

ON THE DISTRIBUTION OF WINNING MOVES IN RANDOM GAME TREES

J.A. FLANIGAN

In a natural way, associated with each rooted tree there exists a pair of two-person games, each game possessing the root as initial position. When a rooted tree is selected at random from the set of all rooted trees which possess $\{1, 2, \dots, n\}$ as vertex set, the number of winning moves available to the first player to move in each of the associated games is a random variable. For fixed n , we determine the distribution of this random variable. As an immediate consequence, we find the probability that the first player to move has no winning move at all. The "saddlepoint method" is applied to a certain contour integral to obtain the asymptotic distribution of the number of winning moves as $n \rightarrow \infty$.

1. Introduction

Let T_n denote the set of rooted trees which have $V_n = \{1, 2, \dots, n\}$ as vertex set. To each $T_n \in T_n$ associate the following two-person game. The vertices of T_n are the positions of the game, the root being the initial position. The edges represent the legal moves, provided that moves are made in the direction away from the root. The players alternate moves. Play ends when a player is unable to move on his turn. In this paper, except in Section 5, the normal play rule is adopted: a player unable to move on his turn is the loser. These rules of

Received 3 April 1981.

play determine a partition of the nodes of T_n into two sets.

One set is composed of those positions from which the first player to move can force a win. The other set is composed of those positions from which the second player to move can force a win. For normal play, a position represents a first-player win if and only if there is a move to a second-player win position. Such a move is a winning move. If the edges of T_n are oriented away from the root, then the kernel of the resulting digraph is the set of second-player win positions ([1], Chapter 14). It is also the set of positions with normal Grundy number equal to zero ([1], Chapter 14). It can be found by a simple inductive procedure.

Suppose that T_n is selected at random from \mathcal{T}_n . Let the random variable $X_n(T_n)$ [$Y_n(T_n)$] equal the number of moves [winning moves] from the root of T_n . The distribution of X_n and its limiting distribution as $n \rightarrow \infty$ are well-known [see formulae (1) and (11)]. The main purpose of this note is to determine the (limiting) joint distribution of X_n and Y_n (Sections 2, 3, 4). From this, some other probabilities of interest can be found immediately (Corollaries 3.1 to 3.4). Finally (Section 5), we determine the (limiting) joint distribution of X_n and Y_n when the normal play rule is replaced by the misère play rule: the player unable to move on his turn is the winner.

2. Notation

For $n \geq 1$ and $0 \leq p \leq k < n$, set

$$t_n(k, p) = |\{T_n \in \mathcal{T}_n \mid \text{there are } k \text{ moves from the root of } T_n, \\ p \text{ of which are winning moves}\}|,$$

$$t_n(k, \cdot) = |\{T_n \in \mathcal{T}_n \mid \text{there are } k \text{ moves from the root of } T_n\}|,$$

$$t_n(\cdot, p) = |\{T_n \in \mathcal{T}_n \mid \text{there are } p \text{ winning moves from the} \\ \text{root of } T_n\}|,$$

$$t_n(\cdot, \cdot) = |\mathcal{T}_n|,$$

$$l(x) = \sum_{n=1}^{\infty} t_n(\cdot, 0)(x^n/n!) , \quad t(x) = \sum_{n=1}^{\infty} t_n(\cdot, \cdot)(x^n/n!) ,$$

$$\Phi(x, y, z) = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} \sum_{p=0}^k t_n(k, p)(x^n/n!)y^kz^p .$$

Note that $t_n(\cdot, 0)$ represents the number of trees in T_n whose roots are second-player wins. It is well-known [2] that $t_n(\cdot, \cdot) = n^{n-1}$ and that, [3], $t_n(k, \cdot) = n \binom{n-2}{k-1} (n-1)^{n-k-1}$ for $1 \leq k \leq n-1$. Thus,

$$(1) \quad P[X_n = k] = \binom{n-2}{k-1} (1-(1/n))^{n-2-(k-1)} (1/n)^{k-1} .$$

$X_n - 1$ has a binomial distribution with parameters $n - 2$ and $1/n$.

3. Enumerative analysis

In this section, first we find expressions for the generating functions $\sum_{n=k+1}^{\infty} t_n(k, p)(x^n/n!)$ (Corollary 1.1) and

$\sum_{n=p+1}^{\infty} t_n(\cdot, p)(x^n/n!)$ (Corollary 1.4) in terms of the generating functions $t(x)$ and $l(x)$ for all k and p . Then an explicit formula for $t_n(\cdot, 0)$ is found (Theorem 2).

THEOREM 1.

$$(2) \quad \Phi(x, y, z) = xe^y [t(x) - (z-1)l(x)] .$$

Proof. For $n > k > 0$ and $0 \leq p \leq k$, consider

$$T_n(k, p) = \{T_n \in T_n \mid \text{there are } k \text{ moves from the root of } T_n, \\ p \text{ of which are winning moves}\} .$$

There is a one-to-one correspondence between members of $T_n(k, p)$ and forests whose vertex sets are subsets of V_n of size $n - 1$ and which consist of k rooted trees, p of which possess roots which are second-player wins. There are n ways to select a subset of V_n containing

$n - 1$ elements. Given any such subset as vertex set, by an elementary counting argument, the number of ways to form a forest of k rooted trees, p of which possess roots which are second-player wins is

$$\sum_{i=p}^{(n-1)-(k-p)} \binom{n-1}{i} \left[\frac{1}{p!} \sum_{j_1+\dots+j_p=i} \binom{i}{j_1 \dots j_p} l_{j_1} \dots l_{j_p} \right] \\ \times \left[\frac{1}{(k-p)!} \sum_{j_1+\dots+j_{k-p}=(n-1)-i} \binom{(n-1)-i}{j_1 \dots j_{k-p}} w_{j_1} \dots w_{j_{k-p}} \right]$$

where $l_n = t_n(\cdot, 0)$ and $w_n = n^{n-1} - l_n$ for $n \geq 1$. Thus

$$t_n(k, p) = (n!/k!) \binom{k}{p} \sum_{i=p}^{(n-1)-(k-p)} \left[\sum_{j_1+\dots+j_p=i} (l_{j_1} \dots l_{j_p}) / (j_1! \dots j_p!) \right] \\ \times \left[\sum_{j_1+\dots+j_{k-p}=(n-1)-i} (w_{j_1} \dots w_{j_{k-p}}) / (j_1! \dots j_{k-p}!) \right] \\ = (1/k!) \{ \text{coefficient of } (x^n/n!)z^p \text{ in } x[t(x)+(z-1)l(x)]^k \},$$

and so

$$(3) \quad x(y^k/k!) [t(x)+(z-1)l(x)]^k = y^k \sum_{n=k+1}^{\infty} \sum_{p=0}^k t_n(k, p) (x^n/n!)z^p.$$

Since $t_1(0, 0) = 1$, (3) holds for all $k \geq 0$. The proof is completed by summing both sides of (3) over $0 \leq k < \infty$ and then changing the order of summation on the right hand side. \square

COROLLARY 1.1. For $k \geq 1$ and $0 \leq p \leq k$,

$$(4) \quad \sum_{n=k+1}^{\infty} t_n(k, p) (x^n/n!) = (x/k!) \binom{k}{p} l(x)^p [t(x)-l(x)]^{k-p}.$$

Proof. Observe that the left hand side of (4) is $(1/k!p!) (\partial^p/\partial z^p) (\partial^k/\partial y^k) \Phi(x, y, z)|_{y=z=0}$ which can be calculated from (2). \square

COROLLARY 1.2 [6].

$$(5) \quad t(x) = xe^{t(x)}.$$

Proof. Set $y = z = 1$ in (2). \square

COROLLARY 1.3.

$$(6) \quad t(x) = \mathcal{L}(x)e^{\mathcal{L}(x)} .$$

Proof. Set $y = 1$, $z = 0$ in (2) and then apply (4). \square

COROLLARY 1.4. For $p \geq 0$,

$$(7) \quad \sum_{n=p+1}^{\infty} t_n(\cdot, p)(x^n/n!) = \mathcal{L}(x)^{p+1}/p! .$$

Proof. Observe that the left hand side of (7) is $(1/p!)(\partial^p/\partial z^p)\Phi(x, 1, z)|_{z=0}$ which can be calculated by using (2). Then use (6) to get (7). \square

The remainder of this section is devoted to determining $\mathcal{L}(x)$. Apply the Lagrange inversion formula ([7], p. 135) to (5) to get Cayley's result:

$$t(x) = \sum_{n=1}^{\infty} n^{n-1}(x^n/n!) .$$

It is well-known ([5], p. 344) that this power series converges for all complex z satisfying $|z| \leq e^{-1}$, where it is also the inverse of ze^{-z} . Thus, for $|z| \leq e^{-1}$, in addition to (5), we have

$$(8) \quad t(ze^{-z}) = z .$$

In the sequel, $t(z)$ will represent the analytic continuation of this power series from $\{z \mid |z| < e^{-1}\}$ to $\Omega = \mathbb{C} - \{r \in \mathbb{R} \mid r \geq e^{-1}\}$. Relations (5) and (8) remain valid in Ω .

LEMMA.

$$(9) \quad \mathcal{L}(x) = -t(-t(x)) .$$

Proof. By (6) and (8), $x = \mathcal{L}(xe^{-x})/e^{-\mathcal{L}(xe^{-x})}$. Use the Lagrange inversion formula to get $\mathcal{L}(xe^{-x}) = \sum_{n=1}^{\infty} (-1)^{n-1} n^{n-1}(x^n/n!) = -t(-x)$. By (5), $\mathcal{L}(x) = -t(-t(x))$. \square

THEOREM 2. For $n \geq 1$, $t_n(\cdot, 0) = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} k^n n^{-k-1}$.

Proof. Note that $t(0) = 0$ and $(d/dz)t(z)|_{z=0} = 1$. There exists an open disc Δ about the origin of \mathbb{C} in which t is a homeomorphism. Let C be any circle contained in Δ with the origin as center. Then $t(C)$ is a simple closed curve around the origin.

For $n \geq 1$, by the Cauchy integral formula, (5) and (9),

$$\begin{aligned} t_n(\cdot, 0) &= (n-1)!/2\pi i \oint_C ((d/dz)t(z)/z^n) dz \\ &= (n-1)!/2\pi i \oint_C [-(d/dz)t(-t(z))/[t(z)e^{-t(z)}]^n] dz . \end{aligned}$$

Make the change of variable $w = t(z)$ to get

$$\begin{aligned} t_n(\cdot, 0) &= (n-1)!/2\pi i \oint_{t(C)} [-(d/dw)t(-w)] e^{nw}/w^n dw \\ &= (n-1)! \left\{ \text{coefficient of } w^{n-1} \text{ in } \left[\sum_{k=0}^{\infty} (-1)^k (k+1)^k (w^k/k!) \right] \left[\sum_{k=0}^{\infty} (n^k/k!) w^k \right] \right\} \\ &= \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} k^n n^{-k-1} . \quad \square \end{aligned}$$

Theorem 2 can also be proved by using (9) and the following result (proved on pp. 181-182 of [7]) which will be needed in Section 5:

$$(10) \quad t(x)^k = \sum_{n=k}^{\infty} kn^{n-k-1} (n)_k (x^n/n!) .$$

It will be seen in Corollary 3.3 that

$$t_n(\cdot, 0) \sim n^{n-1} (d/dx)t(x)|_{x=-1} = (.362^-)n^{n-1} .$$

4. Asymptotic analysis

For large n , it is not practical to use the results of Section 3 to make exact computations. Observe from (1) that

$$(11) \quad \lim_{n \rightarrow \infty} P[X_n = k] = e^{-1}/(k-1)! .$$

The limiting distribution of $X_n - 1$ is Poisson with mean one. The main task of this section is to determine the limiting joint distribution of X_n and Y_n . Here α will denote $-t(-1) = .567^+$.

THEOREM 3. For $k \geq 1$ and $0 \leq p \leq k$,

$$(12) \quad \lim_{n \rightarrow \infty} P[X_n = k, Y_n = p] = (1/k!e) [(k-\alpha p)/(1-\alpha^2)] \binom{k}{p} \alpha^p (1-\alpha)^{k-p} .$$

Proof. Let C be as in the proof of Theorem 2, $n > k > 0$, $0 \leq p \leq k$, and $A = n \binom{k}{p} [(n-2)!]/k!$. Use (4), Cauchy's integral formula, (5), and (9) to get

$$\begin{aligned} t_n(k, p) &= A/2\pi i \oint_C \left\{ (d/dz) \{ [l(z)]^p [t(z)-l(z)]^{k-p} / z^{n-1} \} dz \right. \\ &= A/2\pi i \oint_C \left\{ (d/dz) \{ [-t(-t(z))]^p [t(z)+t(-t(z))]^{k-p} / t(z)^{n-1} e^{-(n-1)t(z)} \} dz \right. . \end{aligned}$$

Now make the change of variable $w = t(z)$ to get

$$t_n(k, p) = A/2\pi i \oint_{t(C)} \varphi(w) e^{h(n,w)} dw$$

where

$$\begin{aligned} \varphi(w) &= p[-t(-w)]^{p-1} [w+t(-w)]^{k-p} (d/dw)[-t(-w)] \\ &\quad + (k-p)[-t(-w)]^p [w+t(-w)]^{k-p-1} [1+(d/dw)t(-w)] \end{aligned}$$

and

$$h(n, w) = (n-1)[w-\ln(w)] .$$

Note that $e^{h(n,w)}$ has a saddlepoint at $w_0(n) = 1$. $\varphi(w)$ and $h(n, w)$ satisfy certain conditions ([4], pp. 46-47) which permit the application of the saddlepoint method to get an asymptotic estimate for $t_n(k, p)$. According to a result of ([4], p. 47),

$$1/2\pi i \oint_{t(c)} \varphi(w) e^{h(n,w)} dw \sim \varphi(w_0(n)) e^{h(n,w_0(n))} / \sqrt{2\pi(d^2/dw^2)h(n,w)|_{w=w_0(n)}} .$$

Thus

$$t_n(k, p)/t_n(\cdot, \cdot) \sim (A/n^{n-1}) (\varphi(1)e^{n-1/\sqrt{2\pi(n-1)}}) = n \binom{k}{p} [(n-1)!] \varphi(1) e^{n-1/n^{n-1}(k!)(n-1)\sqrt{2\pi(n-1)}} .$$

By (5), $(d/dz)t(z)|_{z=-1} = \alpha/(1+\alpha)$. Now apply this fact, Stirling's estimate $(n-1)! \sim (n-1)^{n-1}\sqrt{2\pi(n-1)}e^{-(n-1)}$, and some elementary algebra to complete the proof. \square

COROLLARY 3.1. For $p \geq 0$,

$$(13) \quad \lim_{n \rightarrow \infty} P[Y_n = p] = ((p+1)/(1+\alpha)) (\alpha^{p+1}/p!) .$$

Proof. Sum both sides of (12) over $p \leq k < \infty$ and then note that, by (5), $e^{-\alpha} = \alpha$. \square

We observe from (1) that $EX_n = 2(n-1)/n$, and so $\lim_{n \rightarrow \infty} EX_n = 2$. From

$$(13), \quad \lim_{n \rightarrow \infty} EY_n = \alpha + \alpha/(1+\alpha) = .929^+ .$$

COROLLARY 3.2. For $k \geq 1$ and $0 \leq p \leq k$,

$$(14) \quad \lim_{n \rightarrow \infty} P[Y_n = p | X_n = k] = (1/k) [(k-\alpha p)/(1-\alpha^2)] \binom{k}{p} \alpha^p (1-\alpha)^{k-p} .$$

Proof. Use (11) and (12). \square

From (14) it follows that $\lim_{n \rightarrow \infty} E[Y_n | X_n = k] = k\alpha - \alpha^2/(1+\alpha)$.

Consequently, for large n , suppose T_n is selected at random from T_n , $X_n(T_n) = k$, and the first player to move (perhaps being too lazy to find a winning move) selects a move at random from the k possible moves. Then he will select a winning move with approximate probability $\alpha - \alpha^2/k(1+\alpha)$.

COROLLARY 3.3. $\lim_{n \rightarrow \infty} P[Y_n = 0] = (d/dx)t(x)|_{x=-1}$.

Proof. Set $p = 0$ in (13). Then use $e^{-\alpha} = \alpha$ and $\alpha/(\alpha+1) = (d/dx)t(x)|_{x=-1}$. \square

We note that $(d/dx)t(x)|_{x=-1} = .362^-$.

COROLLARY 3.4. For $k \geq 1$,

$$\lim_{n \rightarrow \infty} P[Y_n = 0 \mid X_n = k] = (1-\alpha)^{k-1}/(1+\alpha) .$$

Proof. Set $p = 0$ in (14). \square

5. Misère analysis

For misère play, a position represents a first-player win if and only if either it is a terminal position or there exists a move to a second-player win position. For the misère analysis, with one exception, we retain the notation of Section 2. Unlike in normal play, in misère play terminal positions are not second-player wins. Thus, in misère play we

take $l(x) = \sum_{n=2}^{\infty} t_n(\cdot, 0)(x^n/n!)$. The misère analysis will parallel that presented for normal play in Sections 3 and 4.

Equations (2) and (4) hold for misère play. Their proofs for misère play are the same as for normal play, except that in the proof of (2), $l_1 = t_1(\cdot, 0) = 1$ is replaced by $l_1 = 0$.

Now set $y = 1$ and $z = 0$ in (2) to get the misère analogue of (6):

$$(15) \quad t(x) = [l(x)+x]e^{l(x)} .$$

To get the misère analogue of (7), use (2) and (15) to get

$$\sum_{n=p+1}^{\infty} t_n(\cdot, p)(x^n/n!) = (1/p!)[l(x)+x]l(x)^p .$$

To determine $l(x)$, rewrite (15) as $e^x t(x) = [l(x)+x]e^{[l(x)+x]}$, or equivalently, as $x = y/e^{-y}$, where $y = l(\varphi^{-1}(x)) + \varphi^{-1}(x)$, $\varphi(x) = e^x t(x)$. Then apply the Lagrange inversion formula to get the

misère analogue of (9):

$$(16) \quad l(x) = -t(-e^x t(x)) - x .$$

THEOREM 4. For $n \geq 2$,

$$t_n(\cdot, 0) = \sum_{k=1}^n (-1)^{k-1} \sum_{l=k}^n \binom{n}{l} \binom{l}{k} k^{n+k-l} l^{l-k-1} .$$

Proof. Apply (10) to (16).

The proofs of the remaining results are similar to the proofs of their normal play analogues. Set

$$\beta = -t(-e^{-1}) - e^{-1} \quad \text{and} \quad \gamma = e^{-1} (d/dx)t(x) \Big|_{x=-e^{-1}} .$$

$$\beta = .342^+ \quad \text{and} \quad \gamma = .415^+ .$$

THEOREM 5. For $k \geq 1$ and $0 \leq p \leq k$,

$$\lim_{n \rightarrow \infty} P[X_n = k, Y_n = p] = (1/k!e) [(\beta(1-\gamma)k + (\gamma-\beta)p) / \beta(1-\beta)] \binom{k}{p} \beta^p (1-\beta)^{k-p} .$$

COROLLARY 5.1. For $p \geq 0$,

$$(17) \quad \lim_{n \rightarrow \infty} P[Y_n = p] = ((\gamma p / \beta) + 1 - \gamma) (\beta^p e^{-\beta} / p!) .$$

From (17) it follows that for misère play $\lim_{n \rightarrow \infty} EY_n = \beta + \gamma = .757^+$.

COROLLARY 5.2. For $k \geq 1$ and $0 \leq p \leq k$,

$$(18) \quad \lim_{n \rightarrow \infty} P[Y_n = p \mid X_n = k] \\ = (1/k) [(\beta(1-\gamma)k + (\gamma-\beta)p) / \beta(1-\beta)] \binom{k}{p} \beta^p (1-\beta)^{k-p} .$$

From (18) it follows that $\lim_{n \rightarrow \infty} E[Y_n \mid X_n = k] = (k-1)\beta + \gamma$.

COROLLARY 5.3. $\lim_{n \rightarrow \infty} P[Y_n = 0] = \gamma = .415^+$.

COROLLARY 5.4. For $k \geq 1$,

$$\lim_{n \rightarrow \infty} P[Y_n = 0 \mid X_n = k] = (1-\gamma)(1-\beta)^{k-1} .$$

References

- [1] Claude Berge, *Graphes et hypergraphes* (Monographies universitaires de mathématiques, 37. Dunod, Paris, 1970).
- [2] [A.] Cayley, "A theorem on trees", *Quart. J. Pure Appl. Math.* 23 (1889), 376-378.
- [3] L.E. Clarke, "On Cayley's formula for counting trees", *J. London Math. Soc.* 33 (1958), 471-474.
- [4] M.A. Evgrafov, *Asymptotic estimates and entire functions* (translated by Allen L. Shields. Russian Tracts on Advanced Mathematics and Physics, 4. Gordon and Breach, New York, 1961).
- [5] А.И. Марнушевич [A.I. Markuševič], *Теория аналитических функций* [*Theory of analytic functions*] (Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, Leningrad, 1950).
- [6] G. Pólya, "Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen", *Acta Math.* 68 (1937), 145-254.
- [7] В.Н. Сачнов [V.N. Sačkov], *Комбинаторные методы дискретной математики* [*Combinatorial methods in discrete Mathematics*] ("Наука", Moscow, 1977).

Department of Mathematics,
University of California,
Los Angeles,
California 90024,
USA.