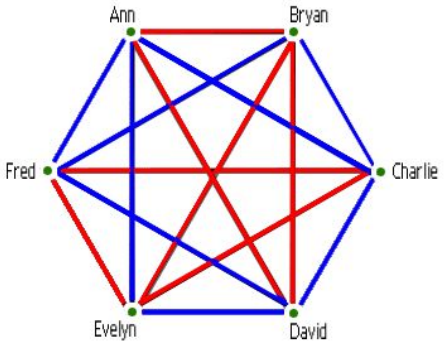
Verified Certificates via SAT & Computer Algebra Systems for the Ramsey R(3, 8) and R(3, 9) Problems

Zhengyu Li[1], Conor Duggan[2], Curtis Bright[3], Vijay Ganesh[1]

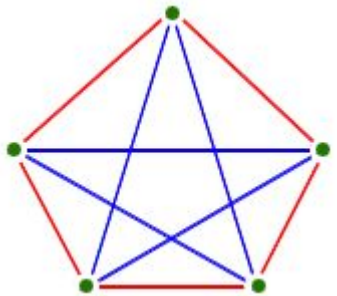
[1]Georgia Institute of Technology, USA
[2]University of Waterloo, Canada
[3]University of Windsor, Canada

Ramsey Numbers

- The Ramsey number R(p,q) solves a classic problem: What's the smallest party size needed to guarantee that either p people all know each other, or q people are all strangers to each other?
- In graph theory terms, R(p,q) is the smallest number of vertices where every 2-coloring of the edges either contains p vertices all connected by the first color or q vertices all connected by the second color.



R(3,3) <= 6



R(3,3) > 5

Visualization of Ramsey Theory concepts. Adapted from "Friends and Strangers" by I. Leader, 2000, Plus Magazine (https://plus.maths.org/content/friends-and-strangers). Copyright 2000 by University of Cambridge.

Ramsey Numbers

COMBINATORICS

Disorder Persists in Larger Graphs, New Math Proof Finds

6 |

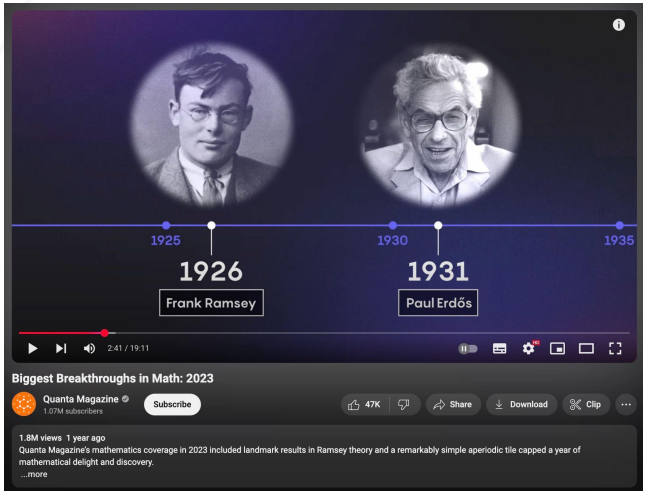
David Conlon and Asaf Ferber have raised the lower bound for multicolor "Ramsey numbers," which quantify how big graphs can get before patterns inevitably emerge.

GRAPH THEORY

A Very Big Small Leap Forward in Graph Theory

8 |

Four mathematicians have found a new upper limit to the "Ramsey number," a crucial property describing unavoidable structure in graphs.

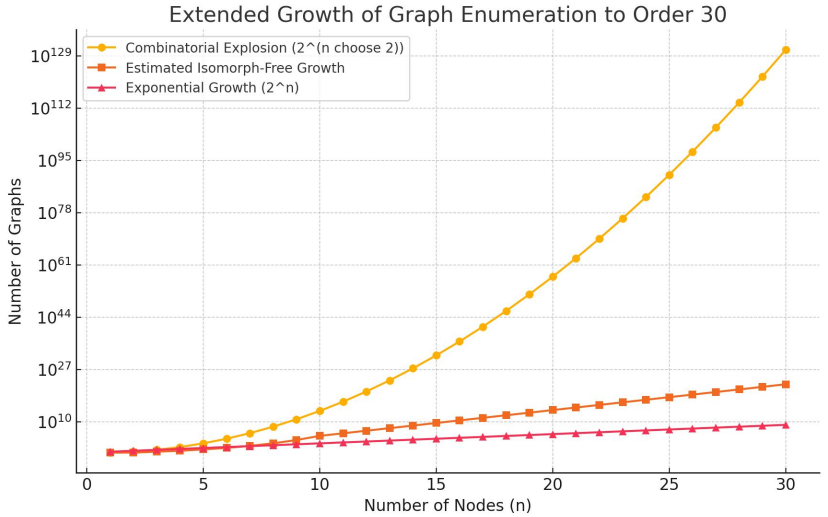


"Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find the Ramsey number for red five and blue five. We could marshal the world's best minds and fastest computers, and within a year we could probably calculate the value. If the aliens demanded the Ramsey number for red six and blue six, however, we would have no choice but to launch a preemptive attack"
– Erdős 1990

University of Windsor
UNIVERSITY OF WATERLOO
Georgia Tech

Ramsey Numbers

$R(p,q)$	$p=3$	$p=4$	$p=5$
$q=3$	6		
$q=4$	9	18	
$q=5$	14	25	43–46
$q=6$	18	36–40	59–85
$q=7$	23	49–58	80–133
$q=8$	**28**	59–79	101–193
$q=9$	**36**	73–105	133–282



- Only 9 non-trivial Ramsey numbers are known.

University of Windsor
UNIVERSITY OF WATERLOO
Georgia Tech

Current Approaches in Ramsey Papers

Graph Theory

Computation using nauty (unverified)

nauty and Traces
Brendan McKay and Adolfo Piperno
GRAPH CANONICAL LABELING AND AUTOMORPHISM GROUP COMPUTATION

University of Windsor
UNIVERSITY OF WATERLOO
Georgia Tech

Our Approaches to prove Ramsey Numbers

SAT Solvers

Computer Algebra Systems

Motivations of SAT+CAS

- SAT solvers are great at solving search problems specified by simple constraints (clauses).
- Computer algebra systems (CASs) are great at many sophisticated mathematical problems involving little search.
- Problems involving both sophisticated mathematics and search are good candidates for a SAT+CAS approach. (developed in 2015 by Zulkoski, Ganesh, and Czarnecki and independently by Erika Ábrahám)

SAT + CAS = efficient search + mathematical knowledge

Zulkoski, E., Ganesh, V., Czarnecki, K.: MathCheck: A Math Assistant via a Combination of Computer Algebra Systems and SAT Solvers. Springer, Cham (2015).
Ábrahám, E.: Building bridges between symbolic computation and satisfiability checking (2015).

An Emerging Paradigm

There has been a lot of research in recent years involving SAT and computer algebra or related methods.

A small and incomplete sample:

- Improving Lower Bound of Kochen–Specker Problem in Quantum Mechanics (Li et al. 2024)
- A SAT-based Resolution of Lam's Problem (Bright et al. 2021).
- A Hybrid SAT and Lattice Reduction Approach for Integer Factorization (Ajani, Bright 2023).
- Proving the correctness of multiplier circuits (Kaufmann, Biere 2020).
- Finding new algorithms for 3 × 3 matrix multiplication (Heule, Kauers, Seidl 2021).
- SAT modulo symmetries for generating combinatorial objects in an isomorph-free way (Kirchweger et al. 2021)
- Making progress on conjectures in geometric group theory (Savela, Oikarinen, Jarvisalo 2020).
- Computing directed Ramsey numbers (Neiman, Mackey, Heule 2020).

Boolean Encoding of the Ramsey Problem

- The encoding can be categorized into three parts:
  - Ramsey Graph Encoding:
    - encoding the formulation of the Ramsey Graph
  - Static Symmetry Breaking Encoding:
    - Break partial symmetries in the graph
  - Cardinality Encoding
    - Restrict degrees of vertices and number of edges in the graph

- We want to encode the problem in conjunctive normal form (CNF), meaning a conjunction (sequence of ANDs) consisting of one or more conjuncts, each of which is a disjunction (OR) of one or more literals.

$$(A \vee \neg C \vee \neg D) \wedge (A \vee B) \vee (\neg B \vee \neg C)$$

Boolean Encoding: Ramsey Graph Encoding

Edge Variables: Each edge e is a Boolean variable
e = True → edge colored blue
e = False → edge colored red

Result:
SAT → (p,q)-graph exists → R(p,q) > n
UNSAT → no valid coloring → R(p,q) ≤ n

$$\bigwedge_{K_p \subseteq K_n} \bigvee_{e \in K_p} \neg e \quad \text{and} \quad \bigwedge_{K_q \subseteq K_n} \bigvee_{e \in K_q} e$$

Every p-clique must have ≥1 red edge

Every q-clique must have ≥1 blue edge

Isomorph-free Orderly Generation

When generating combinatorial objects we only care about generating them up to isomorphism.

The notion of canonicity is defined so that:

- Every isomorphism class has exactly one canonical representative.
- If an adjacency matrix is canonical then its upper-left submatrix of any size is also canonical.

Developed independently by Faradžev and Read in 1978.

Canonicity Examples

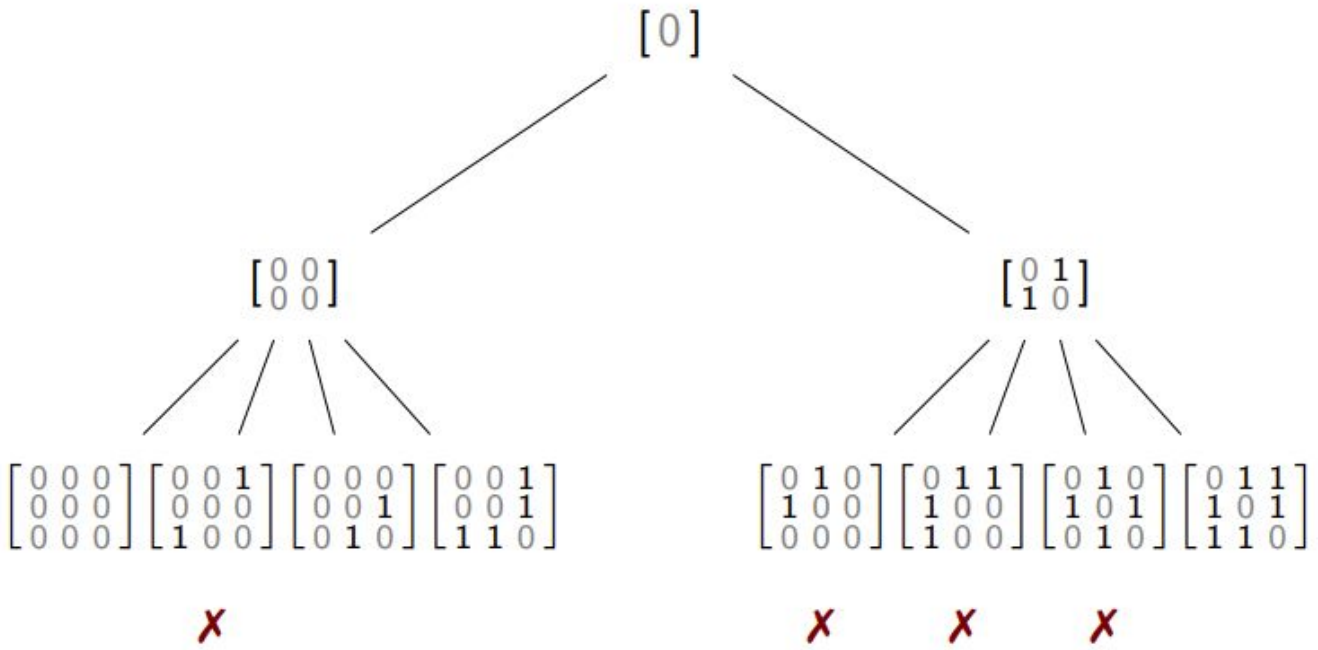
An adjacency matrix is canonical if its “vector representation” is lex-minimal among all matrices in the same isomorphism class.

For example,

Adj. matrix	$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
Vector rep.	$[1\,0\,0] >_{\text{lex}}$	$[0\,1\,0] >_{\text{lex}}$	$[0\,0\,1]$
Canonical?	✗	✗	✓

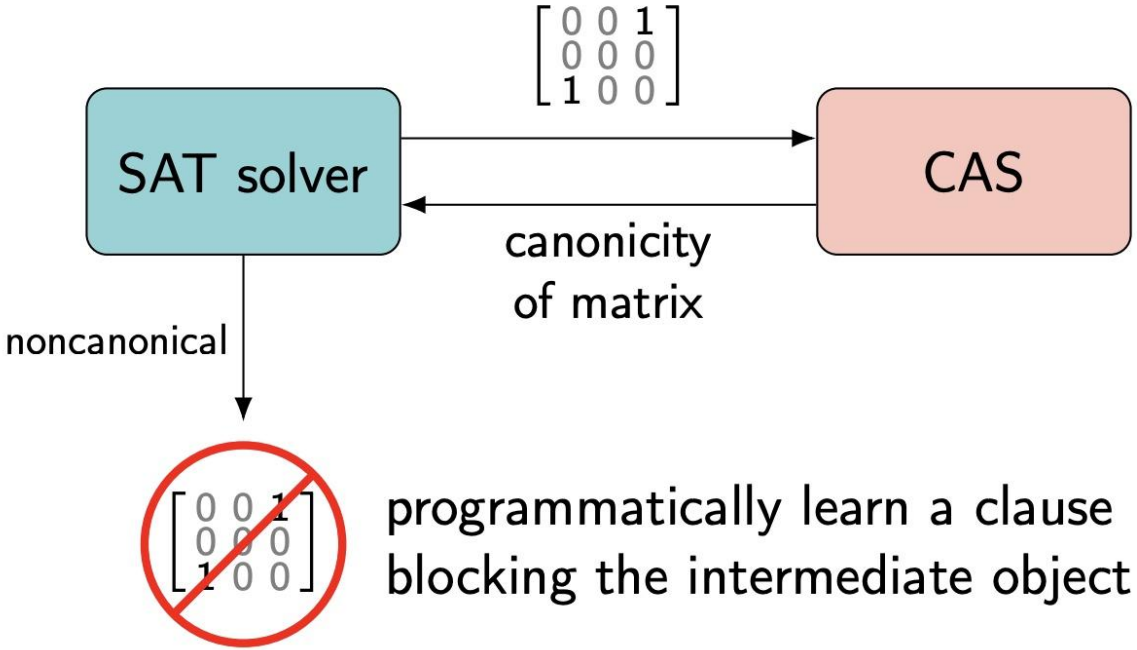
are isomorphic adjacency matrices but only the last is canonical.

Orderly Generation of Graphs

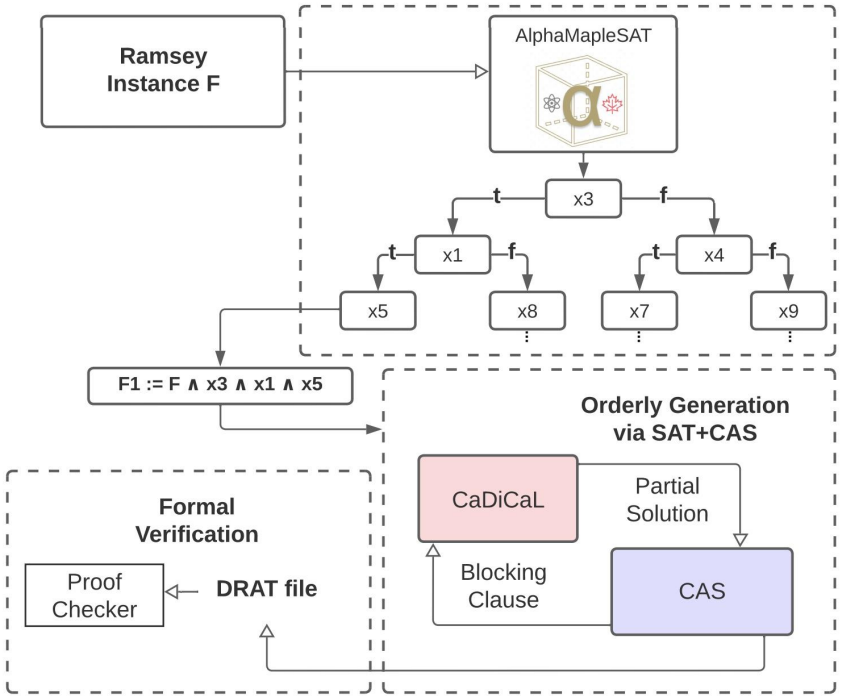


University of Windsor
UNIVERSITY OF WATERLOO
Georgia Tech

Implementing Orderly Generation



The MathCheckRamsey Pipeline



Jha, Piyush, et al. "Alphamaplesat: An MCTS-based cube-and-conquer SAT solver for hard combinatorial problems." arXiv preprint arXiv:2401.13770 (2024).

Results - R(3,8) and R(8,3)

The solving results for R(3,8) and R(8,3) confirm that R(3,8) = 28. The solution was obtained by proving that a 28-vertex R(8,3)-graph does not exist (UNSAT) and that a 27-vertex R(8,3)-graph does exist (SAT).

To efficiently solve these instances, we used cardinality constraints to restrict the degree of each vertex between 20 and 22, significantly reducing the search space.

With cube and conquer, the problem was solved in approximately 8 hours of wall clock time using parallelization.

$$n - R(p, q-1) \leq \deg_b(v) \leq R(p-1, q) - 1$$

Results - R(3,9) and R(9,3)

The solving results for R(3,9) and R(9,3) confirm that R(3,9)=36.

A key theoretical technique used was derived from **Graver and Yackel's results**, which state that any valid (3,9;36)/(9,3;36)-graph must contain a (3,8;27;80)/(8,3;27;271)-graph. Instead of solving R(3,9)/R(9,3) directly, we applied the SAT+CAS approach to an equivalent problem: proving that no (3,8;27;80)/(8,3;27;271)-graph exists. By showing this instance is UNSAT, we established that R(3,9) = (9,3) ≤ 36, and since a (3,9;35)-graph exists, this confirmed R(3,9) = 36.

We applied **cardinality constraints** to limit vertex degrees between 19 and 22 and enforce exactly 271 edges in the R(9,3)-graph. Using AlphaMapleSAT, we split the instance into 2,486 cubes, each solved and verified independently. Cubing was performed until 100 edge variables were eliminated, with further splitting applied if proof files exceeded 7 GiB.

The problem was solved in approximately **26 hours of wall clock time**.

Results and Comparisons

	CaDiCaL + CAS		CaDiCaL only	
k	$R(3,k)$	$R(k,3)$	$R(3,k)$	$R(k,3)$
7	14.3 s	8.2 s	564.3 s	220.7 s
8	112.1 h	18.5 h	>7 days	>7 days

Sequential Experiments with the same SAT instance

Instance	Cubing Time	Simplification Time	Solving Time	Verification Time	Wall Clock Time
$R(8,3)$	1,360 s	1,217 s	19,811 s	22,328 s	8 hrs
$R(9,3)$	15,530 s	42,482 s	697,575 s	473,874 s	26 hrs

Summary of experimental results for solving R(8, 3) and R(9, 3) in parallel.

Verification of Results

SAT Verification: DRAT proof logging enabled to generate certificates for all SAT solver decisions. DRAT-trim checker (modified) verifies proof correctness - only need to trust this simple verifier, not the complex SAT solver. CAS-derived clauses marked with 't' prefix as "trusted" clauses in DRAT proof.

CAS Verification: Witness-based approach where a CAS provides permutation witnesses that prove "non-canonical" partial solutions can be discarded without loss of generality. An independent Python script applies witness permutations to verify each blocked matrix produces lexicographically smaller result. No trust in CAS required - only verify the witness works, not the CAS computation itself.

Cube-and-Conquer: Recursive verification ensures generated cubes collectively partition the entire search space. Exhaustive coverage check verifies that for any cube $\varphi \wedge x$, all extensions of $\varphi \wedge \neg x$ are covered by other cubes.

Trust Assumptions: DRAT-trim verifier correctness (simple, well-established tool), Python verification script correctness (straightforward permutation application), and SAT encoding correctness (not formally verified - limitation).

Conclusion

We provide the first independently verifiable proofs for R(3,8)=28 and R(3,9)=36 using SAT+CAS. When combined with previously known domain knowledge about Ramsey problems, the search space can be reduced and thus the effectiveness of the SAT+CAS method is improved.

We provide a scalable MathCheck pipeline that enables parallelization and can be applied to more combinatorial problems that typically takes years to solve.

We hope this paradigm can open new possibilities for verifying and solving open Ramsey problems such as R(5,5) and R(3,10).

arXiv

https://uwaterloo.ca/mathcheck/

University of Windsor

UNIVERSITY OF WATERLOO

Georgia Tech