Issue Date: 28.06.2013
Version: 2013-07-03 11:35

Exercise Sheet 10 for Combinatorial Algorithms, SS 13

Hand In: *Until Monday, 08.07.2013, 12:00,*
box in the group's hallway or email to `wild@cs.uni....`

Problem 17

$1 + 2 + 2$ points

In this exercise we consider the *Motzkin numbers* M_n once more. As we have already seen, their (ordinary) generating function is

$$M(z) \;=\; \sum_{n\geq 0} M_n z^n \;=\; \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z^2} \,. \tag{1}$$

We are going to use singularity analysis to derive exact asymptotics for M_n.

a) In order to derive asymptotics, we would like to have a generating function that is as simple as possible. Consider the simpler relative of $M(z)$

$$\hat{M}(z) \;=\; \sum_{n\geq 0} \hat{M}_n z^n \;=\; -\tfrac{1}{2}\sqrt{1 - 2z - 3z^2} \,.$$

Prove that for $n \geq 2$, we have $\hat{M}_n = M_{n-2}$.

b) Derive exact asymptotics for $\hat{M}_n$, i. e., find an explicit expression $\hat{m}(n)$ with

$$\lim_{n\to\infty} \frac{\hat{m}(n)}{\hat{M}_n} \;=\; 1 \,.$$

We abbreviate that as $\hat{M}_n \sim \hat{m}(n)$ as $n \to \infty$ and say "$\hat{M}_n$ is *asymptotically equivalent* to $\hat{m}(n)$."

Compute and/or plot the relative error of your asymptotic for some moderate values of n; use your favorite computer algebra system, the online tools on our website `wwwagak.cs.uni-kl.de/mathe-tools.html` or Wolfram Alpha.

Hint: Use the Corollary[1] from Theorem 5.5 of [SF13].

[1]Beware of the typing error in the book: You have to replace ρ^n by ρ^{-n}.

Hint: To compute specific values of the Gamma function $\Gamma(z)$, the following basic properties are handy, see [DLMF]:

$$\Gamma(n+1) = n! \qquad n \in \mathbb{N} \qquad (\Gamma 1)$$

$$\Gamma(z+1) = z\,\Gamma(z) \qquad z \in \mathbb{C} \qquad (\Gamma 2)$$

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi} \qquad (\Gamma\tfrac{1}{2})$$

$$\Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\sin(\pi z)} \qquad z \in \mathbb{C} \qquad (\Gamma 3)$$

c) Recall Problem 16, where you built a top-down random sampler for the combinatorial class $\mathcal{S}$ of RNA secondary structures, given by:

$$\mathcal{S} = \epsilon + \mathcal{Z}_* \times \mathcal{S} + \mathcal{Z}_(\times \mathcal{S} \times \mathcal{Z}_) \times \mathcal{S}\,. \qquad (2)$$

Use your new skills in singularity analysis to verify or disprove your conjecture about the distribution of the number of unpaired bases in a uniformly chosen RNA secondary structure of size n.

Hint: The top-down sampler identifies a single point where it is decided which sort of atom to produce next. Find this point in your sampler and express the probability p_* for a next symbol to be of type $\mathcal{Z}_*$ as the ratio of the coefficients of two generating functions. With a computation very similar to b) you can compute the limit of p_*.

Problem 18

2 + 2 + 1 points

Consider once again the class of secondary structures given in (2).

a) Implement a Boltzmann sampler $\Gamma S(x)$ for $\mathcal{S}$ with parameter $x = 0.33$ as described in Section 3 of [DFLS04].

Keep your implementation adaptable for other choices of x, but for simplicity, you may precompute the needed constants externally and hard-code them into your program.

b) Let N be the (random!) size of a RNA secondary structure generated by $\Gamma S(0.33)$. Compute the expected size $\mathbb{E}\,N$ and its standard deviation $\sigma = \sqrt{\mathbb{V}\,N}$.

Use *Chebychev's inequality* to compute an upper bound $N_{0.99}$, such that with at least 99 % probability, a random structure generated by $\Gamma S(0.33)$ has size at most N; formally

$$\Pr[N \le N_{0.99}] \ge 0.99\,.$$

c) Use your Boltzmann sampler to generate 1 000 random RNA secondary structures and draw a histogram of their sizes.

What can you say regarding the Chebychev tail bound you derived in b)?

Figure 1 of [DFLS04] shows three categories of size distributions for Boltzmann samplers: "bumpy", "flat" and "peaked". In which of these three categories does $\Gamma S(x)$ seem to belong?

References

[DFLS04] Philippe Duchon, Philippe Flajolet, Guy Louchard, and Gilles Schaeffer. Boltzmann Samplers for the Random Generation of Combinatorial Structures. *Combinatorics, Probability and Computing*, 13(4-5):577–625, July 2004. doi:10.1017/S0963548304006315.

[DLMF] NIST Digital Library of Mathematical Functions. Release 1.0.6 of 2013-05-06. Online companion to the Handbook of Mathematical Functions. URL: http://dlmf.nist.gov/.

[SF13] Robert Sedgewick and Philippe Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley Professional, 2nd edition, 2013.