Bounding Ramsey Numbers by Search & Optimization

Dylan Hutchison
Margaret Vulih

10 December 2013

Background

- Named after Frank P. Ramsey. Inspiration/Bio: A British mathematician who published the paper *On a problem of Formal Logic* in 1928 that became the basis of Ramsey Theory.
- Popularized by Erdos and Szekeres in the paper *Happy Ending Problem* [1935]

Frank P. Ramsey

http://en.wikipedia.org/wiki/File:Frank_Plumpton_Ramsey.JPG

Applications

- Graphs and Social Networks
  - In a group of n people, if n >= R(p,q), then there is either
  - a subcollection of p people who are pairwise all friends OR
  - a subcollection of q people who are pairwise all not friends.
- Convex Geometry
  - With n points in general position (no 3 points collinear), if n >= R(5,m; r=4),
  - there MUST BE a subcollection of m points that compose a convex m-gon.
- Information Theory with Dual Cores
  - Goal: Find the maximum number of codes that can be transmitted down a channel *without error* → when 2 codes can be confused for one another.
  - Nodes = codes, edges = pairs of codes confused when transmitted
  - Given n codes, find $\max_{m \geq 1} m \text{ such that } n >= R(2, m)$

Ramsey Number: Graph Interpretation

$$R(p,q)$$

(function $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$)

is the smallest integer such that:

Every graph G with $|G| > R(p,q)$ nodes has, no matter how you arrange the edges,

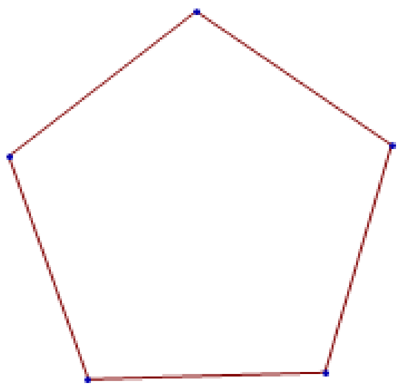
- EITHER a complete (*clique*) subgraph $H \subseteq G$ of size $|H|=p$
- OR an *independent set* subgraph $H \subseteq G$ of size $|H|=q$

“Guaranteed order amidst large enough chaos”

Example: R(3,3)=6

```
GraphPlot[{1 → 2, 2 → 3, 3 → 4, 4 → 5, 5 → 1}]
```

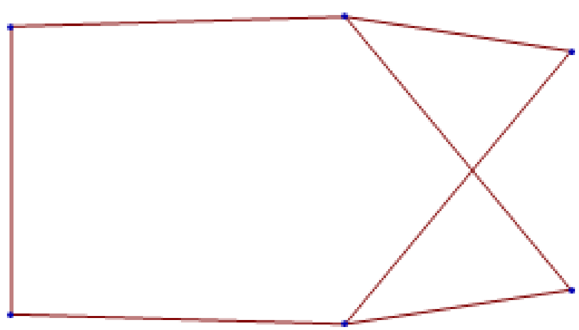
Out[2]=



5 nodes:
May have no 3-clique
or 3-indep subgraph

```
In[4]:= GraphPlot[{1 → 2, 2 → 3, 3 → 4, 4 → 5, 5 → 1, 1 → 6, 6 → 3}]
```

Out[4]=



6 nodes:
We are guaranteed
Either 3-clique
or 3-indep subgraph

Known Ramsey #'s and Bounds

p, q	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7	8	9	10
3	1	3	6	9	14	18	23	28	36	40-43
4	1	4	9	18	25	35-41	49-61	56-84	73-115	92-149
5	1	5	14	25	43-49	58-87	80-143	101-216	125-316	143-442
6	1	6	18	35-41	58-87	102-165	113-298	127-495	169-780	179-1171
7	1	7	23	49-61	80-143	113-298	205-540	216-1031	233-1713	298-2826
8	1	8	28	56-84	101-216	127-495	216-1031	282-1870	317-3583	317-6090
9	1	9	36	73-115	125-316	169-780	233-1713	317-3583	565-6588	580-12677
10	1	10	40-43	92-149	143-442	179-1171	289-2826	317-6090	580-12677	798-23556

Source: http://documents.kenyon.edu/math/BuschurKSenEx2011.pdf

Lower Bounds

- Parity Argument: $\forall m,\ R(m,3) > 2(m-1)$
- Probability Argument: if $\exists\, t \in [0,1]$ such that $\binom{n}{p} t^{\binom{p}{2}} + \binom{n}{q}(1-t)^{\binom{q}{2}} < 1$, then $R(p,q) > n$.
- Constructive Argument: Find a *bad* graph size n such that it has no p-clique or q-indep. subgraph. Then $R(p,q) > n$.

Upper Bounds

- Graph Argument: $\forall p,q \geq 2,\ R(p,q) \leq R(p,q-1) + R(p-1,q)$
- Induction/Generating Functions: $\forall p,q \geq 2, \quad R(p,q) \leq \binom{p+q-2}{p-1}$
- Brute Force: Show that every graph of size n has a p-clique or q-indep. subgraph. Then $R(p,q) \leq n$.

Asymptotics

- Noticed $R(p,q) = O\left(\dfrac{R(p,q-1)^2}{R(p,q-2)}\right)$
- $R(3,t) = \Theta\left(\dfrac{t^2}{logt}\right)$ in literature [Kim 1995]

R(q,q) Search using C++

```
// set n := known number below ub
while (!found_ub) {
    gr_init(&gr, n);
    do {
        gr_search(&gr, qboth);
        gr_incr(&gr);
    } while (gr.foundki != NONE && !gr_allzero(&gr));

    if (gr.foundki != NONE) {
        found_ub = true;
    } else {
        // n is a lower bound
    }
    gr_free(&gr);

    n++;
}
```

R(4,4) Search using Matlab

- While <=10 bad graphs
  - Generate a subgraph
  - If it is not duplicate subgraph
    - Check if it is a bad graph
    - If there are more bad graphs that iterations of the while loop than iterate r.
- End
- Print out final r

Optimization Form

$$\min_{G} C(G) + I(G)$$

- Minimize the totals number of nodes n that guarantees a complete subgraph or an independent subgraph.

Results

- Using C++
- We found R(3,3) = 6
- True number is 6

- Using Matlab
- We found R(4,4) > 16~17 (randomness)
- True number is 18

Summary

- No analytical solutions for *Large* Ramsey Numbers
  - Use brute force numerical techniques.
- We systematically searched through all graphs of size n to bound the Ramsey number R(q,q)
- Partial results on random search through graphs size n
  - If n < R(q,q), random search can find a new lower bound
  - If n >= R(q,q), random search fails – still need to enumerate all graphs for upper bound
- Active research area – new Ramsey # = new paper

Set Interpretation

$$\forall p, q \; \exists R(p,q), \; \forall n \geq R,$$
$$\forall S, \; |S| = n, \; \forall P_2(S) = \mathcal{A}_1 \uplus \mathcal{A}_2$$
$$(\exists U \subseteq S, \; |U| = p, \; P_2(U) \subseteq \mathcal{A}_1) \vee (\exists U \subseteq S, \; |U| = q, \; P_2(U) \subseteq \mathcal{A}_2)$$

Given subset numbers p, q,
there is a minimum Ramsey number R that, for any set S of size $n \geq R$
and for any partitioning of all the $\binom{n}{2}$ pair subsets of S into the sets of pairs
$\mathcal{A}_1, \mathcal{A}_2$ $(\mathcal{A}_1 \cap \mathcal{A}_2 = \varnothing$, but we allow $\mathcal{A}_1 = \varnothing$ or $\mathcal{A}_2 = \varnothing)$,
there exists either a subset $U \subseteq S$ of length p such that
all of the $\binom{p}{2}$ pair subsets of U are contained in $\mathcal{A}_1$,
or a subset $U \subseteq S$ of length q such that
all of the $\binom{q}{2}$ pair subsets of U are contained in $\mathcal{A}_2$.