Computers Meet Combinatorics: SAT Solvers and the Search for Hales–Jewett Bounds

Nayda Farnsworth

Colgate University
Mathematics Department Seminar
April 21, 2026

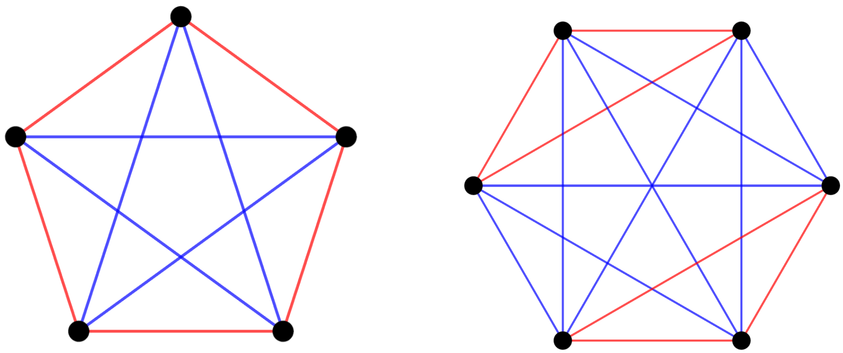
Outline

Ramsey Theory, Van der Waerden, and Hales–Jewett

What is Ramsey Theory?

Ramsey theory studies when structure becomes unavoidable in sufficiently large systems.

Example: Suppose we color each edge of K_n red or blue. How large must n be to force a monochromatic triangle in every coloring of K_n?



Van der Waerden Theorem and Numbers

We now consider **Ramsey Theory in the integers**.

Theorem (Van der Waerden Theorem [2])

For integers $k, r \in \mathbb{Z}^+$, there exists an integer $W(k; r)$ such that for all $N \geq W(k; r)$, every r-coloring of $\{1, \ldots, N\}$ contains a monochromatic k-term arithmetic progression (i.e. a sequence $\{a, a + d, \ldots, a + (k-1)d\}$).

The **Van der Waerden Theorem** guarantees such an N exists, and the **Van der Waerden number $W(k; r)$** is the smallest such N for given k and r.

EXAMPLE: $W(3,2) = 9$

Van der Waerden number: $W(3,2) = 9$

We can 2-color $\{1, \ldots, 8\}$ to avoid any monochromatic 3-term arithmetic progression:

$$\{1, 2, 3, 4, 5, 6, 7, 8\}$$

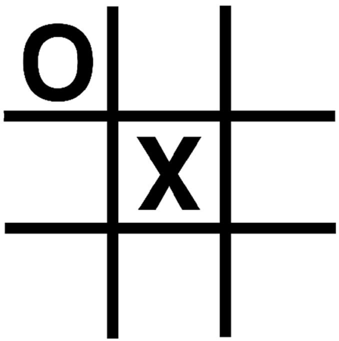
This coloring avoids all monochromatic progressions of the form:

$$\{a, a+d, a+2d\}$$

However, no 2-coloring of $\{1, \ldots, 9\}$ can avoid a monochromatic 3-term arithmetic progression.

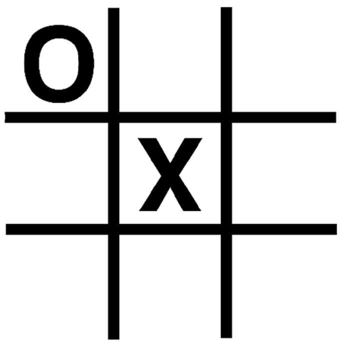
Hales–Jewett in Two Dimensions

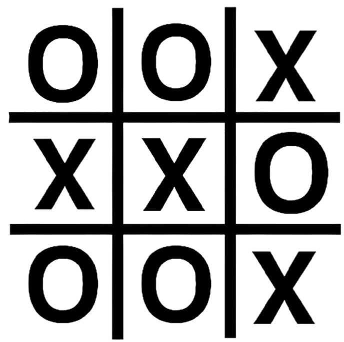
In two dimensions, the Hales–Jewett problem looks like a game of tic-tac-toe. Can we color the board in a way that avoids a monochromatic winning line?



Hales–Jewett in Two Dimensions

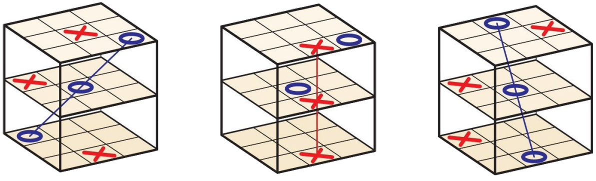
In two dimensions, the Hales–Jewett problem looks like a game of tic-tac-toe. Can we color the board in a way that avoids a monochromatic winning line?





Hales–Jewett in Three Dimensions

What about the $3 \times 3 \times 3$ cube?



This problem becomes significantly harder, as there are

$$2^{27} = 134{,}217{,}728 \text{ possible colorings.}$$

From Grids to Higher Dimensions

To make our question precise, we need to formalize **points and lines on hypercubes**.

Let $[k] = \{1, 2, \ldots, k\}$ be an alphabet.

Words of length n are sequences

$$[k]^n = \{(x_1, \ldots, x_n) : x_i \in [k]\}.$$

We think of $[k]^n$ as a k-ary n-dimensional discrete cube.

Defining Lines in Higher Dimensions

To describe lines, we introduce **variable words**.

A **variable word** is a sequence where some coordinates are replaced by a wildcard $\star$:

$$w = (w_1, \ldots, w_n) \in ([k] \cup \{\star\})^n.$$

Combinatorial Lines

Fix a variable word w. Substituting a letter $a \in [k]$ into all wildcard positions gives

$$w(a) = (u_1, \ldots, u_n), \qquad u_i = \begin{cases} a & \text{if } i \in S(w), \\ w_i & \text{if } i \notin S(w). \end{cases}$$

The set

$$L(w) = \{w(a) : a \in [k]\}$$

is called a **combinatorial line**.

This is our **line** for the Hales–Jewett numbers.

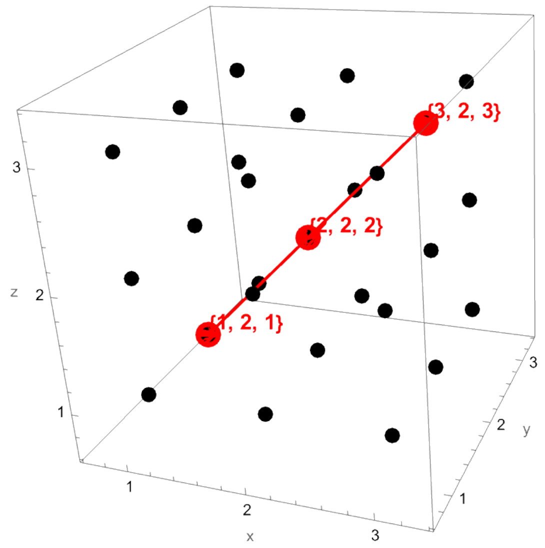
An Example

Example: Let $[3] = \{1, 2, 3\}$ and set

$$w = (\star, 2, \star).$$

Substituting $a \in [3]$ gives

$$L(w) = \{(1, 2, 1), (2, 2, 2), (3, 2, 3)\}.$$



The Hales–Jewett Theorem

Question. Once the cube dimension becomes sufficiently large, do monochromatic combinatorial lines (i.e. winning lines) become unavoidable?

The Hales–Jewett Theorem

Question. Once the cube dimension becomes sufficiently large, do monochromatic combinatorial lines (i.e. winning lines) become unavoidable?

Theorem (Hales–Jewett [1])

For every $k, r \in \mathbb{Z}^+$, there exists an integer n such that for all $d \geq n$, every r-coloring of $[k]^d$ contains a monochromatic combinatorial line.

Hales–Jewett Numbers

Definition

The *Hales–Jewett number* $HJ(k; r)$ is the smallest integer n such that every r-coloring of $[k]^n$ contains a monochromatic combinatorial line.

Hales–Jewett Numbers

Definition

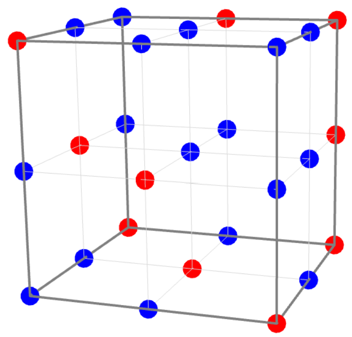
The *Hales–Jewett number* $HJ(k;r)$ is the smallest integer n such that every r-coloring of $[k]^n$ contains a monochromatic combinatorial line.

How **large** are these numbers?

Exact Values

Very little is known about exact Hales–Jewett numbers. The **only** nontrivial value known exactly is

$$HJ(3;2) = 4.$$



Best Known Bounds

For the next simplest cases, we only have bounds:

$$12 \leq HJ(4;2) \leq 10^{11}, \qquad 13 \leq HJ(3;3) \leq 2^{2^{3c}},$$

for some constant c.

Why Brute Force Fails

Case $HJ(4;2) \geq 12$**:**

- Number of colorings of $[4]^{12}$ is $2^{4^{12}} = 2^{16{,}777{,}216}$
- Number of combinatorial lines: 227,363,409

Case $HJ(3;3) \geq 13$**:**

- Number of colorings of $[3]^{13}$ is $3^{3^{13}} \approx 2^{2{,}526{,}942}$
- Number of combinatorial lines is 65,514,541

Exhaustive search is **impossible**, so we need a more efficient way to search.

Boolean Satisfiability (SAT)

Boolean Formulas

A **Boolean formula** is a logical expression built from binary variables. Equivalently, it defines a function

$$f : \{0, 1\}^n \to \{0, 1\}.$$

We combine variables using logical operations:

$$\wedge \text{ (AND)}, \qquad \vee \text{ (OR)}, \qquad \neg \text{ (NOT)}$$

Example:

$$(b_1 \vee \neg b_2) \;\wedge\; (\neg b_3 \vee b_4 \vee \neg b_5)$$

Conjunctive Normal Form (CNF)

We require formulas to be in **conjunctive normal form (CNF).**

A CNF formula is a conjunction of clauses, where each clause is a disjunction of literals.

Example:

$$(b_1 \vee \neg b_2) \ \wedge \ (\neg b_3 \vee b_4 \vee \neg b_5)$$

The SAT Problem

Input: A Boolean formula in conjunctive normal form (CNF)

Question: Does there exist an assignment of values to its variables that makes the formula true?

If such an assignment exists, the formula is **satisfiable**; otherwise, it is **unsatisfiable**.

Why SAT Solvers Work

SAT is NP-complete, yet modern solvers are **remarkably effective** in practice.

They routinely handle formulas with **millions** of variables and clauses, allowing us to explore vast combinatorial spaces efficiently.

How SAT Solvers Work

Example:

$$(\neg b_1 \vee b_3) \ \wedge \ (\neg b_2 \vee b_4) \ \wedge \ (\neg b_3 \vee \neg b_4)$$

Step 1: Assign $b_1 = 1$, $b_2 = 1$

Step 2: Propagation forces

$$b_3 = 1, \quad b_4 = 1$$

Step 3: Conflict!

$(\neg b_3 \vee \neg b_4)$ is false

The Key Idea: Learning

From the conflict, the solver learns:

$$b_1 = 1 \text{ and } b_2 = 1 \Rightarrow \text{contradiction}$$

It adds a new clause:

$$(\neg b_1 \vee \neg b_2)$$

Key insight: The solver learns from conflicts, avoiding repeated mistakes and pruning the search space.

SAT Solvers and Combinatorial Structure

The performance of SAT solvers depends strongly on the **problem structure.**

$HJ(3;3) \geq n$	Incidence-aware	Standard heap
7	77	134
8	307	386
9	1281	3288

Conflict Counts on Hales–Jewett SAT Instances

Key Insight: Combinatorial structure strongly influences solver behavior.

Encoding Hales–Jewett as SAT

Encoding Hales–Jewett as SAT

Goal. For fixed $k, r, n \in \mathbb{Z}^+$, determine whether there exists a coloring

$$[k]^n \to \mathcal{C}_r$$

with no monochromatic combinatorial line.

Encoding Hales–Jewett as SAT

Goal. For fixed $k, r, n \in \mathbb{Z}^+$, determine whether there exists a coloring

$$[k]^n \to \mathcal{C}_r$$

with no monochromatic combinatorial line.

Idea. Create a Boolean formula $\Phi_{[k]^n}$ such that:

- $\Phi_{[k]^n}$ satisfiable $\Rightarrow HJ(k; r) \geq n + 1$
- $\Phi_{[k]^n}$ unsatisfiable $\Rightarrow HJ(k; r) \leq n$

Encoding $HJ(4;2)$: Points

Each point $x \in [4]^n$ is assigned a variable b_x:

$$b_x = 1 \iff x \text{ is red}, \qquad b_x = 0 \iff x \text{ is blue}.$$

Each variable represents the color of a point.

ENCODING $HJ(4;2)$: LINES

Let $L = \{x(1), x(2), x(3), x(4)\} \subseteq [4]^n$ be a combinatorial line.

To ensure L is not monochromatic, we enforce:

$$(\neg b_{x(1)} \vee \neg b_{x(2)} \vee \neg b_{x(3)} \vee \neg b_{x(4)}) \quad \wedge \quad (b_{x(1)} \vee b_{x(2)} \vee b_{x(3)} \vee b_{x(4)})$$

Applying this to all lines $L \in \mathcal{L}$ yields a formula $\Phi_{[4]^n}$, which is satisfiable if and only if $HJ(4;2) > n$.

ENCODING $HJ(3;3)$

In the 3-color case, each point $x \in [3]^n$ is assigned three variables:

$$(b_{x,r}, b_{x,b}, b_{x,g})$$

$$b_{x,r} = 1 \iff x \text{ is red}, \quad b_{x,b} = 1 \iff x \text{ is blue}, \quad b_{x,g} = 1 \iff x \text{ is green}$$

Each triple of variables represents the color of a point. Each point has one (or more) colors.

ENCODING $HJ(3;3)$

Let $L = \{x(1), x(2), x(3)\}$ be a combinatorial line.

To prevent monochromatic lines, we add:

$$\begin{aligned} & (\neg b_{x(1),r} \vee \neg b_{x(2),r} \vee \neg b_{x(3),r}) \\ \wedge \quad & (\neg b_{x(1),b} \vee \neg b_{x(2),b} \vee \neg b_{x(3),b}) \\ \wedge \quad & (\neg b_{x(1),g} \vee \neg b_{x(2),g} \vee \neg b_{x(3),g}) \end{aligned}$$

We also require each point to receive a color:

$$(b_{x,r} \vee b_{x,b} \vee b_{x,g})$$

The resulting formula is satisfiable $\iff HJ(3;3) > n$.

An Improved Hales–Jewett Bound

Main Result

Theorem (Conlon, F-, Robertson)

$$HJ(3;3) \geq 14.$$

Main Result

Theorem (Conlon, F-, Robertson)

$$\boldsymbol{HJ(3;3) \geq 14.}$$

Proof idea: We construct a coloring of $[3]^{13}$ in two stages:

1. Use a van der Waerden coloring to build a structured coloring of $[3]^{12}$
2. Extend this partial coloring to $[3]^{13}$ using a SAT solver

Building the Construction

We begin with a 3-coloring

$$C_{VW} : [1, 26] \to \{\text{red}, \text{blue}, \text{green}\}$$

with no monochromatic 3-term arithmetic progression.

This exists since

$$W(3; 3) = 27.$$

We use this to construct a monochromatic line-avoidant coloring of $[3]^{12}$.

Coloring $[3]^{12}$ from van der Waerden

Define $f : [3]^{12} \to [1, 25]$,

$$f(x) = 1 + \sum_{i=1}^{12} (x_i - 1).$$

We color each point $x \in [3]^{12}$ by

$$C_{12}(x) := C_{VW}(f(x)).$$

Interpretation: Each point in the cube is colored via its image under f.

Why Lines Map to Progressions

Let $L = \{w(1), w(2), w(3)\}$ be a combinatorial line arising from variable word w.

If w has t wildcard positions, then

$$f(w(2)) - f(w(1)) = t, \quad f_(w(3)) - f(w(2)) = t.$$

Thus, the images form a **3-term arithmetic progression** (with common difference t).

Why This Prevents Monochromatic Lines

Suppose a line $L = \{w(1), w(2), w(3)\}$ were monochromatic.

Then

$$C_{VW}(f(w(1))) = C_{VW}(f(w(2))) = C_{VW}(f(w(3))).$$

These these values form a 3-term arithmetic progression - contradiction!

Thus, **no monochromatic lines exist** in $[3]^{12}$ under C_{12}.

Extending to $[3]^{13}$

We view $[3]^{13}$ as three 12-dimensional slices and fix one slice,

$$\{(x_1, \dots, x_{12}, 1) : (x_1, \dots, x_{12}) \in [3]^{12}\}.$$

Color this slice using C_{12}:

$$C_{13}(x_1, \dots, x_{12}, 1) := C_{12}(x_1, \dots, x_{12}).$$

This produces a **partial monochromatic line-avoidant coloring** of $[3]^{13}$.

Completing the Construction via SAT

We have colored one slice of $[3]^{13}$.

We encode these constraints as a SAT instance and use a solver to search for a valid completion.

Result: After 360 CPU-days, we identify 9 full colorings of $[3]^{13}$ with no monochromatic combinatorial lines.

Thus,

$$HJ(3;3) \geq 14.$$

Hales–Jewett Variants

Geometric Lines

A *geometric line* is defined by a geometric variable word,

$$w^* \in ([k] \cup \{\star, \diamond\})^n,$$

containing at least one $\star$ and at least one $\diamond$.

For each $a \in [k]$, define a point $w^*(a)$ by:

$$w^*(a)_i = \begin{cases} a & \text{if } w_i^* = \star, \\ k + 1 - a & \text{if } w_i^* = \diamond, \\ w_i^* & \text{otherwise.} \end{cases}$$

Some coordinates **increase** while others **decrease.**

Example of a Geometric Line

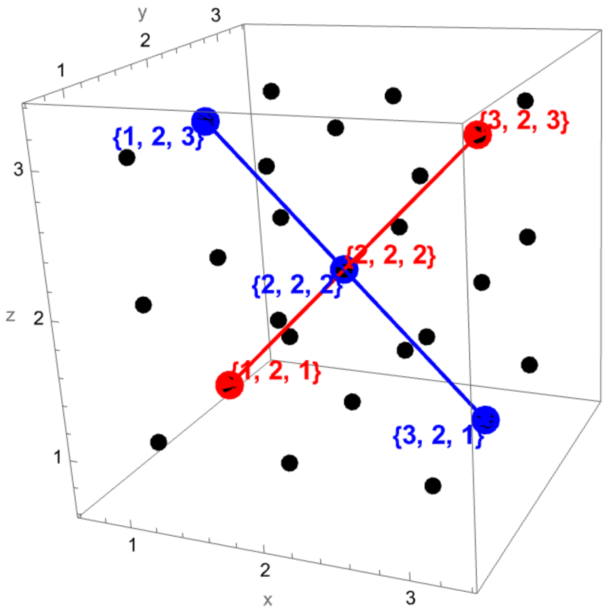
Let $k = 3$, $n = 3$, and consider the geometric variable word

$$w^* = (\star, 2, \diamond).$$

For each $a \in [3] = \{1, 2, 3\}$, we construct points $w^*(a)$:

$$w^*(1) = (1, 2, 3),$$
$$w^*(2) = (2, 2, 2),$$
$$w^*(3) = (3, 2, 1).$$

Visualizing Geometric and Combinatorial Lines



Geometric Variants

Geometric Variants

- $G(k; r)$: Hales–Jewett number for geometric lines
- $HJ^*(k; r)$: Hales–Jewett number for combinatorial or geometric lines

Geometric Variants

Geometric Variants

- $G(k; r)$: Hales–Jewett number for geometric lines
- $HJ^*(k; r)$: Hales–Jewett number for combinatorial or geometric lines

Results:

$\boldsymbol{k}$	$G(k; 2)$	$HJ^*(k; 2)$
2	3	2
3	**5**	**3**
4	$\geq$ **8**	$\geq$ **7**

Off-diagonal Hales–Jewett Numbers

A **partial *m*-line** is a subset of a line consisting of $m \leq k$ consecutive points.

Off-diagonal Hales–Jewett Numbers

A **partial *m*-line** is a subset of a line consisting of $m \leq k$ consecutive points.

Off-diagonal Hales–Jewett Numbers

Let $m \leq k$. Then $HJ(k, m; 2)$ is the minimal n such that every 2-coloring of $[k]^n$ admits:

- a red combinatorial k-line, or
- a blue combinatorial m-line.

Given $j \leq m \leq k$, $HJ(k, m, j; 3)$ guarantees at least one of:

- red: k-line, blue: m-line, green: j-line.

Analogous extensions apply to $G(k; r)$ and $HJ^*(k; r)$.

Off-diagonal Results

k	$HJ(k,2;2)$	$G(k,2;2)$	$HJ^*(k,2;2)$
2	2	**3**	**2**
3	3	**4**	**2**
4	4	**4**	**3**
5	5	**6**	**4**
6	**6**	**6**	**4**
7	**7**	$\geq$ **8**	**5**
8	$\geq$ **7**	**7**	**4**

3-Color Results:

$$\boldsymbol{HJ(3,2,2;3) = 5, \quad HJ(3,3,2;3) = 9}$$

A Conjecture

Conjecture

For all $k \geq 2$,

$$HJ(k, 2; 2) = k.$$

Theorem (Conlon, F-, Robertson)

$$HJ(k, 2; 2) \leq k \quad \text{and} \quad HJ(k, 2; 2) \geq k - t$$

where $k - t$ is the largest prime $\leq k$.

The Remaining Gap

If p is prime, then

$$HJ(p, 2; 2) = p.$$

To prove the conjecture, it suffices to show:

$$HJ(k, 2; 2) \geq k \quad \text{for composite } k.$$

Future Directions

Future Directions

- The resolution of the off-diagonal Hales–Jewett conjecture
- The improvement of SAT solver performance on Hales–Jewett instances
- The application of SAT to related combinatorial invariants

Acknowledgements and References

Acknowledgments

This work was completed under the supervision of Professor Aaron Robertson, whose guidance, feedback, and continued support were essential throughout the project.

Gratitude is extended to Professor Nick Moore and the Colgate Mathematics Seminar for the opportunity to share this work.

References

- Aaron Robertson. *Fundamentals of Ramsey Theory.* CRC Press (2021). DOI: 10.1201/9780429431418.
- B. L. van der Waerden. *Beweis einer Baudetschen Vermutung.* Nieuw Archief voor Wiskunde. Tweede Serie, 15:212–216 (1927).

Appendix A: $HJ(p, 2) \geq p$

Consider a 2-coloring of $[p]^{p-1}$ and color a point x blue iff

$$\sum_{i=1}^{p-1} x_i \equiv 0 \pmod{p}.$$

Along a line the sum becomes

$$a + td \pmod{p},$$

and we obtain a blue point exactly when $a + td \equiv 0 \pmod{p}$.
Since p is prime, d is invertible mod p, so a solution

$$t \equiv -ad^{-1} \pmod{p}$$

always exists.