

Ramsey Theory in Theoretical Computer Science

Yitong Yin

November 12, 2004

Ramsey Theory

- “Of three ordinary people, two must have the same sex.” — *D. J. Kleitman*
- In any collection of six people either three of them mutually know each other or three of them mutually do not know each other.
- Every irregular structure, if it is large enough, contains a regular substructure of some given size.

Ramsey's Theorem — Abridged

Definition 1 *We write*

$$n \rightarrow (l_1, \dots, l_r)$$

if, for every r -coloring of $[n]^2$, there exists i , $1 \leq i \leq r$, and a set $T \subseteq [n]$, $|T| = l_i$ so that $[T]^2$ is colored i .

Thus the above example can be written as:

$$6 \rightarrow (3, 3)$$

Definition 2 *The Ramsey function $R(l_1, \dots, l_r)$ denotes the minimal n such that $n \rightarrow (l_1, \dots, l_r)$.*

Theorem 1 *The function R is well defined. That is, for all l_1, \dots, l_r there exists n so that $n \rightarrow (l_1, \dots, l_r)$.*

Ramsey's Theorem — Unabridged

Similarly, define

$$n \rightarrow (l_1, \dots, l_r)^k$$

on hypergraph $[n]^k$

Write $R_k(l_1, \dots, l_r)$ as the Ramsey Function on k -sets.

Theorem 2 (Ramsey's theorem)

The function R_k is well defined. That is, for all l_1, \dots, l_r there exists n_0 so that, for $n \geq n_0$,

$$n \rightarrow (l_1, \dots, l_r)^k.$$

Ramsey-type Theorem

Except for coloring large complete hypergraphs, one can color other objects (points in spaces, numbers, subsets, etc.) and also, one can look for other “regular” substructures (monochromatic subspaces, arithmetic progressions, configurations, etc.)

1. Van der Waerden's theorem.
2. Schur's theorem.
3. Rado's theorem.
4. Hales–Jewett theorem.
5. Graham–Leeb–Rothschild theorem.

Ramsey Theory Applications

- in Number Theory, Geometry, Topology, Set Theory, Logic, Ergodic Theory, Information Theory
- and **Theoretical Computer Science**

Yao's celebrated paper

Should Tables Be Sorted? [JACM 1981]

The basic information retrieval problem “Is $x \in S$ ” is considered.

Model: S is a set of n keys from a large domain $M = \{1, \dots, m\}$. A table structure \mathcal{T} specifies how any particular set of n keys are to be placed in the table. An algorithm is a sequence of probes.

S : edge; M : vertex set; \mathcal{T} : coloring.

A lower bound: For sufficiently large M , $\lceil \log(n+1) \rceil$ probes are needed.

This is done by:

1. For $m \geq 2n - 1$, if the table is sorted (or in some fixed permutation), then $\lceil \log(n+1) \rceil$ probes are needed. (This is by adversary argument.)
2. According to Ramsey's theorem: for sufficiently large m ,

$$m \rightarrow \underbrace{(2n - 1, \dots, 2n - 1)}_{n!}^n$$

Similar result for Decision Trees

“Applications of Ramsey’s Theorem to Decision Tree Complexity” by Moran, Snir and Manber [JACM 1985]

Model: A decision tree T is a labeled binary tree. Each internal node is associated with a predicate P on S^n , where S is the domain of the decision problem.

Theorem: For sufficiently large S , there exists a set $C \subset S$ such that the decision tree T is order invariant on C^n .

WLOG assume that the arity of P is 2: there are 4 order invariant equivalence classes, and

$$|S| \rightarrow (|C|, |C|, |C|, |C|)^2$$

generally the arity is k and there are multiple P .

Along this line

For the sufficiently large domain, there is some restricted domain, where on this domain the original problem is transformed into a simpler class (e.g. ordered). And some desired property (e.g. lower bound) is proved for this simpler class and then generalized to the original problem.

“Electing a Leader in a Synchronous Ring” by Frederickson and Lynch [JACM 1987]

“One, Two, Three ... Infinity: Lower Bounds for Parallel Computation” by Fich, au der Heid and Widgerson [STOC 1985]

“What Can Be Computed Locally?” by Naor and Stockmeyer [STOC 1993]

Two drawbacks

1. The model is restricted.

This seems to be inevitable.

2. The domain must be extremely large, which makes the results impractical.

De-ramsey: “Removing Ramsey Theory: lower bounds with smaller domain size” by Edmonds [TCS 1997]

Applications with extremal flavor

Ramsey vs. Extremal

Extremal: if a collection of finite objects satisfies certain restrictions, how large or how small can it be?

So this time we can directly get the bound rather than transform the problem into simple type.

An alternative proof to Linial's $\Omega(\log^* n)$ lower bound on the distributed algorithms for MIS of n -cycles. "Locality in Distributed Graph Algorithms" by Linial [SICOMP 1992].

MIS in an n -cycle

In an n -cycle with ID space $[n]$, a t -time bounded local algorithm is nothing but a mapping

$$f : \mathbf{P}_{[n]}^{2t+1} \rightarrow \{0, 1\}$$

where $\mathbf{P}_{[n]}^{2t+1}$ denote the set of all permutations of size $2t + 1$.

From f , we can define another mapping $h : [n]^{2t+1} \rightarrow \{0, 1\}$, in the natural way.

For $n \geq R_{2t+1}(2t + 3, 2t + 3)$, there exists monochromatic clique of $(2t + 3)$ vertices. And thus there are three consecutive nodes in the n -cycle with the same value. The result is not MIS.

Therefore the lower bound is

$$t \geq \Omega(R^{-1}(n)) \geq \Omega(\log^* n)$$

An algorithm named Ramsey

Recall that in the six people example, the Ramsey theory is just a tradeoff between Independent Set and Clique.

Ramsey is an algorithm that enjoys this tradeoff. “Approximating Maximum Independent Sets by Excluding Subgraphs” by Boppana and Halldórsson.

For the approximation of Independent Set and Clique number, a natural heuristic is:

Pick v

$$I(G) \leftarrow \{v\} \cup I(\overline{N}(v))$$

$$C(G) \leftarrow \{v\} \cup C(N(v))$$

It suffers from the bad pivot nodes.

Ramsey algorithm is defined as:

Pick v

$$I(G) \leftarrow \max(\{v\} \cup I(\overline{N}(v)), I(N(v)))$$

$$C(G) \leftarrow \max(\{v\} \cup C(N(v)), C(\overline{N}(v)))$$

Define $r(s, t)$ as the smallest n such that the algorithm produces either a clique of size s or an independent set of size t .

$$R(s, t) \leq r(s, t) \leq \binom{s+t-2}{s-1} \text{ and so } |I||C| \geq \frac{1}{4}(\log n)^2$$