List 3

Introductory to analytic combinatorics course at Wroclaw University of Science and Technology 2020/2021

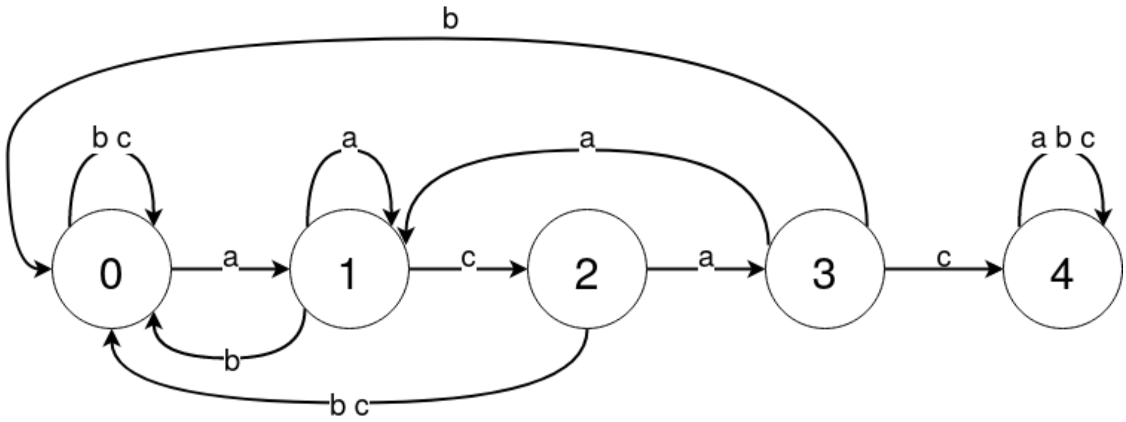
Deadline:

30.11.2020

Exercise 1 (5 points)

Let $\mathcal{L}$ be the language over the alphabet $\{a, b, c\}$ composed of words that contain pattern $acac$. Describe the appropriate finite automata and present the OFG for $\mathcal{L}$.

Having pattern of length 4 finite automa has 5 stages. They are connected by following letters from string, different letters are connected to "lower" stages. After reaching 4th stage we stay in it even after appending additional letters.



To calculate OGF for class $\mathcal{L}$. Let first how languages $\mathcal{L}_0$, .., $\mathcal{L}_4$ are connected together.

$$\begin{aligned}
\mathcal{L}_0 &= a\mathcal{L}_1 + b\mathcal{L}_0 + c\mathcal{L}_0 \\
\mathcal{L}_1 &= a\mathcal{L}_1 + b\mathcal{L}_0 + c\mathcal{L}_2 \\
\mathcal{L}_2 &= a\mathcal{L}_3 + b\mathcal{L}_0 + c\mathcal{L}_0 \\
\mathcal{L}_3 &= a\mathcal{L}_1 + b\mathcal{L}_0 + c\mathcal{L}_4 \\
\mathcal{L}_4 &= a\mathcal{L}_4 + b\mathcal{L}_4 + c\mathcal{L}_4 + \varepsilon
\end{aligned}$$

This gives rise to a set of equations for the associated OGFs

$$
\begin{aligned}
L_0 &= zL_0 + 2zL_0 \\
L_1 &= zL_0 + bzL_1 + zL_2 \\
L_2 &= 2zL_0 + zL_3 \\
L_3 &= zL_0 + zL_1 + zL_4 \\
L_4 &= 3zL_4 + 1
\end{aligned}
$$

And after a lot of algebraic transformations we get

Exercise 2 (4 points)

How many words of length n are over the alphabet $\mathcal{A} = \{a, b, c, d, e, f, g\}$ that contain the block pattern $\mathbb{p} = aabbaa$?

To calculate the number of words containing pattern $aabbaa$ we can subtract the number not containing pattern from all.

We can calculate the number of words that do not contain pattern using formula

$$S(z) = \frac{c(z)}{(z^k + (1 - mz)c(z))}$$

Where:

m is alphabet size $(|\mathcal{A}| = 6)$,

k is pattern length $(|\mathbb{p}| = 6)$,

c(z) is the autocorrelation polynomial of a pattern, equal to $c(z) = \sum_{j=0}^{k-1} c_j z^j$

Where $c_i = ||p_{i+1}p_{i+2}\ldots p_k = p_1p_2\ldots p_{k-i}||$ equals 1 or 0

In [10]:
```
%%latex
$a,a,b,b,a,a$
\begin{align}
a,a,b,b,a,a\phantom{a,a,a,a,a,a,a} &= 1\\
a,a,b,b,a,a\phantom{a,a,a,a,a,a} &= 0\\
a,a,b,b,a,a\phantom{a,a,a,a,a} &= 0\\
a,a,b,b,a,a\phantom{a,a,a,a} &= 0\\
a,a,b,b,a,a\phantom{a,a,a} &= 1\\
a,a,b,b,a,a\phantom{a,a} &= 1
\end{align}
```

a, a, b, b, a, a

$$
\begin{aligned}
a, a, b, b, a, a &= 1\\
a, a, b, b, a, a &= 0\\
a, a, b, b, a, a &= 0\\
a, a, b, b, a, a &= 0\\
a, a, b, b, a, a &= 1\\
a, a, b, b, a, a &= 1
\end{aligned}
$$

$c(z) = 1 + z^4 + z^5$

We need also a number of words of size n, with equals to m^n. So finally the number of words containing pattern equals to

$$[z^n]\mathcal{A}_n(z) = \frac{1}{} - \frac{1 + z^4 + z^5}{}$$

Exercise 3 (3 points)

Apply Lagrange inversion theorem to calculate the coefficient with the term z^n for the function $L(z)$, which is a solution of the equation

$$\frac{L(z)}{z} = 1 + (L(z))^3.$$

Let's start by simplifying the equation

$$L(z) = z\left(1 + (L(z))^3\right)$$
$$z = \frac{L(z)}{1 + (L(z))^3}$$

Now we can use Lagrange Inversion Theorem. It states that:

The coefficients of an inverse function and of all its powers are determined by coefficients of powers of the direct function: if $z = T/\phi(T)$, then one has (with any $k \in \mathcal{Z} \geq 0$):

$$[z^n]T(z) = \frac{1}{n}[\omega^{n-1}]\phi(\omega)^n$$

In our case $L(z) = T(z)$, $\omega = u$ and $\phi(u) = 1 + u^3$.

$[z^n]L(z) = \frac{1}{n}[u^{n-1}](1+u^3)^n = \frac{1}{n}[u^{n-1}]\sum_{k=0}^{n}\binom{n}{k}u^{3k}$

Let us introduce auxiliary variable $3m + 1 = n \rightarrow m = \frac{n-1}{3}$

$[z^n]L(z) = \frac{1}{n}[u^{3m+1-1}]\sum_{k=0}^{3m+1}\binom{3m+1}{k}u^{3k} = \frac{1}{n}[u^{3m}]\sum_{k=0}^{3m+1}\binom{3m+1}{k}u^{3k} = \frac{1}{n}\binom{3m+1}{m}$

Exercise 4 (4 points)

Write the formula OFG $S^{(4)}(z)$ such that the coefficient for z^n will be a Stirling number II of the type $S(n, 4)$ that is $[z^n]S^{(4)}(z) = S(n, 4)$.

$S^{(r)}(z)$ is defined as

$$b_1 \cdot SEQ(b_1) \cdot b2 \cdot SEQ(b_1 + b_2) \cdot \ldots \cdot b_r \cdot SEQ(b_1 + b_2 + \cdots + b_r)$$

The language specification immediately gives the OGF

$$S^{(r)}(z) = \frac{z^r}{(1-z)(1-2z)(1-3z)\cdot \ldots \cdot(1-rz)}$$

The partial fraction expansion of S(r)(z) is then readily computed,

$$S^{(r)}(z) = \frac{1}{r!} \sum_{j=0}^{r} \frac{(-1)^{r-j}}{1 - jz}$$

And for $r = 4$ it evaluates to