An Introduction to
ANALYTIC COMBINATORICS

Computer Science 488
Robert Sedgewick

http://aofa.cs.princeton.edu

AofA REVIEW
Basic facts about review problem sets


**Purpose**
- Check understanding of key facts.
- Build confidence in what you have learned.
- Prepare for the next level.

**We know that you don't have much time to study**
- Questions are like old exam questions, adjusted because you have more time to solve them.
- At most 1 question per lecture
- Review the material and your own work.
- Create a "cheatsheet" of important facts.

**Each problem is a small part (<2%) of the story.**
Q. Solve the following recurrence

\[ F_N = N^2 + \frac{1}{N} \sum_{1 \leq k \leq N} (F_{k-1} + F_{N-k}) \text{ with } F_0 = 0. \]

\[ C_N = N + 1 + \frac{1}{N} \sum_{1 \leq j \leq N} (C_{j-1} + C_{N-j}) \]

\[ = N + 1 + \frac{2}{N} \sum_{1 \leq j \leq N} C_{j-1} \]

\[ NC_N - (N - 1)C_{N-1} = 2N + 2C_{N-1} \]

\[ \frac{C_N}{N + 1} = \frac{C_{N-1}}{N} + \frac{2}{N + 1} \]

\[ \ldots \]
Q. Match each “toll function” at left with the order of growth of the solution at right for the Quicksort recurrence

$$F_N = t_N + \frac{1}{N} \sum_{1 \leq k \leq N} (F_{k-1} + F_{N-k}) \quad \text{with} \quad F_0 = 0$$

$$NC_N - (N + 1)C_{N-1} = 3N^2 - 3N + 1$$

$$\frac{C_N}{N+1} = \frac{C_{N-1}}{N} + 3 + \ldots$$

A good question must avoid answers that depend on detailed calculations.
AofA Chapter 2 Recurrences Q&A example

Q. Solve the recurrence

\[ na_n = (n - 3)a_{n-1} + n \quad \text{for} \quad n \geq 3 \quad \text{with} \quad a_n = 0 \quad \text{for} \quad n \leq 2 \]

A. \[ n(n-1)(n-2)a_n = (n-1)(n-2)(n-3)a_{n-1} + n(n-1)(n-2) \quad \text{for} \quad n \geq 3 \]

\[
\frac{\binom{n}{3} a_n}{\binom{n-1}{3}} = \binom{n-1}{3} a_{n-1} + \binom{n}{3} \quad \text{for} \quad n \geq 3
\]

\[
= \sum_{3 \leq k \leq n} \frac{k}{3} = \binom{n+1}{4}
\]

\[ a_n = \frac{n+1}{4} \]

Borderline suitable for a COS 488 inclass exam, but OK for a review problem.

\[
\text{easier:} \quad na_n = (n - 3)a_{n-1} + 4n
\]
AofA Chapter 3 GFs Q&A example

**Q.** Match each of sequence with its OGF.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>OGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 0, 1, 3, 6, 10, ...</td>
<td>( \frac{1}{1-3z} )</td>
</tr>
<tr>
<td>0, 0, 1/2, 0, 1/4, 0, 1/6, ...</td>
<td>( \frac{z^2}{(1-z)^3} )</td>
</tr>
<tr>
<td>1, 3, 9, 27, 81, 243, ...</td>
<td>( \ln \frac{1}{1-z^2} )</td>
</tr>
<tr>
<td>1, 1 + 1/2, 1 + 1/2 + 1/3, ...</td>
<td>( \frac{3}{1-z} )</td>
</tr>
<tr>
<td>3, 3, 3, 3, 3, ...</td>
<td>( \ln \frac{1}{1-2z} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{1-z} \ln \frac{1}{1-z} )</td>
</tr>
<tr>
<td></td>
<td>( \frac{1}{(1-z)^3} )</td>
</tr>
</tbody>
</table>

\[
\frac{z^M}{(1-z)^M} = \sum_{N \geq M} \binom{N}{M} z^M
\]

\[
\ln \frac{1}{1-z} = \sum_{N \geq 1} \frac{z^N}{N}
\]

Maybe put formulas you won't quickly remember on the cheatsheet.

OK for a COS 488 inclass exam, but too easy for a review problem.
Q. Give an asymptotic approximation of $e^{H_{2N} - H_N}$ to within $O\left(\frac{1}{N^2}\right)$

A.

\[
H_{2N} - H_N = \ln(2N) + \gamma + \frac{1}{4N} + O\left(\frac{1}{N^2}\right)
\]

\[- \ln N - \gamma - \frac{1}{2N} + O\left(\frac{1}{N^2}\right)
\]

\[= \ln 2 - \frac{1}{4N} + O\left(\frac{1}{N^2}\right)\]

\[\exp\left(\ln 2 - \frac{1}{4N}\right) = 2 \exp\left(-\frac{1}{4N}\right) = 2 - \frac{1}{2N} + O\left(\frac{1}{N^2}\right)\]

\[H_N = \ln N + \gamma + \frac{1}{2N} + O\left(\frac{1}{N^2}\right)\]

\[e^x = 1 + x + \frac{x^2}{2} + O\left(x^3\right)\]

Maybe put formulas you won't quickly remember on the cheatsheet.
Q. Match each function with an asymptotic expansion.

<table>
<thead>
<tr>
<th>Function</th>
<th>Asymptotic Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_N$</td>
<td>$1 + \frac{1}{2N} + O\left(\frac{1}{N^2}\right)$</td>
</tr>
<tr>
<td>$\exp(H_{2N} - H_N) - 1$</td>
<td>$N + O(1)$</td>
</tr>
<tr>
<td>$\exp(H_N)$</td>
<td>$1 - \frac{1}{N} + O\left(\frac{1}{N^2}\right)$</td>
</tr>
<tr>
<td>$\exp\left(\frac{1}{N}\right)$</td>
<td>$1 - \frac{1}{2N} + O\left(\frac{1}{N^2}\right)$</td>
</tr>
<tr>
<td>$(1 + \frac{1}{N})^{-1}$</td>
<td>$\ln N + \gamma + \frac{1}{2N} + O\left(\frac{1}{N^2}\right)$</td>
</tr>
</tbody>
</table>
Q. Match each of the topics described in the book with a mathematician’s name.

Approximate a sum with an integral: Laplace
Expand a differentiable function: Ramanujan
Approximate factorials: Euler
Birthday function: Taylor
Approximate a function by swapping tails: Stirling

No, this is not high school, but... You do not want to appear to be ignorant!
Q. Match each expression with an approximation to its value.

<table>
<thead>
<tr>
<th>Expression</th>
<th>Approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.01^{10}$</td>
<td>1.10102</td>
</tr>
<tr>
<td>$1.05^{10}$</td>
<td>1.10462</td>
</tr>
<tr>
<td>$1.01^{20}$</td>
<td>1.22019</td>
</tr>
<tr>
<td>$1.01^{50}$</td>
<td>1.50034</td>
</tr>
<tr>
<td>$1.01^{100}$</td>
<td>1.62889</td>
</tr>
<tr>
<td></td>
<td>1.64463</td>
</tr>
<tr>
<td></td>
<td>2.02300</td>
</tr>
<tr>
<td></td>
<td>2.70481</td>
</tr>
<tr>
<td></td>
<td>2.71828</td>
</tr>
</tbody>
</table>
\[
(1 + x)^t = \sum_{0 \leq k \leq t} \binom{t}{k} x^k
\]

\[
= 1 + tx + \frac{t(t - 1)}{2} x^2 + O(x^3)
\]

\[
(1 + \frac{1}{N})^t = 1 + \frac{t}{N} + \frac{t(t - 1)}{2N^2} + O\left(\frac{1}{N^3}\right)
\]

\[
1.01^{10} = 1 + \frac{10}{100} + \frac{90}{20000} + \ldots
\]

\[
\approx 1.1045
\]

\[
(1 + \frac{1}{N})^{\alpha N} = 1 + \frac{\alpha N}{N} + \frac{\alpha^2 N^2}{2N^2} + \ldots
\]

\[
(1 + \frac{1}{N})^{\alpha N} = \exp\left(\alpha N \ln(1 + 1/N)\right)
\]

\[
= \exp\left(\alpha N \left(1/N + O(1/N^2)\right)\right)
\]

\[
= e^\alpha + O\left(\frac{1}{N}\right)
\]

\[
1.01^{50} \approx \sqrt{e}
\]
Q. Match each expression with an approximation to its value.

<table>
<thead>
<tr>
<th>Expression</th>
<th>Approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.01^{10}$</td>
<td>1.10102</td>
</tr>
<tr>
<td>$1.05^{10}$</td>
<td>1.10462</td>
</tr>
<tr>
<td>$1.01^{20}$</td>
<td>1.22019</td>
</tr>
<tr>
<td>$1.01^{50}$</td>
<td>1.50034</td>
</tr>
<tr>
<td>$1.01^{100}$</td>
<td>1.62889</td>
</tr>
</tbody>
</table>

Great for a COS 488 inclass exam, but no good for a review problem.
Def. A unary-binary tree is a rooted, ordered tree with node degrees all 0, 1, or 2.
Q. How many unary-binary trees?

**construction**

\[ M = Z + Z \times M + Z \times M \times M \]

**GF equation**

\[ M(z) = z + zM(z) + zM(z)^2 \]

**explicit form**

\[
\begin{align*}
M(z) &= \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z} \\
&= \frac{1 - z - \sqrt{(1 + z)(1 - 3z)}}{2z}
\end{align*}
\]

**coefficient asymptotics**

\[
[z^n]M(z) = \frac{\sqrt{4/3}}{(4/3)\sqrt{\pi n^3}} 3^n = \frac{1}{\sqrt{4\pi n^3/3}} 3^n
\]

\[ \frac{f(z)}{(1 - z/\rho)^\alpha} \sim f(\rho) \frac{\Gamma(\alpha)}{\Gamma(1)} \rho^{-n} n^{\alpha-1} \]

\[
\begin{align*}
\Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} \\
\Gamma(1) &= 1 \\
\Gamma(s + 1) &= s\Gamma(s)
\end{align*}
\]

NOT AN EXAM QUESTION (too much calculation)

“Motzkin numbers”
IMPORTANT NOTE: It’s wise to check GF equations before trying to solve them!

\[
M(z) = z + zM(z) + zM(z)^2
\]

\[
M(z) = z + z^2 + 2z^3 + 4z^4 + 9z^5 + \ldots
\]

\[
z = z
\]

\[
zM(z) = z^2 + z^3 + 2z^4 + 4z^5 + \ldots
\]

\[
z(M(z)^2) = z^3 + z^4 + 2z^5 + \ldots
\]

\[
+ z^4 + z^5 + \ldots
\]

\[
+ 2z^5 + \ldots
\]

\[
z + zM(z) + zM(z)^2 = z + z^2 + 2z^3 + 4z^4 + 9z^5 + \ldots
\]
**Q.** How many unary-binary trees?

<table>
<thead>
<tr>
<th>construction</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GF equation</td>
<td></td>
</tr>
<tr>
<td>explicit form</td>
<td>$M(z) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z}$</td>
</tr>
<tr>
<td>coefficient asymptotics</td>
<td></td>
</tr>
</tbody>
</table>
§6.15  Trees

Now, Theorem 4.11 provides an immediate proof that \([z^N]M(z)\) is \(O(3^N)\), and methods from complex asymptotics yield the more accurate asymptotic estimate \(3^N/\sqrt{3/4\pi N^3}\). Actually, with about the same amount of work, we can derive a much more general result.

Errata posted in 2019:

- 335. First sentence should read "The corollary to Theorem 5.5 (page 250) provides an immediate proof that \([z^n]M(z)\sim 3^n/\sqrt{4\pi n^3/3}."

Complex asymptotics is often not needed.

BUT it allows us to address entire classes of problems (stay tuned).
AofA Chapter 5 Trees Q&A example

Q. Match each diagram with its description

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>random Catalan tree</td>
<td><img src="image" alt="A" /></td>
</tr>
<tr>
<td>B</td>
<td>random BST</td>
<td><img src="image" alt="B" /></td>
</tr>
<tr>
<td>C</td>
<td>random AVL tree</td>
<td><img src="image" alt="C" /></td>
</tr>
</tbody>
</table>

Some questions are *very easy* if you’ve watched the lecture (and impossible if not).

Too easy for a review problem.
Q. What is the probability that a random perm of size $n$ has exactly 2 cycles?
Q. What is the probability that a random perm of size $n$ has exactly 2 cycles?

### Construction

$$P_2 = SET_2(CYC(Z))$$

### EGF equation

$$P(z) = \frac{1}{2} \left( \ln \left( \frac{1}{1 - z} \right) \right)^2$$

### Coefficient asymptotics

$$\frac{1}{2} \ln \left( \frac{1}{1 - z} \right)^2 = \sum_{n \geq 0} \frac{p_n}{n!} z^n$$

$$\frac{1}{1 - z} \ln \left( \frac{1}{1 - z} \right) = \sum_{n \geq 1} \frac{p_n}{(n - 1)!} z^{n-1}$$

$$p_n = (n - 1)! H_{n-1}$$

---

**A.** $H_{n-1}/n$
Q. What is the probability that a random perm of size $n$ has exactly 2 cycles?

<table>
<thead>
<tr>
<th>construction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EGF equation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>coefficient asymptotics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_n = (n - 1)!H_{n-1}$</td>
</tr>
</tbody>
</table>
Q. OGF for number of bitstrings not containing 01010?

**constructions**

\[ E + (Z_0 + Z_1) \times B = B + P \]

\[ Z_{01010} \times B = P + Z_{01} \times P + Z_{0101} \times P \]

**GF equations**

\[ 1 + 2zB(z) = B(z) + P(z) \]

\[ z^5 B(z) = (1 + z^2 + z^4)P(z) \]

**explicit form**

\[ B(z) = \frac{1 + z^2 + z^4}{z^5 + (1 - 2z)(1 + z^2 + z^4)} \]
Q. Fill in the blanks in this OGF for the number of bitstrings not containing 01010.

**constructions**

\[ E + (Z_0 + Z_1) \times B = B + P \]

\[ Z_{01010} \times B = \]

**GF equations**

\[ 1 + 2zB(z) = \]

\[ z^5 B(z) = (1 + z^2 + z^4)P(z) \]

**explicit form**

\[ B(z) = \frac{1 + z^2 + z^4}{z^5 + (1 - 2z)(1 + z^2 + z^4)} \]
Q. Rank these patterns by expected wait time in a random bit string.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Wait Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>00000</td>
<td>62</td>
</tr>
<tr>
<td>00001</td>
<td>32</td>
</tr>
<tr>
<td>01000</td>
<td>34</td>
</tr>
<tr>
<td>01010</td>
<td>36</td>
</tr>
<tr>
<td>10101</td>
<td>36</td>
</tr>
</tbody>
</table>

The expected wait time can be calculated using the formula:

\[
c(z) = 1 + z^4
\]

\[
B_p(z) = \frac{c(z)}{z^P + (1 - 2z)c(z)}
\]

Wait time: \(2^Pc(1/2)\)
Q. Find the probability that a random mapping has no singleton cycles.

**Constructions**

\[ C = Z \ast \text{SET}(C) \quad M = \text{SET}(CYC_{\geq 1}(C)) \]

**EGF equations**

\[ C(z) = ze^{C(z)} \quad M(z) = \exp\left(\ln\frac{1}{1-C(z)} - C(z)\right) = \frac{e^{-C(z)}}{1-C(z)} \]

**Coefficients via Lagrange inversion**

NOT AN EXAM QUESTION (too much calculation)

**Asymptotic result**

Lagrange Inversion Theorem (Bürgmann form).

If a GF \( g(z) = \sum_{n \geq 1} g_n z^n \) satisfies the equation \( z = f(g(z)) \) with \( f(0) = 0 \) and \( f'(0) \neq 0 \) then, for any function \( H(u) \),

\[ [z^n]H(g(z)) = \frac{1}{n} [u^{n-1}] H'(u) \left( \frac{u}{f(u)} \right)^n \]

Still, you might want this on your cheatsheet.
Q. Give the EGF for random mappings with no singleton cycles. Express your answer as a function of the Cayley function $C(z) = ze^{C(z)}$

<table>
<thead>
<tr>
<th>Constructions</th>
<th>EGF Equations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C = Z \star SET(C)$</td>
<td>$C(z) = ze^{C(z)}$</td>
</tr>
<tr>
<td>$M = SET(CYC_{\geq 1}(C))$</td>
<td>$M(z) = \exp\left(\ln \frac{1}{1 - C(z)} - C(z)\right)$</td>
</tr>
<tr>
<td>$C(z) = ze^{C(z)}$</td>
<td>$M(z) = \exp\left(\ln \frac{1}{1 - C(z)} - C(z)\right)$</td>
</tr>
<tr>
<td>$M(z) = \exp\left(\ln \frac{1}{1 - C(z)} - C(z)\right)$</td>
<td>$= \frac{e^{-C(z)}}{1 - C(z)}$</td>
</tr>
</tbody>
</table>
Q. Find the probability that a random mapping has no singleton cycles.

A. Each entry can have any value but its own index, so the number of $N$-mappings with no singleton cycles is $(N - 1)^N$

\[
\frac{(N - 1)^N}{N^N} = \left(1 - \frac{1}{N}\right)^N
\]

\[
\sim \frac{1}{e}
\]
<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wednesday March 17</td>
<td>AofA review posted</td>
<td>11:59AM</td>
</tr>
<tr>
<td>Thursday March 18</td>
<td>AofA review due</td>
<td>11:59PM</td>
</tr>
<tr>
<td>Friday March 19</td>
<td>AC week 1 posted</td>
<td></td>
</tr>
<tr>
<td>Wednesday April 28</td>
<td>AC review posted</td>
<td>11:59AM</td>
</tr>
<tr>
<td>Thursday April 29</td>
<td>AC review due</td>
<td>11:59PM</td>
</tr>
</tbody>
</table>