Experimental Asymptotics: How Much Experimentation is Enough?

M. Goldberg, D. Hollinger, M. Magdon-Ismail

Computer Science Department
Rensselaer Polytechnic Institute
Maximum Independent Set

- Random Graphs $G(n, p)$

$$2\log\frac{1}{1-p} n \quad \text{[Bollobás and Erdös]}$$

- Trivial sequential algorithm $Random(G)$:

  $$I = \emptyset \text{ (the independent set).}$$
  
  while ($G \neq \emptyset$) do
  
  select a vertex $v$ at random.

  add $v$ to $I$ and remove from $G$.

  remove all neighbors of $v$ from $G$.

  end while

  It is easy to show that $Random$ finds sets of size $\log\frac{1}{1-p} n$.
Question: Is there any polynomial algorithm better than $Random$? 

$$(1 + \epsilon) \log \frac{1}{1-p} n$$ for some $\epsilon > 0$

Experimental evaluation of simple algorithms should be easy?.

We start with the simplest modification to $Random$ we can think of.
Sequential Greedy algorithm \textit{Greedy}

At each step, select a vertex with minimum degree:

\[ I = \emptyset \text{ (the independent set).} \]

while \( G \neq \emptyset \) do

\quad select a vertex \( v \) from among those in \( G \) with minimum degree.
\quad add \( v \) to \( I \) and remove from \( G \).
\quad remove all neighbors of \( v \) from \( G \).

end while
Experiments with fixed edge probability $p$

- For fixed edge probability $p$ (we used $p = 0.1$) and various graph sizes $n$:
  - run both $Random$ and $Greedy$ on a number of randomly generated graphs.
  - compute the average ratio: $\frac{I_{Greedy}}{I_{Rand}}$

- Plot ratio vs. $n$
Ratio of Greedy to Random
Edge Probability 0.1
Ratio of Greedy to Random
Edge Probability 0.1

ratio

number of vertices

0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000

0 1.23 1.24 1.25 1.26 1.27 1.28 1.29 1.3 1.31

data
Ratio of Greedy to Random
Edge Probability 0.1

Ratio of Greedy to Random
Edge Probability 0.1

1/log
1/sqrt
data

ratio

number of vertices

0  5000  10000  15000  20000  25000  30000  35000  40000  45000  50000
Conjecture

Experiments are not conclusive, but it appears that:

\[ \text{Greedy} = \text{Random} \]

for random graphs with fixed edge probability \( p \).

We tried a number of different edge probabilities \( p \) - all had similar results.
Other Classes of Graphs

- For 3-regular graphs, *Greedy* is better than *Random* [Frieze and Suen].
- For d-regular graphs $d > 3$ the answer is unknown.
- For *average* degree graphs the answer is unknown.
- We worked on average degree $d = 3$. 
Average Degree $d = 3$

*Random* finds independent sets of size:

$$1 + \log_{\frac{1}{1-p}} (1 + (n - 1)p) \approx .46n$$

$$p = \frac{3}{n - 1}$$
Fixed Average Degree Experiments

• For fixed average degree $d = 3$ and various graph sizes $n$:
  – run both $Random$ and $Greedy$ on a number of randomly generated graphs.
  – compute the average ratio: $\frac{I_{Greedy}}{I_{Rand}}$

• Plot ratio vs. $n$
Ratio of Greedy to Random
Average Degree 3

Measured

1/log

ratio

number of vertices

1e+06
Average degree $d = 3$ Combinatorics

\[ I(n) = I_0 + I_1 + I_2 + I_3 + I_4 + \ldots \]

\[ n = I_0 + 2I_1 + 3I_2 + 4I_3 + 5I_4 + \ldots \]

- $I(n)$ is the size of independent set found by *Greedy* on a graph with $n$ vertices and average degree 3.

- $I_k$ is the number of vertices with degree $k$: *when selected for inclusion in the independent set.*
$I_k$ are Dynamic Degrees
Average degree $d = 3$ discovery

Discovery: *Greedy* never selects a vertex of degree 3 (or more).

\[
I(n) = I_0 + I_1 + I_2
\]

\[
n = I_0 + 2I_1 + 3I_2
\]
Degree 0 vertices

\[ I_0 = I_0^{\text{initial}} + I_0^{\text{dynamic}} \]

- \( I_0^{\text{initial}} \) is number of vertices in the initial graph with degree 0.
  \[ I_0^{\text{initial}} = n(1 - p)^{n-1} = n\left(1 - \frac{3}{n - 1}\right)^{n-1} \rightarrow \frac{n}{e^3} \]

- \( I_0^{\text{dynamic}} \) is number of additional vertices removed when their degree is 0
Distribution of Degrees (dynamic degrees)

average degree \( d = 3 \)
Conjecture for \textit{Greedy} on average degree random graphs $d = 3$

Independent set size found by \textit{Greedy} on random, average degree-3 graphs with $n$ vertices:

$$I(n) = I_{0}^{initial} + I_{0}^{dynamic} + I_{1} + I_{2}$$

$$I(n) \rightarrow 0.53n$$
Issues

- Experimental analysis of the asymptotic behavior of algorithms can lead to the “always need to try larger inputs” syndrome.

- Experiments can sometimes be designed to illuminate asymptotic behavior indirectly. In our case, theoretical methods and results for 3-regular graphs provided necessary insights.
Degree 0 +

$I_0^{\text{dynamic}}$ vertices selected
Average Degree 3 Graphs
Degree 1 vertices selected
Average Degree 3 Graphs

fraction of n

n

0 10000 20000 30000 40000 50000 60000 70000 80000 90000 100000
Degree 2 vertices selected
Average Degree 3 Graphs