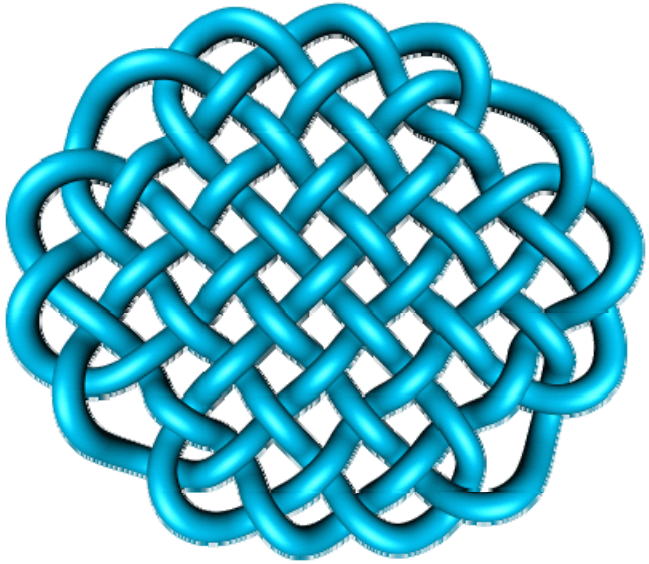
ESF
EIROPAS SOCIĀLAIS FONDS
EIROPAS SAVIENĪBA
IEGULDĪJUMS TAVĀ NĀKOTNĒ



# Nonconstructive Language Recognition by Finite Automata

**Kaspars Balodis, Rūsiņš Freivalds, Lauma Pretkalniņa,**

**Inga Rumkovska, Madars Virza**

European Social Fund project “Datorzinātnes pielietojumi un tās saiknes ar kvantu fiziku”
Nr.2009/0216/1DP/1.1.1.2.0/09/APIA/VIAA/044

# David Hilbert

- **Born: 23 Jan 1862 in Königsberg, Prussia (now Kaliningrad, Russia)**
  **Died: 14 Feb 1943 in Göttingen, Germany**

# Paul Albert Gordan

- **Born: 27 April 1837 in Breslau, Germany (now Wroclaw, Poland)**
  **Died: 21 Dec 1912 in Erlangen, Germany**

# MATHEMATISCHE ANNALEN.

BEGRUNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN.

Unter Mitwirkung der Herren

PAUL GORDAN, CARL NEUMANN, MAX NOETHER,
KARL VONDERMÜHLL, HEINRICH WEBER

gegenwärtig herausgegeben

von

**Felix Klein**
in Göttingen

**Walther Dyck**
in München

**Adolph Mayer**
in Leipzig.

43. Band.






## Ueber die vollen Invariantensysteme.


Von

DAVID HILBERT in Königsberg i./Pr.

# Ernst Friedrich Ferdinand Zermelo

- **Born: 27 July 1871 in Berlin, Germany**
  **Died: 21 May 1953 in Freiburg im Breisgau, Germany**

1904

# Neuer Beweis für die Möglichkeit einer Wohlordnung.

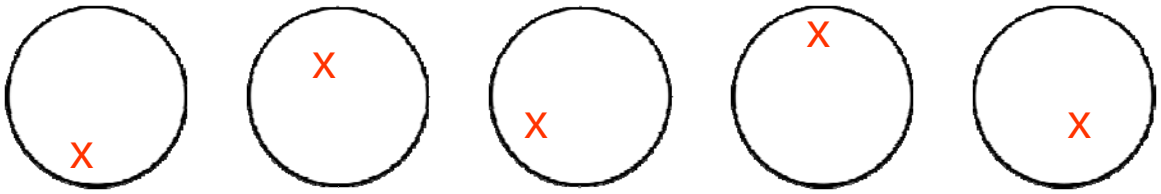
Von

E. Zermelo in Göttingen.

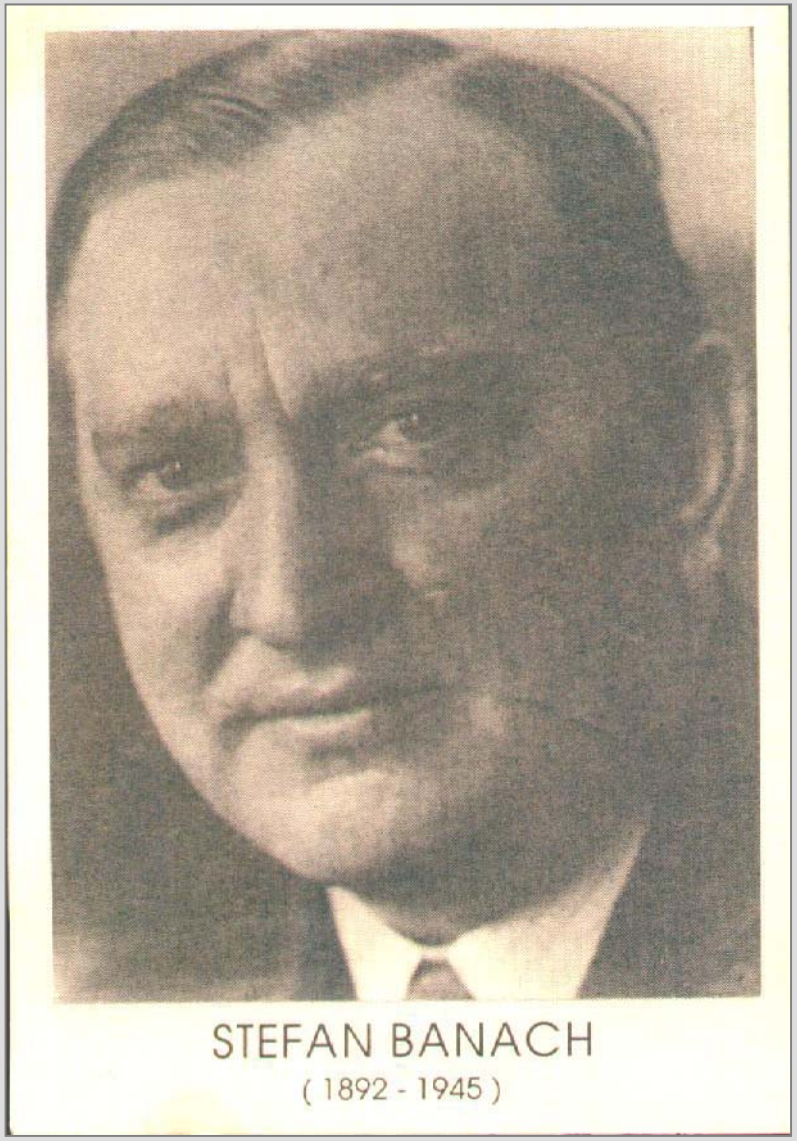
---

Obwohl ich meinen im Jahre 1904 veröffentlichten „Beweis, daß jede Menge wohlgeordnet werden kann“*) gegenüber den verschiedenen im § 2 ausführlich zu besprechenden Einwendungen noch heute vollkommen aufrecht erhalte, dürfte doch der hier folgende neue Beweis desselben Theorems nicht ohne Interesse sein, da er einerseits keine speziellen Lehrsätze der Mengentheorie voraussetzt, andererseits aber den rein formalen Charakter der Wohlordnung, die mit räumlich-zeitlicher Anordnung gar nichts zu tun hat, deutlicher als der erste Beweis hervortreten läßt.

## Axiom of choice



$$\forall a\ (a \neq \varnothing\ \wedge\ \forall b\ (b \in a \rightarrow b \neq \varnothing)\ \wedge$$
$$\wedge\ \forall b_1 \forall b_2\ (b_1 \neq b_2\ \wedge\ \{b_1,\ b_2\} \subseteq a \rightarrow b_1 \cap b_2 = \varnothing) \rightarrow$$
$$\rightarrow \exists d \forall b\ (b \in a \rightarrow \exists c\ (b \cap d = \{c\})))$$




Alfred Tarski (1902 – 1983)

1924

# Sur la décomposition des ensembles de points en parties respectivement congruentes.

Par

St. Banach (Lwów) et A. Tarski (Varsovie).

Nous étudions dans cette Note les notions de *l'équivalence des ensembles de points par décomposition finie*, resp. *dénombrable*. Deux ensembles de points situés dans un espace métrique sont dits équivalents par décomposition finie (ou dénombrable), lorsqu'ils peuvent être décomposés en un nombre fini et égal (ou une infinité dénombrable) de parties disjointes respectivement congruentes.

Les principaux résultats contenus dans le présent article sont les suivants:

*Dans un espace euclidien à $n \geqslant 3$ dimensions deux ensembles arbitraires, bornés et contenant des points intérieurs (p. ex. deux sphères à rayons différents), sont équivalents par décomposition finie.*

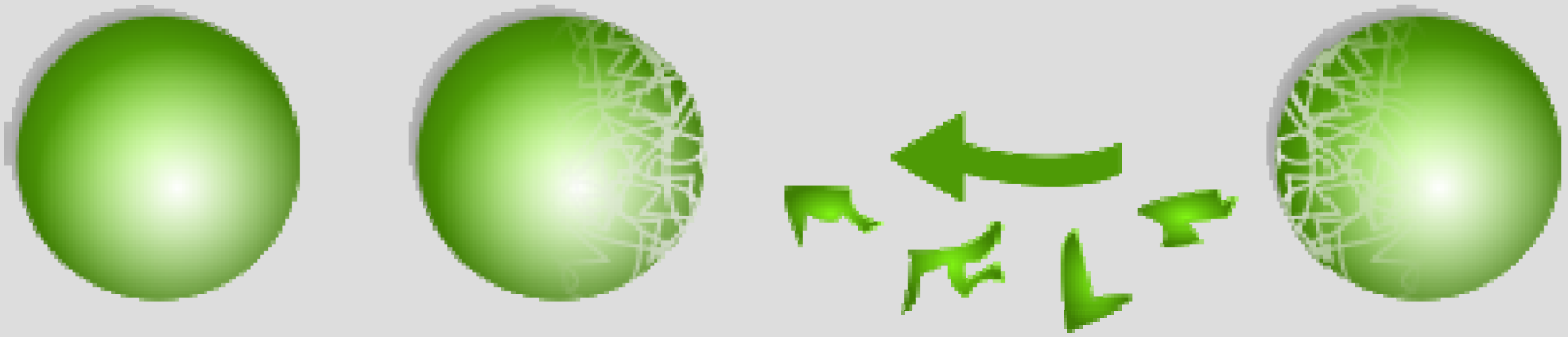
*Un théorème analogue subsiste pour les ensembles situés sur la surface d'une sphère; mais le théorème correspondant concernant l'espace euclidien à 1 ou 2 dimensions est faux.*

D'autre part:

*Dans un espace euclidien à $n \geqslant 1$ dimensions deux ensembles arbitraires (bornés ou non), contenant des points intérieures, sont équivalents par décomposition dénombrable.*

La démonstration des théorèmes précédents s'appuie sur les résultats de MM. Hausdorff, Vitali et Banach [1]), qui concernent le probléme général de mesure; elle fait donc usage de *l'axiome*

[1]) F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig 1914, p. 401 et 469.
G. Vitali, *Sul problema della mesura dei gruppi di punti di una retta*, Bologna 1905.



# Luitzen Egbertus Jan Brouwer

- **Born:** 27 Feb 1881 in Overschie (now a suburb of Rotterdam), Netherlands
- **Died:** 2 Dec 1966 in Blaricum, Netherlands

Original 1912
English 1913

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 37, Number 1, Pages 55–64
S 0273-0979(99)00802-2
Article electronically published on December 21, 1999

# INTUITIONISM AND FORMALISM

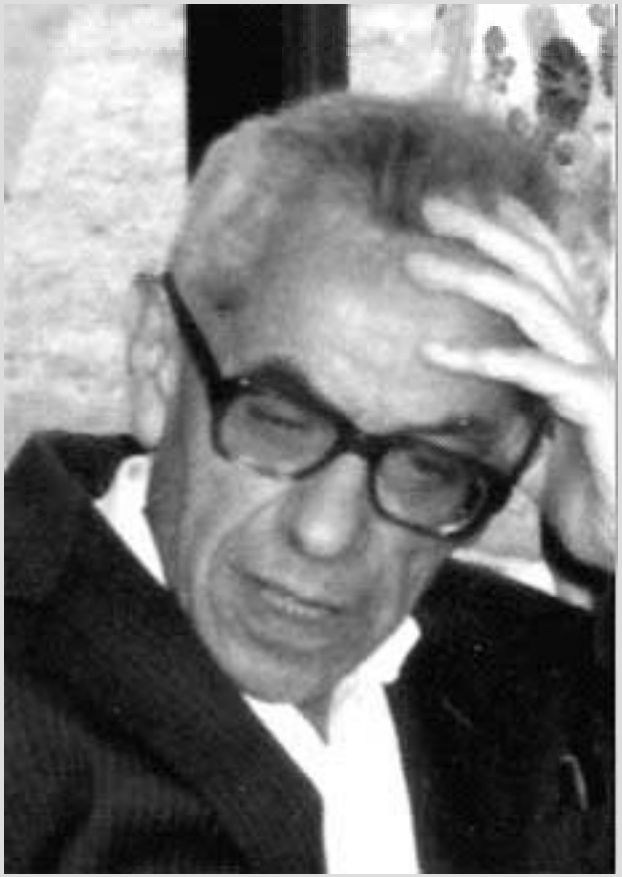
DR. L. E. J. BROUWER

The subject for which I am asking your attention deals with the foundations of mathematics. To understand the development of the opposing theories existing in this field one must first gain a clear understanding of the concept "science"; for it is as a part of science that mathematics originally took its place in human thought.

By science we mean the systematic cataloguing by means of laws of nature of causal sequences of phenomena, i. e., sequences of phenomena which for individual or social purposes it is convenient to consider as repeating themselves identically,—and more particularly of such causal sequences as are of importance in social relations.

That science lends such great power to man in his action upon nature is due to the fact that the steadily improving cataloguing of ever more causal sequences of phenomena gives greater and greater possibility of bringing about desired phenomena, difficult or impossible to evoke directly, by evoking other phenomena connected with the first by causal sequences. And that man always and everywhere creates order in nature is due to the fact that he not only isolates the causal sequences

# Paul Erdös



- **Born: 26 March 1913 in Budapest, Hungary**
- **Died: 20 Sept 1996 in Warsaw, Poland**

1947

# SOME REMARKS ON THE THEORY OF GRAPHS

P. ERDÖS

The present note consists of some remarks on graphs. A graph $G$ is a set of points some of which are connected by edges. We assume here that no two points are connected by more than one edge. The complementary graph $G'$ of $G$ has the same vertices as $G$ and two points are connected in $G'$ if and only if they are not connected in $G$.

A special case of a theorem of Ramsey can be stated in graph theoretic language as follows:

There exists a function $f(k, l)$ of positive integers $k, l$ with the following property. Let there be given a graph $G$ of $n \geqq f(k, l)$ vertices. Then either $G$ contains a complete graph of order $k$, or $G'$ a complete graph of order $l$. (A complete graph is a graph any two vertices of which are connected. The order of a complete graph is the number of its vertices.)

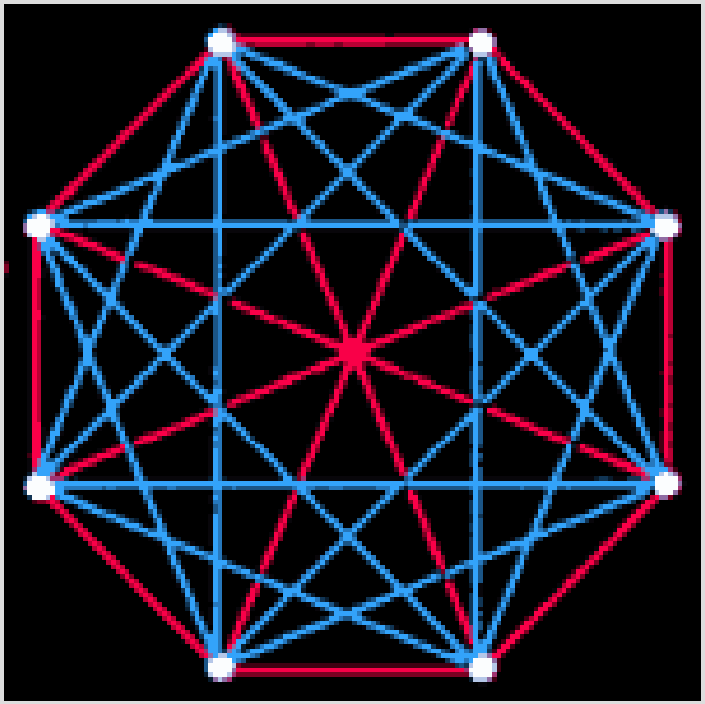
It would be desirable to have a formula for $f(k, l)$. This at present we can not do. We have however the following estimates:

THEOREM I. *Let $k \geqq 3$. Then*

$$2^{k/2} < f(k, k) \leqq C_{2k-2,k-1} < 4^{k-1}.$$

The second inequality of Theorem I was proved by Szekeres,[1] thus we only consider the first one. Let $N \leqq 2^{k/2}$. Clearly the number of different graphs of $N$ vertices equals $2^{N(N-1)/2}$. (We consider the vertices of the graph as distinguishable.) The number of different graphs containing a given complete graph of order $k$ is clearly $2^{N(N-1)/2}/2^{k(k-1)/2}$. Thus the number of graphs of $N \leqq 2^{k/2}$ vertices containing a complete graph of order $k$ is less than

$$(1) \qquad C < \frac{2^{N(N-1)/2}}{N^k} \cdot \frac{2^{N(N-1)/2}}{2^{N(N-1)/2}} = 2^{N(N-1)/2}$$

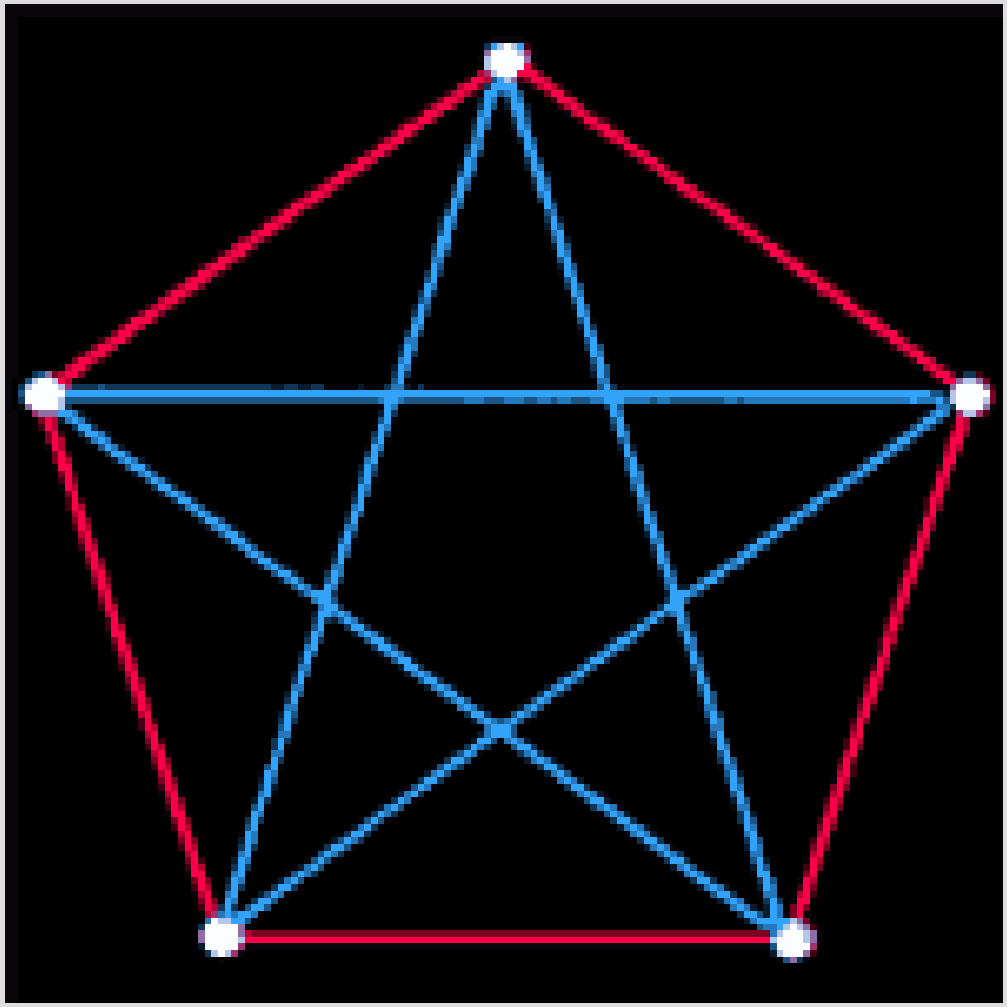


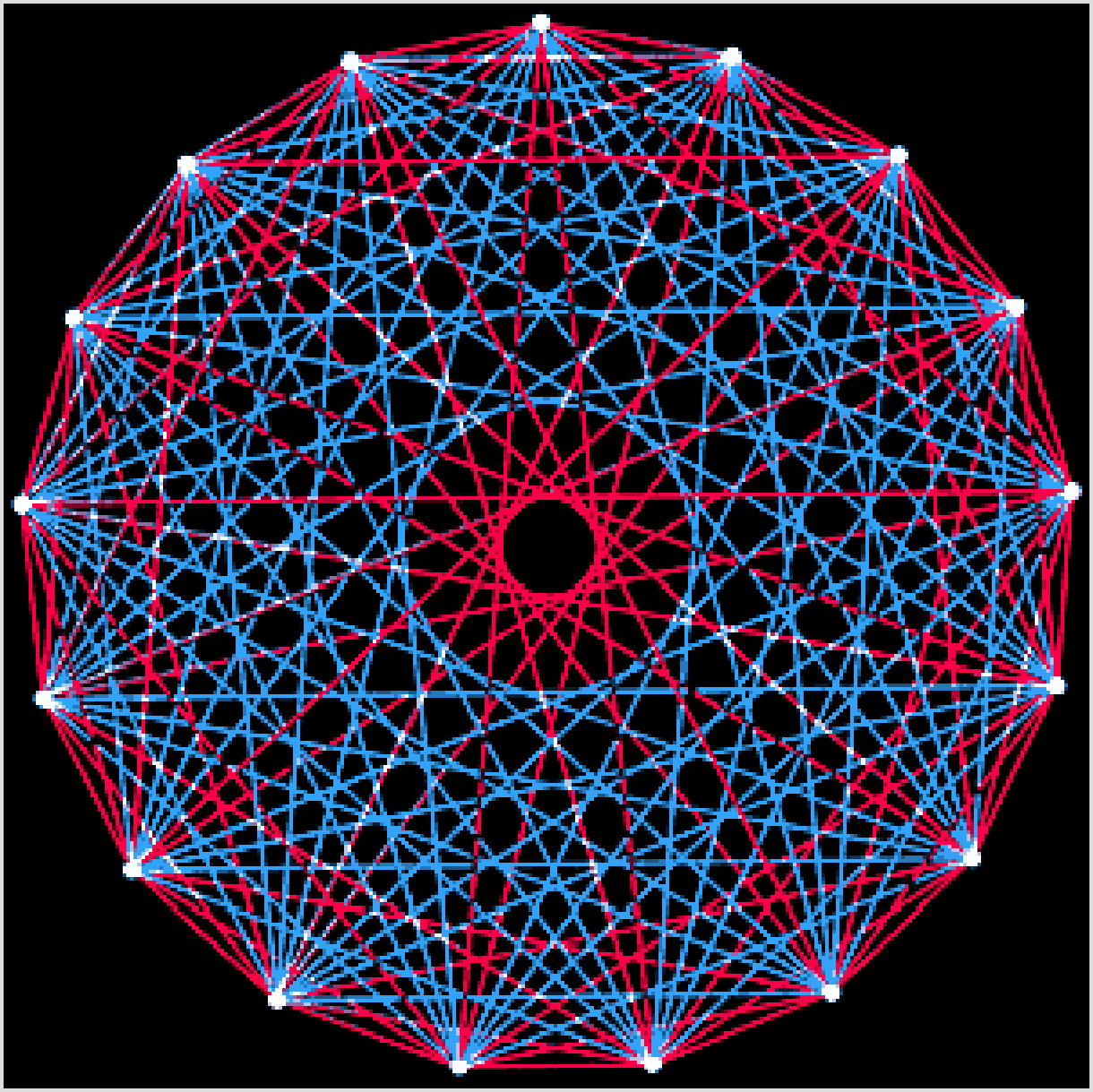
**Ramsey number R(k,k)** is the least natural number n such that for an arbitrary 2-coloring of the edges of a complete graph with n vertices, contains a k-vertice monochromatic subgraph.

Frank Plumpton Ramsey in 1929 proved that for arbitrary natural number k the value of R(k,k) is finite. However, how to find it?

# Frank Plumpton Ramsey

- **Born:** 22 Feb 1903 in Cambridge, England
- **Died:** 19 Jan 1930 in London, England





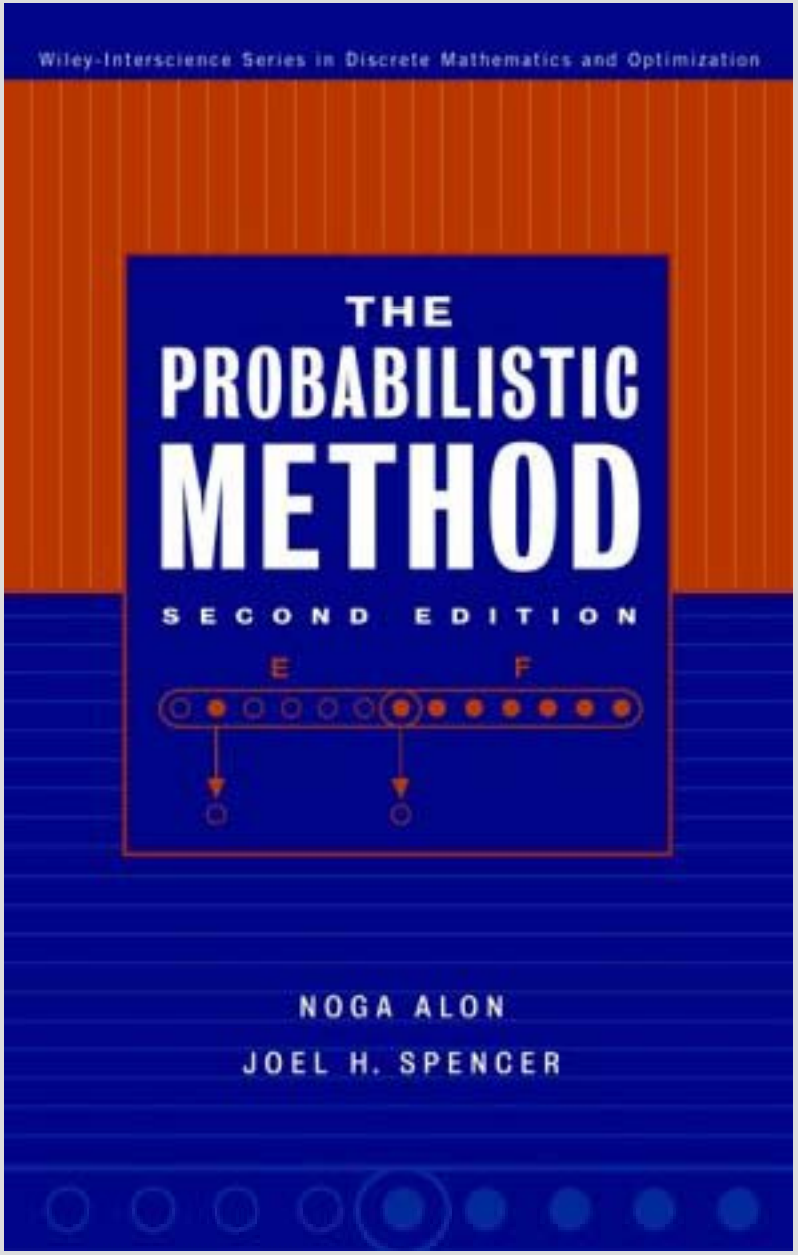
**Theorem (Erdös, 1947)** If $k \geq 3$, then $R(k,k) \geq \lfloor 2^{\frac{k}{2}} \rfloor$.

**Proof.** We can choose $k$ out of $n$ vertices in the graph in $\frac{n!}{k!(n-k)!}$ ways.

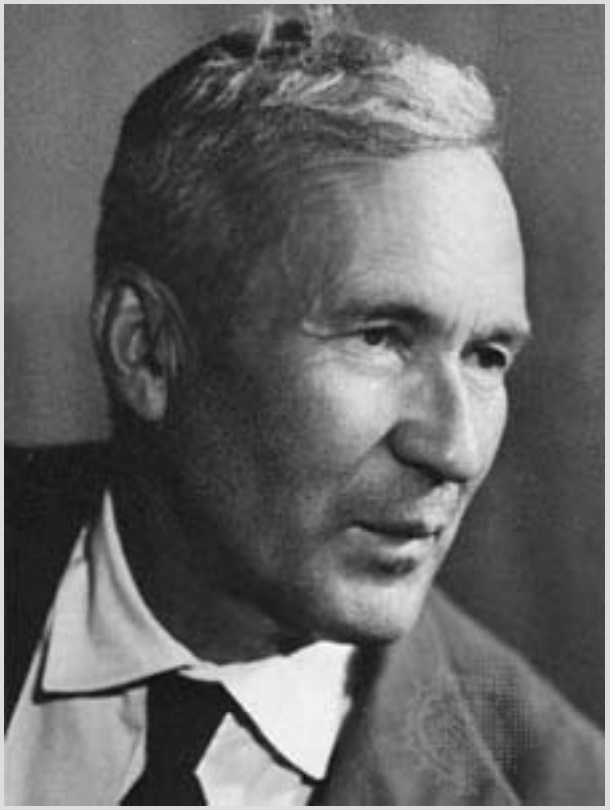
An edge coloring of the graph with $n$ vertices is performed at random. The probability of a fixed graph with $k$ vertices to be monochromatic equals $2^{1-\frac{k(k-1)}{2}}$. Hence the probability to have no monochromatic subgraph with $k$ vertices does not exceed

$$\frac{n!}{k!(n-k)!} 2^{1-\frac{k(k-1)}{2}}$$

If we take $n = \lfloor 2^{\frac{k}{2}} \rfloor$, then this probability is strictly less than 1.


Wiley-Interscience Series in Discrete Mathematics and Optimization
THE
PROBABILISTIC
METHOD
SECOND EDITION
E
F
NOGA ALON
JOEL H. SPENCER

# Andrei Nikolaevich Kolmogorov



- April 12 (25), 1903, Tambov, Russia
- October 20, 1987, Moscow, Soviet Union

**Theorem (Martin-Löf, 1966)** "Nearly all" binary sequences of infinite length are such that Kolmogorov complexity of every its initial fragment of the length $n$ is no less than $n - \log_2 n$ and for infinitely many initial fragments Kolmogorov complexity is no less than $n - \log_2 \log_2 n$ .

**Theorem (J.Bārzdiņš, 1968).** Kolmogorov complexity of arbitrary recursively enumerable set L does not exceed $\log_2 n$ .

His means that for arbitrary r.e. set L and arbitrary natural number n there exists a program whose length does not exceed $\log_2 n$ bits, and this program can output a word $s_1 s_2 s_3 \ldots s_n$ , of length n bits where

$$s_i = \begin{cases} 1, \text{ if } i \in L, \\ 0, \text{ if } i \in L. \end{cases}$$

**Theorem.** Under assumption of Artin's Conjecture there exists an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ in a 2-letter alphabet and an infinite sequence of positive integers $z(1), z(2), z(3), \ldots$ such that for arbitrary $j$:

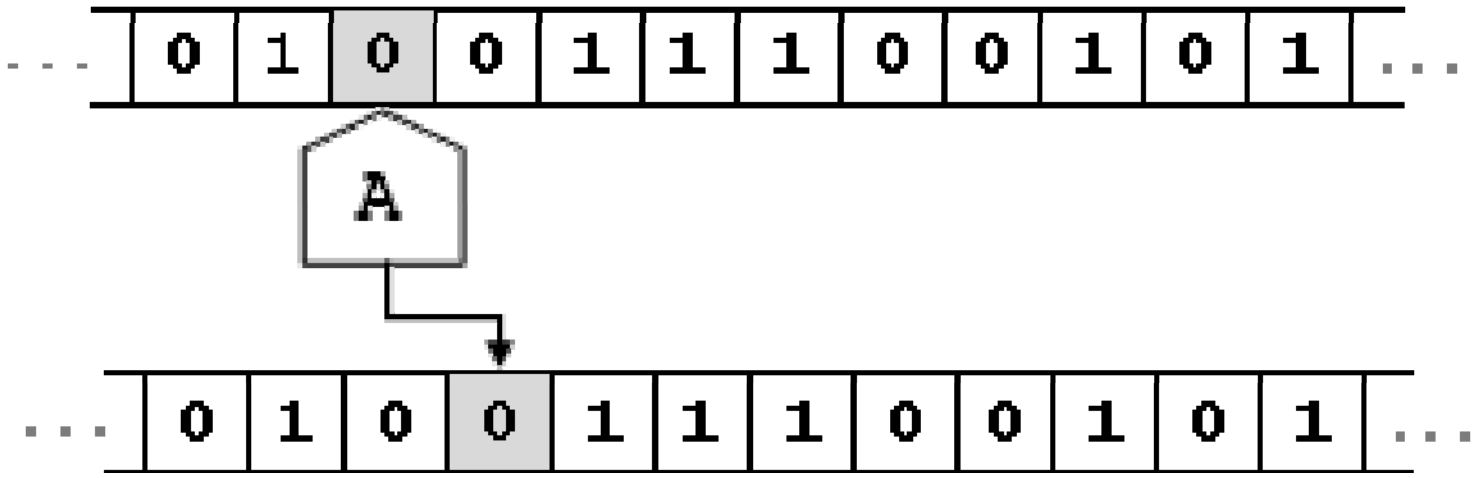
1. there is a probabilistic reversible automaton with $z(j)$ states recognizing $L_j$ with the probability $\frac{19}{36}$,

2. any deterministic finite automaton recognizing $L_j$ has at least $(2^{1/4})^{z(j)} = = (1.1892071115\ldots)^{z(j)}$ states,

   **Theorem.** Even without assumption of Artin's Conjecture there exists an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ in a 2-letter alphabet and an infinite sequence of positive integers $z(1), z(2), z(3), \ldots$ such that for arbitrary $j$:

   (a) there is a probabilistic reversible automaton with $z(j)$ states recognizing $L_j$ with the probability $\frac{68}{135}$,

   (b) any deterministic finite automaton recognizing $L_j$ has at least $(7^{\frac{1}{14}})^{z(j)} = = (1.1149116725\ldots)^{z(j)}$ states,




**Definition.** We say that a finite automaton recognizes the language L with a nonconstructivity g(n) , if for arbitrary natural number n there is a word y of length not exceeding g(n) such that for all input words x of length not exceeding n the automaton working on the pair (x,y) produces a correct result on x being or not being in L .

**Theorem.** There exists a nonrecursive language L ,which can be recognized by a finite automaton with a nonconstructivity n .

**Proof.** An infinite nonrecursive sequence of bits

$$a_0\, a_1\, a_2\, a_3\, a_4 \ldots$$

The language is "input word is an initial fragment of the sequence"

**Theorem.** There exists a nonrecursive language L ,which can be recognized by a finite automaton with a nonconstructivity n but cannot be recognized with a nonconstructivity n – h(n) , where h(n) grows to infinity.

**Proof.** The sequence is a Martin-Löf sequence.

Consider a language L consisting of words

$$0^m 1 0^m 1 0^m 1 \ldots 1 0^m$$

with m arrays of zeros

**Theorem.** The language L is not regular but it can be recognized by a finite automaton with nonconstructivity $n^{1/2}$.

R. Karp and R. Lipton have introduced in a notion *Turing machine that takes advice* which is in fact a usage of a nonconstructive help from outside in a process of computation. Later C. Damm and M. Holzer have adapted this notion of advice for finite automata. A slightly different definition was used by Freivalds. It turns out that for some languages this nonconstructive help can bring zero information about the input word's being or not being in the language considered. Is it equivalent to the automaton's being a probabilistic automaton? We will see that it is not.

What is a random sequence of bits? Martin-Löf's original definition of a random sequence was in terms of constructive null covers; he defined a sequence to be random if it is not contained in any such cover. Leonid Levin and Claus-Peter Schnorr proved a characterization in terms of Kolmogorov complexity: a sequence is random if there is a uniform bound on the compressibility of its initial segments. An infinite sequence S is Martin-Löf random if and only if there is a constant $c$ such that all of $S$'s finite prefixes are $c$-incompressible. Schnorr gave a third equivalent definition in terms of martingales (a type of betting strategy).

**Theorem.** (1)The language

$$L = \{x2x \mid x \in \{0,1\}^*\}$$

cannot be recognized with a bounded error by a probabilistic 2-way finite automaton.
(2) The language $L$ can be recognized by a deterministic non-writing 2-tape finite automaton one tape of which contains the input word, and the other tape contains an infinite Martin-Löf random sequence, the automaton is 2-way on every tape, and it stops producing a the correct result in a finite number of steps for arbitrary input word.

**Proof.** (2) Let the input word be $x(r)2z(s)$ where $r$ and $s$ are the lengths of the corresponding words. At first, the 2-tape automaton finds a fragment $01111\cdots$ which has the length at least $r$ and uses it as a counter to test whether $r = s$. Then the automaton searches for another help-word. If the help-word turns out to be $y$ then the automaton tests whether $x(r) = y$ and whether $z(s) = y$.

**Definition.** A 2-infinite sequence of bits is a sequence $\{a_i\}$ where $i \in (-\infty, \infty)$ and all $a_i \in \{0, 1\}$.

**Definition.** We say that a 2-infinite sequence of bits is Martin-Löf random if for arbitrary $i \in (-\infty, \infty)$ the sequence $\{b_n\}$ where $b_n = a_{i+n}$ for all $i \in \mathbb{N}$ is Martin-Löf random, and the sequence $\{c_n\}$ where $c_n = a_{i-n}$ for all $i \in \mathbb{N}$ is Martin-Löf random.

**Definition.** A deterministic finite automaton with intuition is a deterministic non-writing 2-tape finite automaton one tape of which contains the input word, and the other tape contains a 2-infinite Martin-Löf random sequence, the automaton is 2-way on every tape, and it stops producing a the correct result in a finite number of steps for arbitrary input word. Additionally it is demanded that the head of the automaton never goes beyond the markers showing the beginning and the end of the input word.

**Theorem.** The unary language PERFECT SQUARES= $= \{1^n \mid (\exists m)(n = m^2)\}$ can be recognized by a deterministic finite automaton with intuition.

**Proof.** It is well-known that

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2.$$

The deterministic automaton with intuition searches for a help-word (being a fragment of the given 2-infinite Martin-Löf sequence) of a help-word

$$00101110111110 \cdots 01^{2n-1}00.$$

At first, the input word is used as a counter to test whether each substring of 1's is exactly 2 symbols longer than the preceding one. Then the help-word is used to test whether the length of the input word coinsides with the number of 1's in the help-word.

**Theorem.** The relation

$$\text{SQUARE ROOTS} = \{(1^n, 1^m) \mid (n = m^2)\}$$

can be computed by a deterministic finite-state transducer with intuition.

**Theorem.** The unary language PERFECT CUBES=
$= \{1^n \mid (\exists m)(n = m^3)\}$ can be recognized by a deterministic finite automaton with intuition.

**Proof.** In a similar manner the formula

$$1 + 3(n-1) + 3(n-1)^2 = n^3$$

suggests a help-word

$$000[1]00[101110111]00[10111111011111111111]00\cdots00[101^{n-1}01^{(n-1)^2}]000$$

where symbols [,] are invisible. At first, the input word is used as a counter to test whether the help-word is correct but not whether its length is sufficient. Then the help-word is used to test whether the length of the input word coinsides with the number of 1's in the help-word.

**Theorem.** The relation

$$\text{CUBE ROOTS} = \{(1^n, 1^m) \mid (n = m^3)\}$$

can be computed by a deterministic finite-state transducer with intuition.

**Theorem.** The unary language PRIMES= $= \{1^n \mid n \text{ is prime}\}$ can be recognized by a deterministic finite automaton with intuition.

**Theorem.** The relation

$$FACTORISATION =$$

$$= \{(1^n, 1^m) \mid (m \text{ divides } n \wedge m \neq 1) \vee (m = 1 \text{ if } m \text{ is prime})\}$$

can be computed by a deterministic finite-state transducer with intuition.

We define a language UNARY 3-SAT as follows. The term $term_1 = x_k$ is coded as $[term_1]$ being $21^k$, the term $term_2 = \neg x_k$ is coded as $[term_2]$ being $31^k$, the subformula $f$ being $(term_1 \vee term_2 \vee term_3)$ is coded as $[f]$ being $[term_1] \vee [term_2] \vee [term_3]$. The $CNF$ being $f_1 \wedge f_2 \wedge \cdots \wedge f_m$ is coded as $[f_1] \wedge [f_2] \wedge \cdots \wedge [f_m]$.

**Theorem.** The language UNARY 3-SAT can be enumerated by a deterministic finite automaton with intuition.

**Theorem.** Every $L \in NP$ is reducible by a deterministic $log$-space bounded Turing machine to a language $L'$ such that $L'$ is enumerable by a deterministic finite automaton with intuition.

**Proof.** 3-SAT is $NP$-complete. Hence $L$ is reducible by a deterministic $log$-space bounded Turing machine to 3-SAT. The language 3-SAT is reducible by a deterministic $log$-space bounded Turing machine to $UNARY 3-SAT$. The language UNARY 3-SAT is enumerable by a deterministic finite automaton $B$ with intuition.

We define a relation UNARY 3-SATISFIABILITY as follows. The term $term_1 = x_k$ is coded as $[term_1]$ being $21^k$, the term $term_2 = \neg x_k$ is coded as $[term_2]$ being $31^k$, the subformula $f$ being $(term_1 \vee term_2 \vee term_3)$ is coded as $[f]$ being $[term_1] \vee [term_2] \vee [term_3]$. The $CNF$ being $f_1 \wedge f_2 \wedge \cdots \wedge f_m$ is coded as $[f_1] \wedge [f_2] \wedge \cdots \wedge [f_m]$. The string of the values $x_1 = a_1, x_2 = a_2, \cdots, x_n = a_n$ is coded as $a_1 a_2 \cdots a_n$. The relation UNARY 3-SATISFIABILITY consists of all the pairs $(CNF, a_1 a_2 \cdots a_n)$ such that the given CNF with these values of the arguments takes the value TRUE.

**Theorem 1.** *The relation UNARY 3-SATISFIABILITY can be computed by a deterministic finite-state transducer with intuition.*

**Theorem.** If a language $L$ is recognizable by a nondeterministic finite automaton with intuition then
$L \in NP \cap co - NP$.

**Theorem.** Every language enumerable by a deterministic finite automaton with intuition is also recognizable by a nondeterministic finite automaton with intuition if and only if $P = NP$.

**Theorem.** If a language $L$ is recognizable by a nondeterministic finite automaton with intuition then $L$ is also recognizable by a deterministic finite automaton with intuition.

**Theorem.** Every language enumerable by a deterministic finite automaton with intuition is also recognizable by a deterministic finite automaton with intuition if and only if $P = NP$.

# Thank you