<u>The Hales-Jewett Theorem</u>

1. **Quote.** Last year I went fishing with Salvador Dali. He was using a dotted line. He caught every other fish.

   Steven Alexander Wright, American comedian, actor and writer, 1955-

2. **Alphabet.** For $m \in \mathbb{N}$ any set A such that $|A| = m$ is called an alphabet on m symbols.

   Example:

3. **Words.** Let A be an alphabet on m symbols. For $n \in \mathbb{N}$, any function $w : [1, n] \to A$ is called a word of length n on the alphabet A. If $w(i) = a_i$, $i \in [1, n]$ then we write

$$w = a_1 a_2 \cdots a_n.$$

   The set of all words of length n on the alphabet A is denoted as A^n. We say that A^n is *the n-dimensional cube on alphabet A*.

   Example:

4. **Roots.** Let A be an alphabet (on m symbols) and let $*$ be a symbol such that $* \notin A$. We consider the alphabet $A_* = A \cup \{*\}$. Any word on the alphabet A_*, i.e, any element of $(A_*)^n = A_*^n$, for some $n \in \mathbb{N}$, that contains the symbol $*$ is called a root.

   Example:

5. **Words From Roots.** For a root $\tau \in A_*^n$ and a symbol $a \in A$ we define the word $\tau_a \in A^n$ in the following way. For $i \in [1, n]$

$$\tau_a(i) = \begin{cases} \tau(i) & \text{if} \quad \tau(i) \neq *, \\ a & \text{if} \quad \tau(i) = *. \end{cases}$$

6. **Example.** Let $A = \{a, b, c\}$ and let $\tau = * \ b \ c \ b \in A_*^4$ be a root. Then

$$\begin{aligned}
\tau_a &= \square \ b \ c \ b \\
\tau_b &= \square \ b \ c \ b \\
\tau_c &= \square \ b \ c \ b.
\end{aligned}$$

7. **Example:** Let $A = [1, 4]$ and let $\tau = * \ 1 \ 3 \ * \ 4 \ * \in A_*^6$ be a root. Then

$$\tau_2 = \square \ 1 \ 3 \ \square \ 4 \ \square.$$

8. **Combinatorial Line:** Let A be an alphabet, let $n \in \mathbb{N}$, and let $\tau \in A_*^n$ be a root. A combinatorial line in A^n rooted in τ is the set of words

$$L_\tau = \{\tau_a : a \in A\}.$$

Note: $L\tau \subseteq A^n$.

9. **Example.** Let $A = \{1, 2, 3\}$ and $n = 2$. Find all combinatorial lines in A^2.

(a) All roots in A_*^2:

$$\begin{aligned}
\tau &= * \ 1 \\
\sigma &= * \ 2 \\
\theta &= * \ 3 \\
\rho &= 1 \ * \\
\chi &= 2 \ * \\
\phi &= 3 \ * \\
\mu &= * \ * \, .
\end{aligned}$$

(b) All combinatorial lines:

$$\begin{aligned}
L_\tau &= \{1\ 1, 2\ 1, 3\ 1\} \\
L_\sigma &= \{1\ 2, 2\ 2, 3\ 2\} \\
L_\theta &= \{1\ 3, 2\ 3, 3\ 3\} \\
L_\rho &= \{1\ 1, 1\ 2, 1\ 3\} \\
L_\chi &= \{2\ 1, 2\ 2, 2\ 3\} \\
L_\phi &= \{3\ 1, 3\ 2, 3\ 3\} \\
L_\mu &= \{1\ 1, 2\ 2, 3\ 3\}.
\end{aligned}$$

(c) All combinatorial lines - Another view:

L_τ	L_σ	L_θ	L_ρ	L_χ	L_ϕ	L_μ
1 1	1 2	1 3	1 1	2 1	3 1	1 1
2 1	2 2	2 3	1 2	2 2	3 2	2 2
3 1	3 2	3 3	1 3	2 3	3 3	3 3

(d) All combinatorial lines - Another view:

10. **Loks Like. . .**

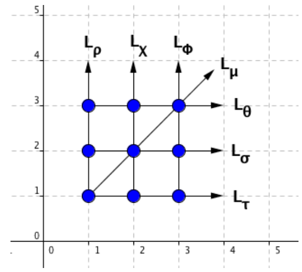


Figure 1: All combinatorial lines in $[1,3]^2$.

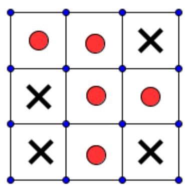


Figure 2: Tic-Tac-Toe: it's a win!

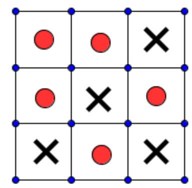


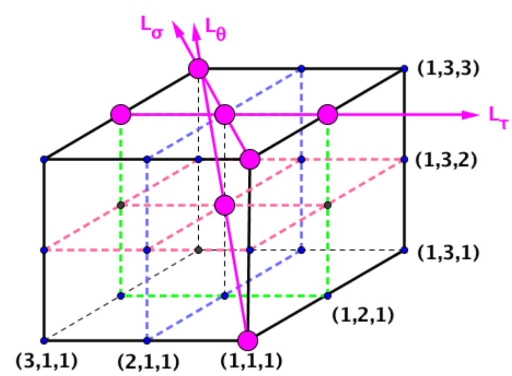
Figure 3: Tic-Tac-Toe: it's a win!

11. **What About Combinatorial Lines in** $[1,3]^3$**?** Consider roots:

$$\tau = *\,2\,3, \quad \sigma = *\,*\,3, \quad \theta = *\,*\,*.$$

Then

L_τ	L_σ	L_θ
123	113	111
223	223	222
323	333	333



We observe that the points in $\mathbb{R}^3$ that correspond to the elements of the combinatorial line L_τ lie on the Euclidean line ℓ_τ with the parametric equations $x = t, y = 2, z = 3, t \in \mathbb{R}$. Similarly, the combinatorial line L_σ corresponds to a set of points on the line ℓ_σ with the parametric equations $x = t, y = t, z = 3, t \in \mathbb{R}$, and L_θ corresponds to a set of points on the line ℓ_θ with the parametric equations $x = t, y = t, z = t, t \in \mathbb{R}$.

In general, for $m \in \mathbb{N}$, the combinatorial line determined by a root $\mu = a_1\, a_2\, a_3 \in [1,m]^3_*$ corresponds to a set of points on the Euclidean line ℓ_μ with the parametric equations $x = b_1 + \alpha_1 \cdot t, y = b_2 + \alpha_2 \cdot t, z = b_3 + \alpha_3 \cdot t, t \in \mathbb{R}$, where $b_i = 0$ and $\alpha_i = 1$ if $a_i = *$, and $b_i = a_i$ and $\alpha_i = 0$ if $a_i \in [1,m]$. This set of points is obtained for the values $t \in [1,m]$.

12. $4 \times 4 \times 4$ **Tic-Tac-Toe**

"A placement game played on the 64 positions in a 3D board with four levels and a four by four grid on each level. The object is to get four in a row, whether on one level or between all four levels, straight or diagonal. It is essentially a Tic-Tac-Toe game in a 4x4x4 cube.

The cube needs to be emptied before start. Each player choose a colour of their pieces and take turn to place their pieces in vacant positions, one piece at a time. The player wins who first gets four in a row of his own pieces - either horizontal, vertical or diagonal. If all positions are occupied without anyone having four in a row, then the game is called a tie."

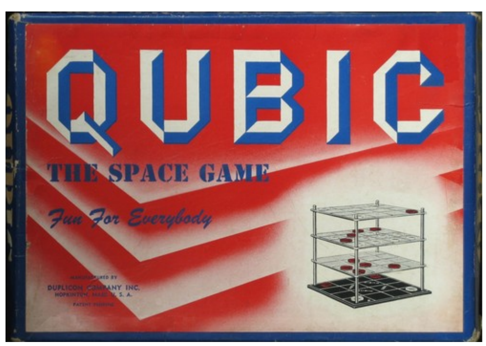


Figure 4: The first edition of Qubic by any company was produced by Duplicon in 1946 or 1947.

13. **Hint.** If you create a monochromatic combinatorial line in $[1,4]^3$ you win!

14. **Example:** We illustrate the fact that each of the combinatorial lines in $[1,4]^3$ corresponds to a winning position in the Qubic game by the following example.

    For $a, b \in [1,4]$, consider the root $\tau = *\ a\ b \in [1,4]^3_*$ and the Euclidean line ℓ_τ in $\mathbb{R}^3$ given by its parametric equations $x = t, y = a, z = b, t \in \mathbb{R}$. Hence, ℓ_τ is the line that passes through the point $(0, a, b)$ and is parallel to the x–axis. Now, the combinatorial line $L_\tau = \{1\ a\ b, 2\ a\ b, 3\ a\ b, 4\ a\ b\}$ corresponds to the winning position $\{(1, a, b), (2, a, b), (3, a, b), (4, a, b)\} = \{(t, a, b) : t \in [1,4]\} \subseteq \ell_\tau \cap Q(4,3)$.

    Observe that the line $x = t, y = 5 - t, z = 1, t \in \mathbb{R}$, contains the winning position

    $$\{(1,4,1), (2,3,1), (3,2,1), (4,1,1)\} = \{(t, 5-t, 1) : t \in [1,4]\}$$

    that does not correspond to any of the combinatorial lines in $[1,4]^3$.

15. **Two exercises** Let $m, n \in \mathbb{N}$ and let $|A| = m$, i.e. let A be an alphabet on m symbols.

    (a) Prove that the number of Euclidean lines that intersect the cube A^n at m points is $\frac{(m+2)^n - m^n}{2}$.

    (b) Prove that the number of combinatorial lines in A^n equals to $(m+1)^n - m^n$.

16. **Generalization:** But what if one considers a k–player game played on the "board" $Q(m,n) = \{(x_1, x_2, \ldots, x_n) : x_1, x_2, \ldots, x_n \in [1,m]\}$?

    Observe that $(x_1, x_2, \ldots, x_n) \in Q(m,n)$ is a point in $\mathbb{R}^n$ with $x_1, x_2, \ldots, x_m \in [1,m] = \{1,2,\ldots,m\}$. With $[1,m]^n = \{x_1 x_2 \cdots x_n : x_1, x_2, \ldots, x_n \in [1,m]\}$ we denote the n–cube on the alphabet $[1,m]$.

    Similar to theTic–Tac game–Toe and the Qubic game, each player is given one of k different colours and tasked to colour one point in $Q(m,n)$ at each turn. The player who first completes a monochromatic line wins. Here "a line" means a set of m collinear points, i.e. a set of m points in $Q(m,n)$ that lie on a line in $\mathbb{R}^n$ given by its parametric equations $x_1 = a_1 + \alpha_1 \cdot t, x_2 = a_2 + \alpha_2 \cdot t, \ldots, x_n = a_n + \alpha_n \cdot t$, $t \in \mathbb{R}$, for some fixed real numbers a_i, α_i, $i \in [1,n]$.

    Graham, Rothschild, and Spencer called the above generalization of Tic–Tac–Toe a "k–person n dimensional Tic–Tac–Toe m–in a row" game.

17. **Proposition.**Any combinatorial line in $[1,m]^n$ corresponds to a winning position in a k–person n dimensional Tic–Tac–Toe m–in a row game.

    Proof For given $m,n \in \mathbb{N}$, we consider a root $\tau = a_1\, a_2 \ldots a_n \in [1,m]^n_*$. Let the line ℓ_τ in $\mathbb{R}^n$ be given by its parametric equations

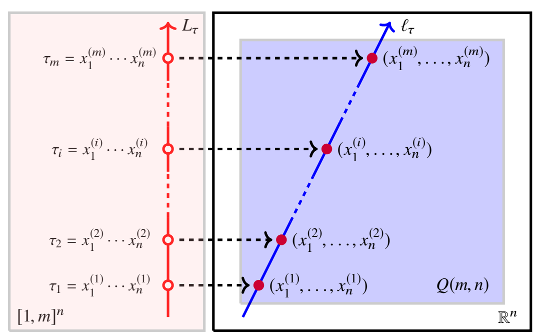
    $$x_1 = b_1 + \alpha_1 \cdot t, x_2 = b_2 + \alpha_2 \cdot t, \ldots, x_n = b_n + \alpha_n \cdot t, t \in \mathbb{R},$$

    where $b_i = 0$ and $\alpha_i = 1$ if $a_i = *$, and $b_i = a_i$ and $\alpha_i = 0$ if $a_i \in [1,m]$.

    Recall that for $i \in [1,m]$, the word $\tau_i = x_1^{(i)}\, x_2^{(i)} \cdots x_n^{(i)} \in [1,m]^n$ is such that $x_j^{(i)} = i$ if $a_j = *$ and $x_j^{(i)} = a_j$ and if $a_j \in [1,m]$.

    It follows that, by taking $t = i$ in the parametric equation for ℓ_τ, the point $(x_1^{(i)}, x_2^{(i)}, \ldots, x_n^{(i)}) \in \ell_\tau \cap Q(m,n)$.

    This establishes an injection between the combinatorial line $L_\tau = \{\tau_i : i \in [1,m]\}$ and the set $\ell_\tau \cap Q(m,n)$. Since $|L_\tau| = m$, the combinatorial line L_τ corresponds to a winning position.

    Suppose that a k–person n dimensional Tic–Tac–Toe m–in a row game ended up in a tie. Hence, the set $Q(m,n)$ was partitioned into k mutually disjunct parts, $Q(m,n) = P_1 \cup \ldots \cup P_k$, $P_i \cap P_j = \emptyset$ if $i \neq j$, in such a way that none of the parts contained a set of m collinear points.

    Let us define a k–colouring C of the cube $[1,m]^n$ in the following way: for $a_1 \cdots a_n \in [1,m]^n$, $C(a_1 \cdots a_n) = i$ if and only if $(a_1, \ldots, a_k) \in P_i$. By Proposition and our assumption that the game ended up in a tie, it follows that the colouring C of the n–dimensional cube on alphabet $[1,m]$ does not contain a monochromatic combinatorial line.

18. **Question:** In the spirit of Ramsey theory, this observation leads us to the following question: Let A be an alphabet on m symbols and let A^n be the n–dimensional cube on alphabet A, i.e. let $A^n = \{a_1 a_2 \cdots a_n : a_i \in A, i \in [1, n]\}$. If A^n is k–coloured, under which conditions can we be sure that A^n contains a monochromatic combinatorial line?

19. **More precisely . . .** Let $m, k \in \mathbb{N}$ and let A be an alphabet on m symbols. Does there exist an $n \in \mathbb{N}$ such that whenever A^n is k-coloured there exists a monochromatic line?

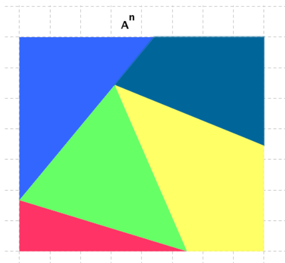


Figure 5: A^n is k-coloured

$$\begin{array}{ccccc} a_1 & 1 & a_3 & \cdots & a_n \\ a_1 & 2 & a_3 & \cdots & a_n \\ \vdots & & & & \\ a_1 & m & a_3 & \cdots & a_n \end{array}$$

Figure 6: A red combinatorial line

20. **The Hales-Jewett Theorem.** A nice answer to Question **??** would be that, for any $k, m \in \mathbb{N}$ and for a big enough natural number n, the k–person n dimensional Tic–Tac–Toe m–in a row game cannot end up in a tie.

    This is exactly what Hales and Jewett discussed in their paper *Regularity and position games*, published in 1963. Their famous Hales–Jewett theorem establishes that, if the dimension is sufficiently large, a generalized Tic–Tac–Toe game never ends up in a tie.

    Let $m, k \in \mathbb{N}$ and let A be an alphabet on m symbols. There exists an $n \in \mathbb{N}$ such that whenever A^n is k-coloured there exists a monochromatic line.

    Note: The smallest such n is denoted by $HJ(m, k)$.

21. **Alfred W. Hales and Robert I. Jewett:** Alfred Washington Hales and Robert Israel Jewett, 1937–2022, are American mathematicians. Both of them had long and distinguished academic careers, Hales at the University of California Los Angeles and Jewett at the University of Western Washington.

    When they submitted the *Regularity and position games* paper in 1961, Hales was 23 years old and Jewett was 24. Both were doctoral students at the time. Hales was working under the supervision of Robert P. Dilworth at the California Institute of Technology (Caltech) and Jewett was supervised by Karl Stromberg at the University of Oregon.

    The pair knew each other from their time as undergraduate students at Caltech.

    In 1971, Hales and Jewett, together with Ronald Graham, Klaus Leeb, and Bruce Rothschild, were the first recipients of the George Pólya Prize.