BERLIN 2005

ERDŐS MAGIC

Erdős 1947: If $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ there **exists** a two coloring of the edges of K_n with no monochromatic K_k.

Proof: Color Randomly!

Theorem (Turán). Any graph G has an independent set S with

$$|S| \geq \sum_{v \in G} \frac{1}{d_v + 1}$$

Randomized Algorithm

- Order Vertices Randomly.
- Place v in S greedily.

If v comes before its neighbors *then* it goes into S.

$$\Pr[v \in S] \geq \frac{1}{d_v + 1}$$

Linearity of Expectation:

$$E[|S|] = \sum_{v \in G} \Pr[v \in S] \geq \sum_{v \in G} \frac{1}{d_v + 1}$$

Erdős Magic: Such S **MUST** exist.

$|A_i| = n$, $1 \le i \le m = 2^{n-1}k$

Seek Red/Blue χ with no A_i monochromatic

Erdős [1963]: $k < 1 \Rightarrow \exists \chi$

Beck [1978]: $k < cn^{1/3} \Rightarrow \exists \chi$

Radhakrishnan-Srinivasan[2000]

$$k < c[n/\ln n]^{1/2} \Rightarrow \exists \chi$$

Erdős [1964]: There exists family with $k = cn^2$ with no χ

Coloring Algorithm(s)

1. Color Randomly

2. Order Vertices Randomly.

3. Consider sequentially. If v “still dangerous” switch $\chi(v)$ with probability p

Still Dangerous: $v \in A_i$ which has *always* been monochromatic

FAIL: Some A_i monochromatic at end

Erdős Magic: If Pr$[FAIL] < 1$ χ **MUST** exist

Two Failure Modes

FAILI: A_i was “Red” and stayed Red

FAILII: A_i wasn’t Red and became Red

$$\Pr[FAILI(A_i)] = 2^{1-n}(1-p)^n$$

$$\Pr[FAILI] \leq (2^{n-1}k)(2^{1-n}(1-p)^n) = k(1-p)^n$$

A_i *blames* A_j if

- $A_i \cap A_j = \{v\}$
- A_j Blue at start
- A_i Red at end
- v **LAST** point of A_i to change
- When v reached A_j all Blue

Theorem:

If FAILII then some A_i blames some A_j

Corollary:

$$\Pr[FAILII] \leq \sum_{i \neq j} \Pr[A_i \text{ blames } A_j]$$

Bounding Pr[A_i blames A_j]

Fix ordering.

- Factor 2 for Red/Blue symmetry
- v Blue and Flips: $p/2$
- $w \in A_j$ after v: 1/2
- $u \in A_i$ after v: 1/2
- $w \in A_j$ before v: $1/2 - p/2$
- $u \in A_i$ before v: $1/2 + p/2$

I: Number of $w \in A_i$ before v

J: Number of $u \in A_j$ before v

$$\Pr[A_i \text{ blames } A_j | I, J] = 2^{2-2n} p (1+p)^{|I|} (1-p)^{|J|}$$

A Bad Gamble

$n-1$ Red Cards, $n-1$ Blue Cards, Joker

Shuffle. Start with 1000 Euro

Red: Multiply funds by $1+p$

Blue: Multiply funds by $1-p$

Joker: Cash In.

Theorem: Expectation less than initial

Corollary

$$\Pr[A_i \text{ blames } A_j] \leq 2^{2-2n}p$$

Corollary

$$\Pr[FAILII] \leq (2^{n-1}k)^2 2^{2-2n} p = k^2 p$$

Asymptotic Calculus

Pr$[FAIL] < k(1-p)^n + k^2p$

Erdős Magic: If for some $p \in [0,1]$

$$k(1-p)^n + k^2p < 1 \text{ (*)}$$

then χ **MUST** exist.

What is max $k = k(n)$ so that (*) holds for some $p \in [0,1]$?

Answer: $k \sim c\sqrt{n/\ln n}$

Liar Game

Paul seeks $x \in \{1, \ldots, 100\}$.

Ten Queries. Carole may lie once.

Theorem: Carole Wins!

Carole plays randomly

At end of game:

Pr[x possible] $= \frac{11}{1024}$

Expected number of possible $100\frac{11}{1024} > 1$

When > 1 possible Carole wins

Carole sometimes wins

Erdős Magic: Carole always wins!

Counting Connected Graphs

and the
Giant/Dominant Component

Joint with

Remco van der Hofstad (EURRANDOM)

Nitin Arora (Courant)

Complexity = E-V+1

$C(k,l)$ = Number of

CONNECTED Labelled Graphs

k Vertices

Complexity l

$C(k,0) = k^{k-2}$ Cayley

$C(k,l) \sim c_l k^{3l/2} k^{k-2}$ Wright

$l > (\frac{1}{2}+\epsilon)k \ln k$ Erdős-Rényi

$k, l \to \infty$ Bender, Canfield, McKay

The Łuczak Gem

$X =$ number of (k, l) components in $G(n, p)$

$E[X] = C(k, l)\binom{n}{k}p^{k+l-1}(1-p)^{k(n-k)+\binom{k}{2}-(k+l-1)}$

$X \leq \frac{n}{k}$ tautologically

$C(k, l) \leq \frac{n}{k}[\binom{n}{k}p^{k+l-1}(1-p)^{k(n-k)+\binom{k}{2}-(k+l-1)}]^{-1}$

for *all* n, p.

Minimize using Asymptotic Calculus!

If $l = \Theta(k)$, $p \sim cn^{-1}$, $c > 1$.

Pick n, p so Giant most likely (k, l)

$E[X] = \Omega(n^{-2})$ (even better)

$C(k, l)$ within n^2 Factor

Tilted Balls in Bins

$k-1$ balls, k bins, $p \in (0,1]$

Truncated Geometric

Ball j in Bin T_j

$$\Pr[T_j = i] = \frac{p(1-p)^{i-1}}{1-(1-p)^k}$$

Z_i balls in bin i

$Y_0 = 1$, $Y_i = Y_{i-1} + Z_i - 1$ (so $Y_k = 0$)

TREE: $Y_t > 0$, $0 \le t < k$

$$M := \sum_{i=0}^{k} (Y_i - 1) = \binom{k}{2} - \sum_{j=1}^{k} T_j$$

$G(k,p)$ Random Graph

Vertices $0,1,\ldots,k-1$

Adjacency Prob p

THM: Prob $G(k,p)$ Connected & Complexity l

$= C(k,l)p^{k+l-1}(1-p)^{k(k-1)/2-k-l+1} = A_1 A_2 A_3$

with

$A_1 = (1-(1-p)^k)^{k-1}$

$A_2 = \Pr[\text{TREE}]$

$A_3 = \Pr[BIN[M,p] = l|\text{TREE}]$

Strategy: A_2, A_3 determine $C(k,l)$

Breadth First Search

1	2	3	4	5
N	N	Y	Y	N
N	N	-	-	N
Y	N	-	-	N
-	Y	-	-	Y
-	-	-	-	-
-	-	-	-	-

$T_3 = T_4 = 1$, $T_1 = 3$, $T_2 = T_5 = 4$

A_1: All T_j defined

$\vec{Z} = (2, 0, 1, 2, 0, 0)$

Walk $\vec{Y} = (1, 2, 1, 1, 2, 1, 0)$

TREE: BFS doesn't terminate early

Tree Edges $03, 04, 41, 12, 15$

$M = 2$ Unexposed $34, 25$

Setting the Tilt p

$\mu := E[M]$, $\sigma^2 := Var[M]$

$$p\mu = l$$

Three Regimes

`Small` $l = o(k)$, $k^{-3/2} \ll p \ll k^{-1}$

`Large` $l = \Theta(k)$, $p = \Theta(k^{-1})$

`Very Large` $l \gg k$, $p \gg k^{-1}$

($l > ck \ln k$; $p > c' \frac{\ln k}{k}$ Erdős-Rényi)

`Small`: $k^{-3/2} \ll p \ll k^{-1}$

$\epsilon = \frac{1}{2}pk$

Left Z_i Poisson $1+\epsilon$

Galton-Watson Pr[ESC] $\sim 2\epsilon$

Right $Z_i^* = Z_{k-i}; Y_i^* = Y_{k-i}$

$Y_0^* = 0$, $Y_i^* = Y_{i-1}^* + 1 - Z_i^*$

Z_i^* Poisson $1-\epsilon$

Pr[ESC*] $\sim \epsilon$

*** Scaling for ESC, ESC* is $\epsilon^{-2} \ll k$

Pr[TREE] $\sim$ Pr[ESC] Pr[ESC^*] $\sim 2\epsilon^2$

`Large`: $p \sim \frac{c}{k}$

Left Z_i Poisson $\frac{c}{1-e^{-c}}$

Galton-Watson Pr[ESC] $\sim 1 - e^{-c}$

Right $Z_i^* = Z_{k-i}; Y_i^* = Y_{k-i}$

$Y_0^* = 0$, $Y_i^* = Y_{i-1}^* + 1 - Z_i^*$

Z_i^* Poisson $\frac{ce^{-c}}{1-e^{-c}}$

Pr[ESC*] $\sim 1 - \frac{ce^{-c}}{1-e^{-c}}$

Chernoff: $Y_i > 0$ in middle

Pr[TREE] $\sim$ Pr[ESC] Pr[ESC^*] $\rightarrow 1-(c+1)e^{-c}$

`Very Large` $p \gg k^{-1}$

Chernoff: $Y_i > 0$ all i

Pr[TREE] $\rightarrow 1$

Gaussian M

$M := \binom{k}{2} - \sum_{j=1}^{k} T_j$

Esseen: $\Pr[M < \mu + u\sigma] \rightarrow \Pr[N(0,1) < u]$

*** Still holds *conditional on* $TREE$

Hardest when p barely $\gg k^{-3/2}$

Easy when p `Large`

Trivial when p `Very Large`

From CLT to Local Stats

$E[W_k] = \mu_k$, $\text{Var}[W_k] = \sigma_k^2$, $V_k = BIN[W_k, p_k]$,

$l_k = \mu_k p_k = E[V_k]$

Assume ** $p_k^2 \sigma_k^2 = O(p_k \mu_k)$ **

$(\sigma_k^+)^2 := p_k^2 \sigma_k^2 + p_k \mu_k$

Assume $\sigma_k^{-1}(W_k - \mu_k) \to N(0,1)$

THM: V_k Local CLT, Mean l_k, Var $(\sigma_k^+)^2$

$$\Pr[V_k = l_k] \to \frac{1}{\sqrt{2\pi}\sigma_k^+}$$

Apply to $M|$TREE

$$A_3 = \mathsf{Pr}[BIN[M,p] = l|TREE]$$

`Small`: $p\mu = l = p^2\sigma^2$, $A_3 \sim (4\pi l)^{-1/2}$

`Very Large`: $p\mu = l \gg p^2\sigma^2$, $A_3 \sim (2\pi l)^{-1/2}$

`Large`: $p \sim \frac{c}{k}$. $p^2\sigma^2 = \Theta(l)$, $A_3 \sim g(c)l^{-1/2}$

The Giant/Dominant Component

$G(n, p)$

$p = \frac{c}{n}$, $c > 1$ Erdős-Rényi Giant

$p = \frac{1}{n} + \lambda n^{-4/3}$, $\lambda \to +\infty$

Supercritical: Dominant Component

THM: Probability $C(v)$ has k vertices, complexity l is $\sim A_1^* A_2 A_3$ with

$$A_1^* = \Pr[BIN[n-1, 1-(1-p)^k] = k-1]$$

Corollary: Local Stats for k, l of Giant/Dominant Component. Correlated Gaussians

(Caution: Work in Progress!)

Generating Random Connected Graph

Time $\Theta(K+L)$ (!!) for $L=\Omega(\ln K)$

p with $p\mu = L$

Tilted Balls into Boxes

$L=\Omega(K)$ get BFS Tree with prob. $\Omega(1)$

$L=o(K)$ use Fast Abort.

Add precisely L of M unexposed with prob.

$$\frac{\Pr[\mathrm{BIN}[M,p]=L]}{\max_m \Pr[\mathrm{BIN}[m,p]=L]}$$

Games Mathematicians Play

Paul versus Carole

N Possibilities

Q Yes/No Paul Queries

K (or fewer) Carole Lies

Try it with $N = 100$, $Q = 10$, $K = 1$

Carole plays Adversary Strategy

$\Rightarrow$ Perfect Information

$\Rightarrow$ Winning Strategy for Paul or Carole

$B_K(Q) =$ maximal N so that Paul Wins

Theorem:

$$B_K(Q) \sim \frac{2^Q}{\binom{Q}{K}}$$

Carole Strategy

Notation

$$\binom{Q}{\leq K} = \sum_{I=0}^{K} \binom{Q}{I}$$

Theorem: $N\binom{Q}{\leq K} > 2^Q \Rightarrow$ Carole Wins

Proof 1: Preserve Ministrategies

Proof 2: Random Play

Proof 1 ⇒ Proof 2: Derandomization

Paul Strategy

$(K = 1$, General Case similar)

Weight = Number Viable Ministrategies

Initial Weight $W_Q > (1+\epsilon)2^Q$

Paul splits ministrategies as evenly as possible

$W_{i-1} \leq \frac{1}{2}(W_i + i + 1)$

(worst case: $(2L+1, 0)$)

Errors don't accumulate!

When reach $(1, S)$, Endgame

Halflie: No False Negatives

N Possibilities

Q Queries

K Halflies

$A_K(Q) =$ maximal N, Paul Wins

Theorem (Cicalese/Mundici): $A_1(Q) \sim 2^{Q+1}/Q$

Dumitriu/JS:

$$A_K(Q) \sim 2^K B_K(Q) \sim 2^K \frac{2^Q}{\binom{Q}{K}}$$

Position $\vec{x} = (x, y)$ $((x_0, \ldots, x_K))$

Paul Query: (a, b) $((a_0, \ldots, a_K))$

Yes $(a, b + x - a)$; No $(x - a, y - b)$

Perfect Split $(\frac{x}{2}, \frac{y}{2} - \frac{x}{4})$

Yes/No $L\vec{x} := (\frac{x}{2}, \frac{y}{2} + \frac{x}{4})$

Problems: Integrality, Nonnegativity

Weight $W_Q(\vec{x}) = L^Q(\vec{x}) \cdot \vec{1}$

$W_Q(x, y) = 2^{-Q}(x(1 + \frac{Q}{2}) + y)$

$2^{-Q}(x_0 p_K(Q) + \ldots + x_{K-1}(1 + \frac{Q}{2}) + x_K)$

Paul Strategy

Start $(N, 0)$, $N < (1 - \epsilon)2^{Q+1}/Q$

- Give Ground to (N, N)

$T := \lfloor \lg N \rfloor$

- Roundoff so $2^T | N$
- T perfect splits to $L^T(N\vec{1})$
- Endgame: Win in R from

$(0, 2^R)$; $(1, 2^R - 1)$; $(2, 2^R - 3)$; $(3, 2^R - 5)$

A Combinatorial Approach

1-Set: Subset of $\{Y,N\}^Q$ with

stem	$YNNYNY$
child	$Y\underline{Y}YNNY$
child	$YN\underline{Y}YYN$
child	$YNNY\underline{Y}N$

0-Set: Any Singleton

K-Set: Depth K tree with marked "lies."

parent	$Y\underline{Y}YNNYN$
child	$Y\underline{Y}YN\underline{Y}NN$
grandchild	$Y\underline{Y}YN\underline{YY}Y$

Theorem: Paul Wins from $(x_0,\ldots,x_K)$ in Q

$\Leftrightarrow$ Can Pack x_i $K-i$-Sets in $\{Y,N\}^Q$

Bound Packing of K-Sets

- When all have $\geq L$ N, Size $> \binom{L}{\leq K}$

$L \sim \frac{Q}{2}$ Volume Bound $2^Q / \binom{Q/2}{K}$

$o(2^Q Q^{-K})$ have any $L < (1 - o(1))\frac{Q}{2}$

$A_K(Q) < (1 + o(1)) 2^Q / \binom{Q/2}{K}$

Careful Cutoff

Set $L = \frac{Q}{2} + c\sqrt{Q}\sqrt{\ln Q}\ Y$

$$A_K(Q) \leq \frac{2^Q}{\binom{Q/2}{K}} (1 + cQ^{-1/2}\sqrt{\ln Q})$$

Yan/JS: Remove $\sqrt{\ln Q}$

Two Batch Strategy

Deppe/Ahlswede/Cicalese/Mundici/Dumitriu/JS

$\{Y,N\}^{r*}$: Number Y within $r^{0.6}$ of $\frac{r}{2}$

$|\{Y,N\}^{r*}| \sim 2^r$

"Assume" $N = |\{Y,N\}^{r*}| \sim 2^Q/(2Q)$

Associate $\sigma \in \{Y,N\}^{r*}$ with possibility

Batch 1: $1 \le i \le r$: Is $\sigma_i = Y$?

Carole *must* say No about half the time!

Endgame from $(1, \sim \frac{r}{2})$ in One Batch

Arbitrary Channel

T-ary queries

E lie patterns

Example with $T = 3$, $E = 4$

Ternary Answers A/B/C.

Carole may lie B for A, A for B, A or B for C.

Theorem (Dumitriu, JS):

$$A_K^*(Q) \sim \frac{T^K}{E^K} \frac{T^Q}{\binom{Q}{K}}$$

Open Question

What is the maximum number $G(R)$ of disjoint 1-Shadows in $\{Y, N\}^R$?

$$\frac{2^R}{R+1} \leq G(R)$$

$$G(R) \leq 2\frac{2^R}{R}(1 + o(1))$$

Asymptotic Factor of Two Gap.

Jim Propp's
Random Walk
Simulator

The P-Machine on Z

Initially: Arbitrary chips on even positions

Every position x has “arrow” $\epsilon_x = \pm 1$.

Initially: All ϵ_x arbitrary

Each round every chip moves one position

If $2a$ at x then a to $x \pm 1$.

If $2a+1$ at x then a to $x \pm 1$.

Then “odd” chip to $x + \epsilon_x$ and,

critically, reset $\epsilon_x \leftarrow -\epsilon_x$.

The L-machine

Chips infinitely divisible

If b at x then $\frac{1}{2}b$ to $x \pm 1$.

Chips at (x,t) is Expected Number if every chip takes Random Walk

Fix initial start, arrows

$P(x,t)$ chips at (x,t) in P-machine

$L(x,t)$ chips at (x,t) in L-machine

Theorem:

$$|P(x,t) - L(x,t)| \leq 3$$

Generalizations

Any bipartite graph G with Finite Degrees

Initially: All chips on “even” positions

Each x has arrow toward neighbor

For each x ordering of neighbors.

Put “extra” i chips to next i neighbors

And, critically, readjust arrow

Sometimes

$$|P(x,t) - L(x,t)| \le K_G$$

JS/Cooper: Yes for Z^d

Outline of Argument (for Z)

Time backwards from T to zero

$X_T, \ldots, X_0$. $X_T = P(x, 0)$, $X_0 = L(x, 0)$.

X_t: Do P until t then L until zero

$$F(d, t) := \Pr[S_{t-1} = d - 1] - \Pr[S_t = d]$$

$$-F(d, t) = \Pr[S_{t-1} = d + 1] - \Pr[S_t = d]$$

$$X_{t-1} - X_t = \textstyle\sum_d A(x - d, t)$$

$$A(x-d,t) \begin{cases} = 0 & \text{if even chips at } w \\ = F(d,t) & \text{if odd chips, } \rightarrow \\ = -F(d,t) & \text{if odd chips, } \leftarrow \end{cases}$$

$X_0 - X_T = \sum_t \sum_w A(x-d,t) = \sum_w \sum_t A(x-d,t)$

Fixing d

$F(d,t)$ *unimodal*, same sign

$\Rightarrow |\sum_t A(x-d,t)| \leq \max_t |F(d,t)| = O(d^{-2})$

$X_0 - X_T \leq \sum_d O(d^{-2}) = O(1)$