Courant, March 2005

Counting and Generating

CONNECTED

Graphs

using
ERDŐS MAGIC

Joel Spencer

Joint with

Remco van der Hofstad (EURRANDOM)

Nitin Arora (Courant)

How many (asymptotically) labelled connected graphs are there with n vertices and $2n-1$ edges?

How can we generate one such graph uniformly?

We can in expected linear time!

Look at *random graph* $G(n,p)$

Apply BFS to check connectivity, then check unexposed pairs.

If connected with $2n-1$ edges take it

Problem: Chance of success exponentially small

Two Speedups

• Exponential Speedup: So each vertex has chance of joining BFS tree.

• Polynomial Speedup: So good chance of getting precisely n additional edges. Needs particular $p = \frac{c}{n}$. (Surprise: $c \neq 2$)

BFS on $G(n,p)$

Root 0; Nonroots $1, \ldots, n-1$

`Quantum Principle:` Adjacency not determined until examined.

`Key Idea:` Nonroots try to get *in* to BFS Tree

Nonroot j flips coin until head at T_j^*.
Interpretation: Enters BFS Tree at stage T_j^*

`Fictional Continuation:` Defined even if BFS has aborted

`Exponential Speedup:` Condition on $T_j^* \leq n$.

Otherwise would not be connected.

Replace T_j^* by *truncated geometric* T_j

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

`Implementation:` $O(1)$ time to find each T_j

Tilted Balls in Bins

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

Truncated Geometric

Ball j in Bin T_j

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

Z_i balls in bin i

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

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

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

Breadth First Search

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

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

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

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 $34, 25$

Pr[TREE] with $p = \frac{c}{n}$

Significant Drift

Left Bins $Po(\lambda^L)$ balls

Right Bins $Po(\lambda^R)$ balls

$$\lambda^L = \frac{c}{1-e^{-c}} > 1 > \frac{ce^{-c}}{1-e^{-c}} = \lambda^R$$

Left Side has Positive Drift, OK $\Omega(1)$

Right Side has Negative Drift, OK $\Omega(1)$

Middle OK (Chernoff . . .)

Pr[TREE] $= \Omega(1)$ $(= 1 - (c+1)e^{-c})$

`Implementation:` Check TREE in $O(n)$ time.

If TREE fails, `ABORT`.

Exposing the Unexposed Edges

$M = \binom{n}{2} - \sum_{j=1}^{n} T_j$ unexposed

$E[M] = \Theta(n^2)$, $Var[M] = \Theta(n^3)$

Conditioning on TREE negligible effect on M

M Gaussian

Choose $p = \frac{c}{n}$ **so that** $pE[M] = n$

Second Speedup

Problem: $\Pr[BIN[M,p]=n]=o(1)$

Set $f(m) := \Pr[BIN[m,p]=n]$.

Probability $f(M)/\max_m f(m)$: `SUCCESS`

Otherwise: `ABORT`.

$M = E[M] + \Theta(n^{3/2})$

$f(M) = \Omega(\max_m f(m))$

$Pr[\texttt{SUCCESS}] = \Omega(1)$

`Implementation:` Select precisely n unexposed.

Clever data structure: $\Theta(n)$ time.

n vertices $n-1+l$ edges

$l = \Theta(n)$: As described

$l > cn \ln n$. Most G connected.

Random G and verify.

$n \ll l = O(n \ln n)$: As described. $\Pr[\texttt{ABORT}] \to 0$

$l \ll \ln n$???

$\ln n < l \ll n$ `EARLY ABORT`

Check balls in bins going in from edges

Counting Connected Graphs

Event A: $G(n,p)$ connected with $2n-1$ edges

$C(n,n) :=$ number of such graphs

$\Pr[A] = C(n,n)p^{2n-1}(1-p)^{\binom{n}{2}-(2n-1)} = A_1 A_2 A_3$

$A_1 = (1-(1-p)^n)^{n-1}$ nonroots throw heads

$A_2 = \Pr[\text{TREE}] \sim 1-(c+1)e^{-c}$ Left and Right walk analyses.

$A_3 = \Pr[BIN[M,p] = n]$

M Gaussian, mean np^{-1}, variance $c'n^3$

A_3 computable