*Discrete Mathematics and Theoretical Computer Science* DMTCS vol. **12:3**, 2010, 75–80

# *On the asymptotic enumeration of accessible automata*

Elcio Lebensztayn[†]

*Institute of Mathematics and Statistics, University of São Paulo, CEP 05508-090, São Paulo, Brazil.*

*received 5th April 2010, revised 7th July 2010, accepted 13th July 2010.*

We simplify the known formula for the asymptotic estimate of the number of deterministic and accessible automata with $n$ states over a $k$-letter alphabet. The proof relies on the theory of Lagrange inversion applied in the context of generalized binomial series.

**Keywords:** asymptotic enumeration, finite automata, Lagrange series, generalized binomial series.

## 1 Main result

Let $\mathcal{A}_n$ be the set of non-isomorphic accessible (also called initially connected) complete and deterministic automata with $n$ states over a $k$-letter alphabet. By using suitable combinatorial transformations, Bassino and Nicaud (2007) prove that the order of magnitude of the cardinality of $\mathcal{A}_n$ is related to the Stirling number of the second kind $\left\{ {kn \atop n} \right\}$. Then they reformulate the asymptotic estimate of $|\mathcal{A}_n|$ due to Korshunov (1978) in terms of this Stirling number. Our purpose is to state a simplification of this formula.

First we need to recall a few definitions. For $m$, $n$ non-negative integers, the *Stirling number of the second kind*, denoted by $\left\{ {n \atop m} \right\}$, is the number of ways of partitioning a set of $n$ elements into $m$ non-empty subsets. Let $W_0$ denote the principal branch of the so-called *Lambert $W$ function* (which is the inverse of the function $x \mapsto x\, e^x$). See Graham et al. (1994) and Corless et al. (1996) for more details about these definitions. Also define the constants

$$v_k = 1 + \frac{W_0(-k\, e^{-k})}{k},$$

$$\zeta_k = k\, v_k = k + W_0(-k\, e^{-k}),$$

$$\beta_k = \frac{k^k}{e^{k-1}} \frac{(e^{\zeta_k} - 1)}{\zeta_k^k}.$$

Except for $v_k$, these constants are considered in Bassino and Nicaud (2007). Notice that $v_k$ is the unique root in $(0, 1)$ of the function $\phi_k(v) = k\, v + \log(1 - v)$ and it satisfies $v_k > (k-1)/k$. See Figure 1.

[†]Research supported by CNPq (311909/2009-4).

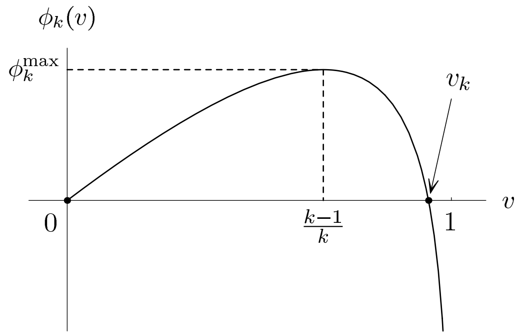


**Fig. 1:** Graph of $\phi_k(v) = k\,v + \log(1-v)$.

The next result, due to Bassino and Nicaud (2007), reformulates in terms of the Stirling numbers the asymptotic estimate for $|\mathcal{A}_n|$ previously obtained by Korshunov (1978).

**Theorem 1 (Bassino and Nicaud (2007))** *The number $|\mathcal{A}_n|$ of accessible complete and deterministic automata with $n$ states over a $k$-letter alphabet satisfies*

$$|\mathcal{A}_n| \sim E_k\, n\, 2^n \left\{ {kn \atop n} \right\},$$

*where*

$$E_k = \frac{1 + \sum_{r=1}^{\infty} \frac{1}{r}\binom{kr}{r-1}(e^{k-1}\beta_k)^{-r}}{1 + \sum_{r=1}^{\infty} \binom{kr}{r}(e^{k-1}\beta_k)^{-r}}.$$

Our main result presents a simplified formula for $E_k$.

**Theorem 2** *For any $k \geq 2$,*

$$E_k = k - \frac{(k-1)}{v_k}.$$

**Remark 1** *We observe that $0 < E_k < 1$ for all $k \geq 2$. Since the principal branch of $W$ is analytic at $0$, we have that*

$$\lim_{k\to\infty} k\,(v_k - 1) = \lim_{k\to\infty} W_0(-k\,e^{-k}) = W_0(0) = 0.$$

*Consequently, $\lim_{k\to\infty} v_k = 1$ and*

$$\lim_{k\to\infty} E_k = \lim_{k\to\infty} \frac{k\,(v_k-1)+1}{v_k} = 1.$$

*Table 1 exhibits the approximate values of $v_k$ and $E_k$ for $k = 2, \ldots, 7$.*

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| $v_k$ | 0.796812 | 0.940480 | 0.980173 | 0.993023 | 0.997484 | 0.999082 |
| $E_k$ | 0.744999 | 0.873426 | 0.939315 | 0.971895 | 0.987386 | 0.994488 |

**Tab. 1:** Values of $v_k$ and $E_k$.

## 2 Proof of Theorem 2

The proof relies on the theory of Lagrange series, which constitutes a fundamental tool for the solution of implicit equations and for the reversion of series. We refer to Henrici (1988) for an extensive treatment of this subject. To recall the two basic forms of Lagrange series, suppose that the dependence between the variables $y$ and $z$ is implicitly given by the equation

$$y = \frac{z}{\varphi(z)},$$

where $\varphi$ is a regular function in a neighborhood of the point $z = 0$ and $\varphi(0) \neq 0$. If $f$ is a regular function in a neighborhood of $z = 0$, then

$$\begin{aligned} f(z) &= f(0) + \sum_{r=1}^{\infty} \frac{y^r}{r!} \left[ \frac{d^{r-1} \left( f'(z) \left[\varphi(z)\right]^r \right)}{dz^{r-1}} \right]_{z=0}, \\ \frac{f(z)}{1 - y\,\varphi'(z)} &= \sum_{r=0}^{\infty} \frac{y^r}{r!} \left[ \frac{d^{r} \left( f(z) \left[\varphi(z)\right]^r \right)}{dz^{r}} \right]_{z=0}, \end{aligned} \tag{1}$$

where the primes denote derivatives.

To prove Theorem 2, we study a classical application of this theory, namely, the solutions of Problems 212 and 216 in Pólya and Szegő (1998, Chapter 5 of Part III). Let $a$ and $b$ be real numbers and consider the equation $1 - x + y\,x^b = 0$, that is, $y = (x-1)\,x^{-b}$. We define $z = x - 1$ and apply the formulae in (1) with $\varphi(z) = (1+z)^b$ and $f(z) = (1+z)^a$. Notice that

$$\frac{1}{r!} \left[ \frac{d^{r-1} \left( f'(z) \left[\varphi(z)\right]^r \right)}{dz^{r-1}} \right]_{z=0} = \frac{1}{r!} \left[ \frac{d^{r-1} \left( a\,(1+z)^{a+br-1} \right)}{dz^{r-1}} \right]_{z=0} = \frac{a}{a+br} \binom{a+br}{r},$$

$$\frac{1}{r!} \left[ \frac{d^{r} \left( f(z) \left[\varphi(z)\right]^r \right)}{dz^{r}} \right]_{z=0} = \frac{1}{r!} \left[ \frac{d^{r} \left( (1+z)^{a+br} \right)}{dz^{r}} \right]_{z=0} = \binom{a+br}{r}.$$

Hence, (1) yields

$$x^a = \sum_{r=0}^{\infty} \frac{a}{a+br} \binom{a+br}{r} y^r \tag{2}$$

and

$$\frac{x^{a+1}}{(1-b)x + b} = \sum_{r=0}^{\infty} \binom{a+br}{r} y^r, \tag{3}$$

with $y = (x-1)\,x^{-b}$. Next, we consider the case $b > 1$ and investigate the neighborhood in $\mathbb{R}$ of the point $x = 1$ in which (2) and (3) hold.

For $a = 1$, the power series appearing in (2)

$$B_b(y) := \sum_{r=0}^{\infty} \binom{br+1}{r} \frac{y^r}{br+1} \tag{4}$$

is called the generalized binomial series. Thus, $B_b$ is the solution of the functional equation $x(y) = 1 + y\,x(y)^b$ that satisfies $x(0) = 1$. Moreover, (2) and (3) can be rewritten as

$$B_b(y)^a = \sum_{r=0}^{\infty} \frac{a}{a+br}\binom{a+br}{r} y^r \quad \text{and} \quad \frac{B_b(y)^{a+1}}{(1-b)B_b(y)+b} = \sum_{r=0}^{\infty} \binom{a+br}{r} y^r.$$

In Graham et al. (1994, Sections 5.4 and 7.5), $B_b$ is introduced as a formal power series, these properties are proved and used to derive some combinatorial identities. See also Heggie and Nicklason (1996) for a representation of the function $B_b$ as a contour integral in the complex plane.

Now let $b > 1$ be fixed. We are interested in studying the set of values of $x$ for which the identity $B_b((x-1)\,x^{-b}) = x$ holds. To this end, we observe that by the ratio test the series in (4) converges for $|y| < R$, where

$$R := \frac{(b-1)^{b-1}}{b^b}.$$

**Lemma 1** *There exists $\bar{x}_b \in (1/2, 1)$ such that*

$$B_b((x-1)\,x^{-b}) = x$$

*for all $x \in (\bar{x}_b, b\,(b-1)^{-1})$.*

**Proof:** It is not difficult to show that the function $g_b(x) = (x-1)\,x^{-b}$ is continuous and strictly increasing in the interval $(0, b\,(b-1)^{-1})$. The result follows since $g_b(b\,(b-1)^{-1}) = R$, $g_b(1) = 0$ and $|g_b(1/2)| > 1 > R$. □

Consequently, identities (2) and (3) are valid for every $x \in (\bar{x}_b, b\,(b-1)^{-1})$.

Now define $\bar{v}_b = (\bar{x}_b)^{-1} \in (1, 2)$. From Lemma 1, we obtain the following result.

**Lemma 2** *For any $v \in ((b-1)/b, \bar{v}_b)$,*

$$\frac{1}{v^a} = \sum_{r=0}^{\infty} \frac{a}{a+br}\binom{a+br}{r} y^r \quad \textit{and} \quad \frac{1}{v^a\,[b\,v-(b-1)]} = \sum_{r=0}^{\infty} \binom{a+br}{r} y^r,$$

*with $y = (1-v)\,v^{b-1}$.*

**Proof:** For each $v \in ((b-1)/b, \bar{v}_b)$, define $x = 1/v$, so $x \in (\bar{x}_b, b\,(b-1)^{-1})$ and $(x-1)\,x^{-b} = (1-v)\,v^{b-1}$. The result follows from (2) and (3). □

We are ready to prove our main result.

**Proof of Theorem 2:** Note that

$$e^{k-1}\beta_k = \frac{k^k\,(e^{\zeta_k}-1)}{\zeta_k^k} = \frac{1}{(1-v_k)\,v_k^{k-1}}.$$

Defining $y_k = (1-v_k)\,v_k^{k-1}$, we conclude from Lemma 2 that the numerator of $E_k$ equals

$$1+\sum_{r=1}^{\infty}\frac{1}{r}\binom{kr}{r-1}y_k^r = \sum_{r=0}^{\infty}\frac{1}{kr+1}\binom{kr+1}{r}y_k^r = \frac{1}{v_k},$$

whereas the denominator simplifies to

$$\sum_{r=0}^{\infty}\binom{kr}{r}y_k^r = \frac{1}{k\,v_k-(k-1)}.$$

The result follows. □

## Acknowledgements

I would like to thank the referees for their valuable remarks and suggestions.

## References

F. Bassino and C. Nicaud. Enumeration and random generation of accessible automata. *Theoret. Comput. Sci.* 381 (1-3): 86–104, 2007.

R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey and D. E. Knuth. On the Lambert $W$ function. *Adv. Comput. Math.* 5 (4): 329–359, 1996.

R. L. Graham, D. E. Knuth and O. Patashnik. *Concrete mathematics: a foundation for computer science.* 2nd. ed., Addison-Wesley, Reading, MA, 1994.

M. Heggie and G. R. Nicklason. An integral representation for the generalized binomial function. *Canad. Math. Bull.* 39 (1): 59–67, 1996.

P. Henrici. *Applied and computational complex analysis, vol. 1.* Reprint of the 1974 original. Wiley, New York, 1988.

D. Korshunov. Enumeration of finite automata. *Probl. Kibern.* 34: 5–82, 1978 (in Russian).

G. Pólya and G. Szegő. *Problems and theorems in analysis, vol. I.* Translated from the German by D. Aeppli. Springer-Verlag, New York, 1998.