Algorithms Seminar 2000–2001,
F. Chyzak (ed.), INRIA, (2002), pp. 15–18.

Available online at the URL
http://algo.inria.fr/seminars/.

The Tennis Ball Problem

Donatella Merlini
DSI, Università degli Studi di Firenze (Italy)

March 19, 2001

Summary by Cyril Banderier

Abstract

Our object is to explore the "s-tennis ball problem" (at each round s balls are available and we play with one ball at a time). This is a natural generalization of the case $s = 2$ considered by Mallows and Shapiro. We show how this generalization is connected with s-ary trees and employ the notion of generating trees to obtain a solution expressed in terms of generating functions. Then, we present a variation in which at each round we have 4 balls and play with 2 balls at a time. To solve this problem we use the concepts of Riordan arrays and stretched Riordan arrays, and a generalization of generating trees. This is a joint work by D. Merlini with D. G. Rogers, R. Sprugnoli and M. C. Verri.

1. Introduction

Let $1 \leq t < s$ be two integer numbers. A tennis player begins a match with 0 ball in the pocket. At each round, he is given s new balls, that he puts in the pocket, and throws away t balls, and so on until the nth round. The balls are labelled from 1 to sn and are served in increasing order. The tn balls thrown away form a sequence of tn labels. Two sequences which are equal once sorted are considered equivalent. The tennis ball problem consists in evaluating the following two quantities: the number f_n of nonequivalent configurations after n rounds and the cumulative sum Σ_n (i.e., the sum—over all the possible configurations—of the labels of the tn balls that the player threw away).

Turns	Balls received	Balls in the pocket	Balls thrown away
$n = 1$	1 and 2	1 and 2	1
$n = 2$	3 and 4	2, 3, and 4	3
$n = 3$	5 and 6	2, 4, 5, and 6	2
$n = 4$	7 and 8	4, 5, 6, 7, and 8	6
			sum $= 1 + 3 + 2 + 6 = 12$

Turns	Balls received	Balls in the pocket	Balls thrown away
$n = 1$	1 and 2	1 and 2	2
$n = 2$	3 and 4	1, 3, and 4	3
$n = 3$	5 and 6	1, 4, 5, and 6	4
$n = 4$	7 and 8	1, 5, 6, 7, and 8	1
			sum $= 2 + 3 + 4 + 1 = 10$

FIGURE 1. Two scenarios for the $(s = 2, t = 1)$-tennis ball player.

The configuration after 4 rounds is $(1,2,3,6)$ for the first example and $(1,2,3,4)$ for the second example. In fact, for the $(2,1)$-case, one has $f_1 = 2$, $f_2 = 5$, $f_3 = 14$, $f_4 = \ldots$ do you guess what? There is indeed 42 different configurations (after 4 rounds), and if one adds all the sums, one gets $\Sigma_1 = 1+2 = 3$, $\Sigma_2 = (1+2)+(1+3)+(1+4)+(2+3)+(2+4) = 23$, $\Sigma_3 = 131$, $\Sigma_4 = 664, \ldots$

In the next section, it is shown how the $(s,1)$-case can be solved in terms of s-ary trees (by symmetry, this also solves the $(s,s-1)$-case). Then, the last section is dedicated to the $(4,2)$-case, that the authors solved with Riordan arrays and a bilabelled generating tree technique.

Turns	Balls received	Balls in the pocket	Balls thrown away
$n=1$	1, 2, 3, 4	1, 2, 3, 4	2, 3,
$n=2$	5, 6, 7, 8	1, 4, 5, 6, 7, 8	1, 7
$n=3$	9, 10, 11, 12	4, 5, 6, 8, 9, 10, 11, 12	10, 12
$n=4$	13, 14, 15, 16	4, 5, 6, 8, 9, 11, 13, 14, 15, 16	5, 16
			$2+3+1+7+10+12+5+16=56$

FIGURE 2. A scenario for the $(4,2)$-tennis ball problem.

This is the only case solved with $t \neq 1$. The general (s,t)-tennis ball problem remains open.

2. The $(s,1)$-Tennis Ball Problem

Generating trees are a convenient way to reexpress the problem. Consider an infinite rooted tree $\mathcal{T}$. The root (labelled 0 and corresponding to level 0) has t children (labelled $1,\ldots,t$). Each path in this tree corresponds to a scenario, thus each node at level n has a label which corresponds to the ball thrown away at round n. As we are counting the *sorted* configurations (that is, one does not care for the order of the balls thrown away), we can without loss of generality suppose that the labels increase with the depth.

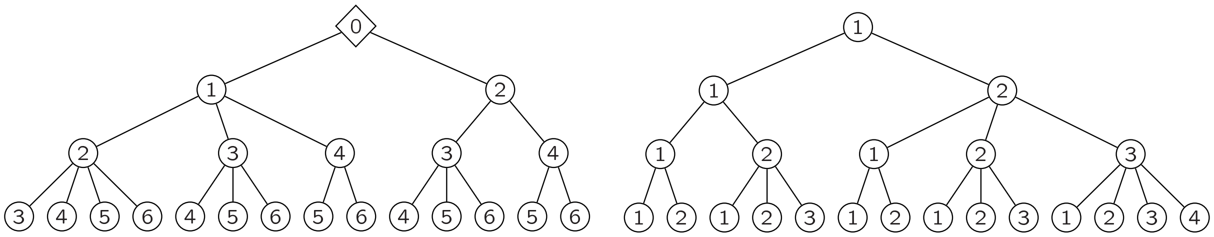

FIGURE 3. The generating tree $\mathcal{T}$ for the $(2,1)$-case and an isomorphic tree $\widetilde{\mathcal{T}}$.

More generally, the rewriting rule $\begin{cases} \text{root} : (1) \\ \text{rule} : (k) \mapsto (1)\ldots(k+s-2)\ldots(k+s-1) \end{cases}$ describes the formation of a tree $\widetilde{\mathcal{T}}$ which is isomorphic to the generating tree $\mathcal{T}$ of the $(s,1)$-case: a node with label b at level i in the generating tree $\mathcal{T}$ becomes a node with label $si-b+1$ in the tree $\widetilde{\mathcal{T}}$.

Theorem 1. *The number f_n of configurations for the $(s,1)$-tennis ball problem is the number T_{n+1} of s-ary trees with $n+1$ nodes. One has*

$$T_n = \frac{\binom{sn}{n}}{1+(s-1)n} \quad \text{and} \quad T(z) = 1 + zT(z)^s.$$

Proof. The problem can be seen as the enumeration of walks on the integers (with an unbounded set of jumps described by the rewriting rule), for which the generating function can be made explicit [1]. Merlini *et al.* used Riordan array techniques [4]. □

Theorem 2. *The cumulative sum (i.e., the sum over all the configurations of the labels of the thrown balls) is*

$$\Sigma_{n-1} = \frac{sn^2 + (s-1)n + 1}{2} \frac{\binom{sn}{n}}{(s-1)n+1} - \frac{1}{2} \sum_{k=0}^{n} \binom{sk}{n} \binom{s(n-k)}{n-k}.$$

Proof. Consider $A_n = \sum_{i=0}^{n} \ell_{i,n}$, the sum of all the labels (with multiplicity) at level i in the tree $\mathcal{T}$. The cumulative sum Σ_n satisfies

$$\Sigma_n = A_n - \frac{(sn+2)(n+1)}{2} T_{n+1}.$$

The generating function for the sequence A_n is:

$$A(z) = \frac{s(s-1)zT'(z)^2}{2T(z)} + T'(z).$$

From these two equations, one gets the almost closed form of the theorem.

Note that the asymptotics of Σ_n can easily be deduced from the asymptotics of A_n. □

These theorems are consistent with the fact that the $(2,1)$-case leads to Catalan numbers $f_n = \frac{\binom{2n}{n}}{n+1}$ (proven in [2]) and to $\Sigma_n = \frac{2n^2+5n+4}{n+2}\binom{2n+1}{n} - 2^{2n+1}$ (as it was found in [3] by hand manipulations of sums of binomial coefficients).

3. The $(4,2)$-Tennis Ball Problem

Here again, as one does not care for the order (of the balls thrown away), one can without loss of generality suppose that any configuration is represented by the smallest equivalent sequence with respect to lexicographical order. Thus the configuration $(1,4),(5,8),(2,10)$ is considered to be the same as the configuration $(1,2),(4,5),(8,10)$.

Let $M_m^{[n]}$ be the number of pairs at level n (in the bilabelled generating tree of the $(4,2)$-case) with larger element equal to m; one has the recurrence

$$M_m^{[n+1]} = \sum_{r=2n}^{m-2} (m-r-1) M_r^{[n]}.$$

Defining $f_{n,k} = M_{4n+1-k}^{[n]}$ gives an infinite lower-triangular array:

n/k	1	2	3	4	5	6	7	8	9
0	1								
1	3	2	1						
2	22	16	10	4	1				
3	211	158	105	52	21	6	1		
4	2306	1752	1198	644	301	116	36	8	1

One has the relation $f_{n+1,k+2} = \sum_{j=0}^{\infty} (j+1) f_{n,k+j}$. The sums $f_n = \sum_{k\geq 1} f_{n,k}$ give the sequence $(1, 6, 53, 554, \dots)$, the number of configurations for the $(4,2)$-tennis ball problem.

It is convenient to transform the above array into a proper Riordan array. A proper Riordan array is an infinite lower triangular array $(D_{n,k})_{n,k\in\mathbb{N}}$ which satisfies

$$d_{n+1,k+1} = \sum_{j=0}^{\infty} a_j d_{n,k+j} \qquad \text{for all } n \text{ and } k \text{ in } \mathbb{N}.$$

The generating function $A(z) = \sum_j a_j z^j$ allows to express $d_{n,k}$ by a Lagrangean-like formula

$$d_{n,k} = [z^n] g(z) \big(zh(z)\big)^k \text{ where } h(z) = A\big(zh(z)\big).$$

The above array can be embedded in the array

n/k	0	1	2	3	4	5	6	7
0	1							
1	0	1						
2	1	1	1					
3	0	**3**	**2**	**1**				
4	6	6	6	3	1			
5	0	**22**	**16**	**10**	**4**	**1**		
6	53	53	53	31	15	5	1	

which satisfies $A(z) = \frac{1}{1-z}$, $h(z) = C(z)$ (the generating function of Catalan numbers), and $g(z) = \frac{2}{2-zC(z)+zC(-z)}$. In particular, the function $g(z)$ generates the first column of this array and corresponds to the number of nonequivalent configurations one wants to enumerate: $f_n = g_n = \frac{3(n+2)}{(n+3)((2n+3)}\binom{2n+4}{n+2} - \frac{4^{(n+1)/2}}{n+2}\binom{n+3}{(n+3)/2}$ (for even n). The cumulative sum in the tree with root $(0,0)$ and rewriting rule $(k_1,k_2) \mapsto (0,0)(0,1)\dots(k_1+2,k_1+2)$ is then given by

$$\Sigma_n = \sum_{h=0}^{n} \sum_r \mu_r^{[2n-2h]} w_r^{[2h]} = \sum_r \sum_{h=0}^{n} \mu_r^{[2n-2h]} w_r^{[2h]}.$$

Here, $\mu_r^{[2n-2h]}$ is the number of nodes at level $n-h$ in the subtree starting with $(r,*)$, and $w_r^{[2h]}$ the total weight that the couples $(r,*)$ have at level h. (Note that a label (k_1,k_2) at level n in the new tree corresponds to a label $(4n-k_2-1, 4n-k_1)$ in the generating tree of the $(4,2)$-case.) The Riordan array property yields $\mu_r(z) = g(z)C(z)^{r+2}$ and $w_r(z) = g(z)z^r C(z)^{r+1}(zC(z)^2+2r)$, thus

$$\Sigma(z) = \frac{1}{4}\sum_r \big(\mu_r(z)+\mu_r(-z)\big)\big(w_r(z)+w_r(-z)\big) = 12z^2 + 284z^4 + 5436z^6 + 96768z^8 + O(z^{10}).$$

The next nontrivial open cases are the $(5,2)$- and $(5,3)$-tennis ball problems. This is related to the enumeration of 2- and 3-dimensional constrained discrete random walks for which no closed form (or even recurrence) is known. Articles and slides related to this summary can be found at Donatella Merlini's web page `http://www.dsi.unifi.it/~merlini/Publications.html`.

Bibliography

[1] Banderier (C.), Bousquet-Mélou (M.), Denise (A.), Flajolet (P.), Gardy (D.), and Gouyou-Beauchamps (D.). – Generating functions for generating trees. *Discrete Mathematics*. – 25 pages. To appear.

[2] Grimaldi (Ralph P.) and Moser (Joseph G.). – The Catalan numbers and a tennis ball problem. In *Proceedings of the Twenty-eighth Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1997)*, vol. 125, pp. 65–71. – 1997.

[3] Mallows (Colin L.) and Shapiro (Lou). – Balls on the lawn. *Journal of Integer Sequences*, vol. 2, 1999. – Article 99.1.5.

[4] Merlini (D.), Rogers (D. G.), Sprugnoli (R.), and Verri (M. C.). – The tennis ball. – 2001. Submitted.