$ th term in the summation of a Stirling number $S(n, i)$, i.e.,

$$
S^{j}(n, i)=\frac{1}{i!}(-1)^{j}\binom{i}{j}(i-j)^{n} .
$$

Lemma 2.1. For all $n$ and $0 \leq i \leq n$,

$$
\begin{equation*}
S^{0}(n, i)=-S^{1}(n, i+1) \tag{5}
\end{equation*}
$$

Proof.

$$
S^{1}(n, i+1)=\frac{1}{(i+1)!}(-1)\binom{i+1}{1} i^{n}=(-1) \frac{1}{i!} i^{n}=-S^{0}(n, i)
$$

Applying (5) in the summation (4) of $\rho_{L_{c}}(n)$, each term $S(n, i)$ simplifies the subterm $S^{1}(n, i)$ with the subterm $S^{0}(n, i-1)$, for $i \geq 2$. We obtain

$$
\begin{aligned}
\rho_{L_{1}}(n) & =S^{0}(n, 1), \\
\rho_{L_{2}}(n) & =S^{0}(n, 2), \\
\rho_{L_{c}}(n) & =S^{0}(n, c)+\sum_{i=3}^{c} \sum_{j=2}^{i-1} S^{j}(n, i), \quad \text { for } c>2 ; \\
& =\frac{c^{n}}{c!}+\sum_{i=3}^{c} \sum_{j=2}^{i-1} S^{j}(n, i), \quad \text { for } c>2 .
\end{aligned}
$$

| c | $\rho_{L_{c}}(n)$ | OEIS |
| :---: | :---: | :---: |
| 1 |  |  |
| 2 | $\begin{aligned} & \frac{1}{2} 2^{n} \\ & 1,2,4,8,16,32,64,128,256,512,1024,2048,4096,8192, \ldots \end{aligned}$ |  |
| 3 | $\begin{aligned} & \frac{1}{6} 3^{n}+\frac{1}{2} \\ & 1,2,5,14,41,122,365,1094,3281,9842,29525,88574,265721, \ldots \end{aligned}$ | $\underline{\text { A007051 }}$ |
| 4 | $\begin{aligned} & \frac{1}{24} 4^{n}+\frac{1}{4} 2^{n}+\frac{1}{3} \\ & 1,2,5,15,51,187,715,2795,11051,43947,175275,700075,2798251, \ldots \end{aligned}$ | $\underline{\text { A007581 }}$ |
| 5 | $\begin{aligned} & \frac{1}{120} 5^{n}+\frac{1}{12} 3^{n}+\frac{1}{6} 2^{n}+\frac{3}{8} \\ & 1,2,5,15,52,202,855,3845,18002,86472,422005,2079475,10306752, \ldots \end{aligned}$ | $\underline{\text { A056272 }}$ |
| 6 | $\begin{aligned} & \frac{1}{720} 6^{n}+\frac{1}{48} 4^{n}+\frac{1}{18} 3^{n}+\frac{3}{16} 2^{n}+\frac{11}{30} \\ & 1,2,5,15,52,203,876,4111,20648,109299,601492,3403127,19628064, \ldots \end{aligned}$ | $\underline{\text { A056273 }}$ |
| 7 | $\begin{aligned} & \frac{1}{5040} 7^{n}+\frac{1}{240} 5^{n}+\frac{1}{72} 4^{n}+\frac{1}{16} 3^{n}+\frac{11}{60} 2^{n}+\frac{53}{144} \\ & 1,2,5,15,52,203,877,4139,21110,115179,665479,4030523,25343488, \ldots \end{aligned}$ | $\underline{\text { A099262 }}$ |
| 8 | $\begin{aligned} & \frac{1}{40320} 8^{n}+\frac{1}{1440} 6^{n}+\frac{1}{360} 5^{n}+\frac{1}{64} 4^{n}+\frac{11}{180} 3^{n}+\frac{53}{288} 2^{n}+\frac{103}{280} \\ & 1,2,5,15,52,203,877,4140,21146,115929,677359,4189550,27243100, \ldots \end{aligned}$ | A099263 |

Table 1: Density functions of $L_{c}$, for $c=1 . .8$.

If the sums are rearranged such that $i=k+j$, we have

$$
\begin{equation*}
\rho_{L_{c}}(n)=\frac{c^{n}}{c!}+\sum_{k=1}^{c-2} \sum_{j=2}^{c-k} S^{j}(n, k+j) \tag{6}
\end{equation*}
$$

where

$$
\begin{align*}
S^{j}(n, k+j) & =\frac{k^{n}}{(k+j)!}(-1)^{j}\binom{k+j}{j}  \tag{7}\\
& =\frac{k^{n}}{k!j!}(-1)^{j} . \tag{8}
\end{align*}
$$

Replacing (8) into equation (6) we get

$$
\begin{equation*}
\rho_{L_{c}}(n)=\frac{c^{n}}{c!}+\sum_{k=1}^{c-2} \frac{k^{n}}{k!}\left(\sum_{j=2}^{c-k} \frac{(-1)^{j}}{j!}\right) . \tag{9}
\end{equation*}
$$

In equation (9), the coefficients of $k^{n}, 1 \leq k \leq c$, can be calculated using the following recurrence relation:

$$
\begin{aligned}
\gamma_{1}^{1} & =1 \\
\gamma_{1}^{c} & =\gamma_{1}^{c-1}+\frac{(-1)^{c-1}}{(c-1)!}, \text { for } c>1 \\
\gamma_{k}^{c} & =\frac{\gamma_{k-1}^{c-1}}{k}, \text { for } c>1 \text { and } 2 \leq k \leq c .
\end{aligned}
$$

And, we have
Theorem 2.2. For all $c \geq 1$,

$$
\begin{equation*}
\rho_{L_{c}}(n)=\sum_{k=1}^{c} \gamma_{k}^{c} k^{n} . \tag{10}
\end{equation*}
$$

From the expression (10), the closed forms in Table 1 are easily derived.
Finally, we can obtain the limit frequency of $L_{c}$ in $\left(\mathbb{N}_{c}\right)^{\star}$. Since

$$
\rho_{\left(\mathbb{N}_{c}\right)^{\star}}(n)=c^{n}
$$

and $\lim _{n \rightarrow \infty}\left(\frac{k}{c}\right)^{n}=0$, for $1 \leq k \leq c-2$, we have
Theorem 2.3. For all $c \geq 1$,

$$
\lim _{n \rightarrow \infty} \frac{\rho_{L_{c}}(n)}{\rho_{\left(\mathbb{N}_{c}\right)^{*}}(n)}=\frac{1}{c!} .
$$

## 3 A bijection between strings of $L_{c}$ and partitions of finite sets

The connection between the density of $L_{c}$ and Stirling numbers of second kind is not accidental. Each string of $L_{c}$ with length $n$ corresponds to a partition of $\mathbb{N}_{n}$ with no more than $c$ parts. This correspondence can be made explicit as follows.

Let $a_{1} a_{2} \cdots a_{n}$ be a string of $L_{c}$. This string corresponds to the partition $\left\{A_{j}\right\}_{j \in \mathbb{N}_{c^{\prime}}}$ of $\mathbb{N}_{n}$ with $c^{\prime}=\max \left\{a_{1}, \ldots, a_{n}\right\}$, such that for each $i \in \mathbb{N}_{n}, i \in A_{a_{i}}$. For example, the string 1123 corresponds to the partition $\{\{1,2\},\{3\},\{4\}\}$ of $\mathbb{N}_{4}$ into 3 parts.

This defines a bijection. That each string corresponds to a unique partition is obvious. Given a partition $\left\{A_{j}\right\}_{j \in \mathbb{N}_{c^{\prime}}}$ of $\mathbb{N}_{n}$ with $c^{\prime} \leq c$, we can construct the string $b_{1} \cdots b_{n}$, such that for $i \in \mathbb{N}_{n}, b_{i}=j$ if $i \in A_{j}$. For the partition $\{\{1,2\},\{3\},\{4\}\}$, we obtain 1123.

## 4 Counting the strings of $L_{c}$ of length equal or less than a certain value

Although strings of $L_{c}$ of arbitrary length represent game configurations, for computational reasons ${ }^{2}$ we consider all game configurations with the same length, padding with zeros the positions of integers not yet in one of the $c$ columns. In this way, we obtain the languages $L_{c}^{0}=L_{c}\left\{0^{\star}\right\}$. So, determining the number of strings of length equal or less than $n$ that are in $L_{c}$ is tantamount to determining the density of $L_{c}^{0}$, i.e.,

$$
\rho_{L_{c}^{0}}(n)=\left|L_{c}^{0} \cap\left(\{0\} \cup \mathbb{N}_{c}\right)^{n}\right| .
$$

As seen in Section 2, and because $L_{c}\left\{0^{\star}\right\}=L\left(\alpha_{c} 0^{\star}\right)$, the generating function $T_{c}^{\prime}(z)$ of $\rho_{L_{c}^{0}}(n)$ can be obtained as the product of a generating function for $\rho_{L_{c}}(n), T_{c}(z)$, by a generating function for $\rho_{\{0\}^{\star}}(n)$, e.g., $\frac{1}{1-z}$. Thus, the generating function for $\left\{\rho_{L_{c}^{0}}(n)\right\}$ is

$$
T_{c}^{\prime}(z)=T_{c}(z) \frac{1}{1-z}=\sum_{i=1}^{c} \frac{z^{i}}{(1-z) \prod_{j=1}^{i}(1-j z)}
$$

and a closed form for $\rho_{L_{c}^{0}}(n)$ is (as expected)

$$
\begin{equation*}
\rho_{L_{c}^{0}}(n)=\sum_{m=1}^{n} \sum_{i=1}^{c} S(m, i), \tag{11}
\end{equation*}
$$

where $m$ starts at 1 because $S(m, i)=0$, for $i>m$.

[^1]Using expression (9) in (11) we have

$$
\begin{aligned}
\rho_{L_{1}^{0}}(n) & =n ; \\
\rho_{L_{2}^{0}}(n) & =2^{n}-1 ; \\
\rho_{L_{c}^{0}}(n) & =\sum_{m=1}^{n}\left(\frac{c^{m}}{c!}+\sum_{k=1}^{c-2} \frac{k^{m}}{k!} \sum_{j=2}^{c-k} \frac{(-1)^{j}}{j!}\right), \quad \text { for } c>2 ; \\
& =\frac{c^{n+1}-c}{(c-1) c!}+n \sum_{j=2}^{c-1} \frac{(-1)^{j}}{j!}+\sum_{k=2}^{c-2} \frac{k^{n+1}-k}{(k-1) k!}\left(\sum_{j=2}^{c-k} \frac{(-1)^{j}}{j!}\right) \\
& =\frac{c^{n}-1}{(c-1)(c-1)!}+n \sum_{j=2}^{c-1} \frac{(-1)^{j}}{j!}+\sum_{k=2}^{c-2} \frac{k^{n}-1}{(k-1)(k-1)!}\left(\sum_{j=2}^{c-k} \frac{(-1)^{j}}{j!}\right) .
\end{aligned}
$$

If we use the equation (10), we have
Theorem 4.1. For all $c \geq 1$,

$$
\rho_{L_{c}^{0}}(n)=n \gamma_{1}^{c}+\sum_{k=2}^{c} \frac{\gamma_{k}^{c}\left(k^{n+1}-k\right)}{k-1} .
$$

Proof.

$$
\rho_{L_{c}^{0}}(n)=\sum_{m=1}^{n} \sum_{k=1}^{c} \gamma_{k}^{c} k^{m}=\sum_{k=1}^{c} \gamma_{k}^{c} \sum_{m=1}^{n} k^{m}=n \gamma_{1}^{c}+\sum_{k=2}^{c} \frac{\gamma_{k}^{c}\left(k^{n+1}-k\right)}{k-1} .
$$

In the Table 2 we present the values of $\rho_{L_{c}^{0}}(n)$, for $c=1 . .8$ and $n=1 . .13$. As before, the limiting sequence as $c \rightarrow \infty$ is the sequence of partial sums of Bell numbers.

Finally, we determine the limit frequency of $L_{c}^{0}$ in $\left(\mathbb{N}_{c}\right)^{\star}\{0\}^{\star}$. Notice that

$$
\rho_{\left(\mathbb{N}_{c}\right)^{\star}\{0\}^{\star}}(n)=\frac{c^{n+1}-1}{c-1},
$$

as it is a sum of the first $n$ terms of a geometric progression of ratio $c$.
We have,
Theorem 4.2. For all $c \geq 1$,

$$
\lim _{n \rightarrow \infty} \frac{\rho_{L_{c}^{0}}(n)}{\rho_{\left(\mathbb{N}_{c}\right)^{\star}\{0\}^{\star}}(n)}=\frac{1}{c!} .
$$

| c | $\rho_{L_{c}^{0}}(n)$ | OEIS |
| :---: | :---: | :---: |
| 1 | $\begin{aligned} & n \\ & 1,2,3,4,5,6,7,8,9,10,11,12,13, \ldots \end{aligned}$ |  |
| 2 | $\begin{aligned} & 2^{n}-1 \\ & 1,3,7,15,31,63,127,255,511,1023,2047,4095,8191, \ldots \end{aligned}$ | A000225 |
| 3 | $\begin{aligned} & \frac{1}{4} 3^{n}+\frac{1}{2} n-\frac{1}{4} \\ & 1,3,8,22,63,185,550,1644,4925,14767,44292,132866,398587, \ldots \end{aligned}$ | $\underline{\text { A047926 }}$ |
| 4 | $\begin{aligned} & \frac{1}{18} 4^{n}+\frac{1}{2} 2^{n}+\frac{1}{3} n-\frac{5}{9} \\ & 1,3,8,23,74,261,976,3771,14822,58769,234044,934119,3732370, \ldots \end{aligned}$ |  |
| 5 | $\begin{aligned} & \frac{1}{96} 5^{n}+\frac{1}{8} 3^{n}+\frac{1}{3} 2^{n}+\frac{3}{8} n-\frac{15}{32} \\ & 1,3,8,23,75,277,1132,4977,22979,109451,531456,2610931,12917683, \ldots \end{aligned}$ | A099265 |
| 6 | $\begin{aligned} & \frac{1}{600} 6^{n}+\frac{1}{36} 4^{n}+\frac{1}{12} 3^{n}+\frac{3}{8} 2^{n}+\frac{11}{30} n-\frac{439}{900} \\ & 1,3,8,23,75,278,1154,5265,25913,135212,736704,4139831,23767895, \ldots \end{aligned}$ | $\underline{\text { A099266 }}$ |
| 7 | $\begin{aligned} & \frac{1}{4320} 7^{n}+\frac{1}{192} 5^{n}+\frac{1}{54} 4^{n}+\frac{3}{32} 3^{n}+\frac{11}{30} 2^{n}+\frac{53}{144} n-\frac{31}{64} \\ & 1,3,8,23,75,278,1155,5294,26404,141583,807062,4837585,30181073, \ldots \end{aligned}$ |  |
| 8 | $\frac{1}{35280} 8^{n}+\frac{1}{1200} 6^{n}+\frac{1}{288} 5^{n}+\frac{1}{48} 4^{n}+\frac{11}{120} 3^{n}+\frac{53}{144} 2^{n}+\frac{103}{280} n-\frac{57023}{117600}$ <br> $1,3,8,23,75,278,1155,5295,26441,142370,819729,5009279,32252379, \ldots$ |  |

Table 2: Density functions of $L_{c}^{0}$, for $c=1 . .8$.

Proof.

$$
\begin{aligned}
\lim _{n \rightarrow \infty} \frac{\rho_{L_{c}^{0}}(n)}{\rho_{\left(\mathbb{N}_{c}\right)^{\star}\{0\}^{\star}}(n)}= & \lim _{n \rightarrow \infty}\left(\frac{\left(c^{n+1}-c\right)(c-1)}{(c-1) c!\left(c^{n+1}-1\right)}+\frac{n(c-1)}{c^{n+1}-1} \sum_{j=2}^{c-1} \frac{(-1)^{j}}{j!}\right) \\
& +\lim _{n \rightarrow \infty}\left(\sum_{k=2}^{c-2} \frac{\left(k^{n+1}-k\right)(c-1)}{(k-1) k!\left(c^{n+1}-1\right)}\left(\sum_{j=2}^{c-k} \frac{(-1)^{j}}{j!}\right)\right) \\
= & \frac{1}{c!}\left(\lim _{n \rightarrow \infty} \frac{1}{1-\frac{1}{c^{n+1}}}-\lim _{n \rightarrow \infty} \frac{c}{c^{n+1}-1}\right) \\
& +\sum_{j=2}^{c-1} \frac{(-1)^{j}}{j!} \lim _{n \rightarrow \infty}\left(\frac{n(c-1)}{c^{n+1}-1}\right) \\
& +\sum_{k=2}^{c-2} \frac{(c-1)}{(k-1)(k-1)!} \sum_{j=2}^{c-k} \frac{(-1)^{j}}{j!} \lim _{n \rightarrow \infty}\left(\frac{\left(\frac{k}{c}\right)^{n}}{c-\frac{1}{c^{n}}}-\frac{1}{c^{n+1}-1}\right) \\
= & \frac{1}{c!} .
\end{aligned}
$$

## 5 Conclusion

In this note we presented a family of regular languages representing finite set partitions and studied their densities. Although it is well-known that the number of partitions of a set of $n$ elements into no more than $c$ nonempty sets is given by partial sums of Stirling numbers of second kind, we determined explicit formulas for their closed forms, as linear combinations of $k^{n}$, for $k \in \mathbb{N}_{c}$. We also determined the limit frequency of those languages, which gives an estimate of the space saved with those representations.

## Acknowledgements

We thank Miguel Filgueiras and the referee for their comments in previous versions of the manuscript. We also thank J. Shallit for pointing us a connection between the density of $L_{c}$ and the Bell numbers.

## References

[BHR04] Peter Blanchard, Frank Harary, and Rogério Reis, Partitions into sum-free sets, Integers: Electronic Journal of Combinatorial Number Theory, (to appear), 2004.
[GKP94] Roland Graham, Donald Knuth, and Oren Patashnik. Concrete Mathematics. Addison-Wesley, 1994.
[Hec03] Andri Heck. Introduction to Maple. Springer, 3rd edition, 2003. http://remote.science.uva.nl/\~heck/Maplebook/.
[HMU00] John E. Hopcroft, Rajeev Motwani, and Jeffrey D. Ullman. Introduction to Automata Theory, Languages and Computation. Addison-Wesley, 2nd edition, 2000.
[RMP04] Rogério Reis, Nelma Moreira, and João Pedro Pedroso. Educated brute-force to get $\mathrm{h}(4)$. Technical Report DCC-2004-04, DCC-FC\& LIACC, Universidade do Porto, July 2004.
[SF96] Robert Sedgewick and Philippe Flajolet. Analysis of Algorithms. Addison-Wesley, 1996.
[Slo03] N.J.A. Sloane. The On-line Encyclopedia of Integer Sequences, 2003. http://www.research.att.com/~njas/sequences.

2000 Mathematics Subject Classification: Primary 05A15; Secondary 05A18, 11B73, 68Q45. Keywords: Partions of sets, Stirling numbers, regular languages, enumeration
(Concerned with sequences A000110 A000225 A007051 A007581 A008277 A047926 A056272 A056273 A099262 A099263 A099265 A099266 .)

Received October 14 2004; revised version received April 17 2005. Published in Journal of Integer Sequences, May 182005.

Return to Journal of Integer Sequences home page.


[^0]:    ${ }^{1}$ Work partially funded by Fundação para a Ciência e Tecnologia (FCT) and Program POSI.

[^1]:    ${ }^{2}$ The data structures used in the programs are arrays of fixed length.

