Proof and Computation in Coq

Maxime Dénès, Benjamin Grégoire, Chantal Keller,
Pierre-Yves Strub, Laurent Théry

Motivations

Functional programming language

```
Compute prime 31.
   = true
```

Theorem prover

```
Check Euclid_dvdX.
   forall m n p :  nat,
   prime p -> (p %| m ^ n) = (p %| m) && (0 < n)
```

Motivations

```
Theorem Name :  Statement.
Proof.
```

```
Qed.
```

Outline

Toy Example

Erdös Discrepancy

Ternary Goldbach Conjecture

Toy Example

Magic Trick

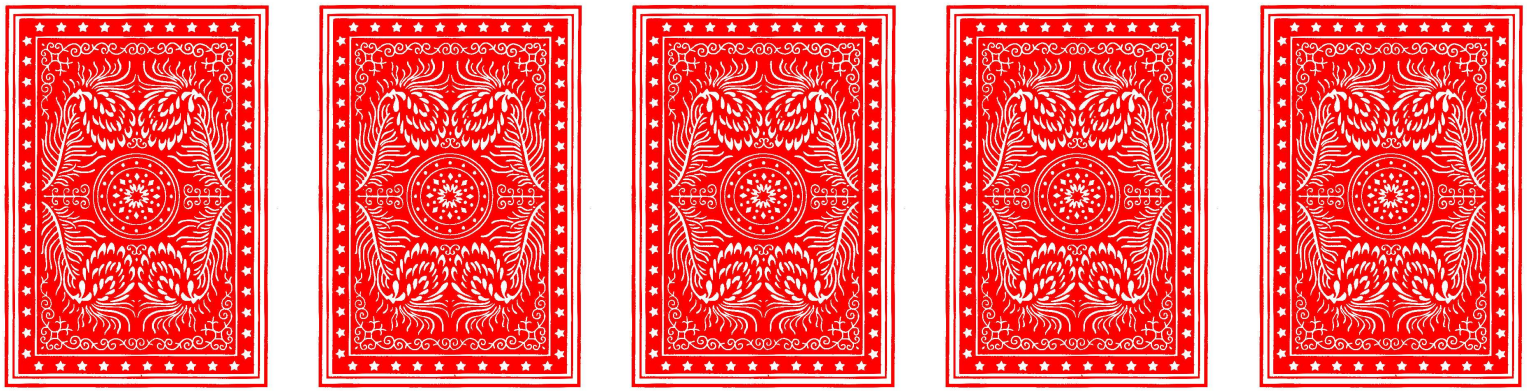
32-card deck

card guessing trick

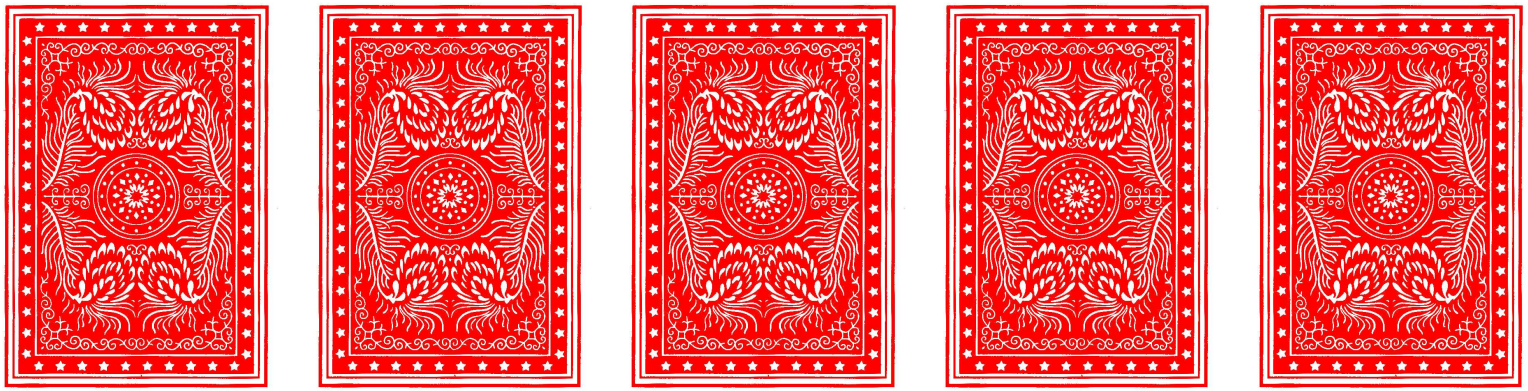
Steps

1. Shuffle the deck by multiple cuts 
2. Pick five people in the audience
3. Give a card to each participant
4. Test mind-reading
5. Guess cards

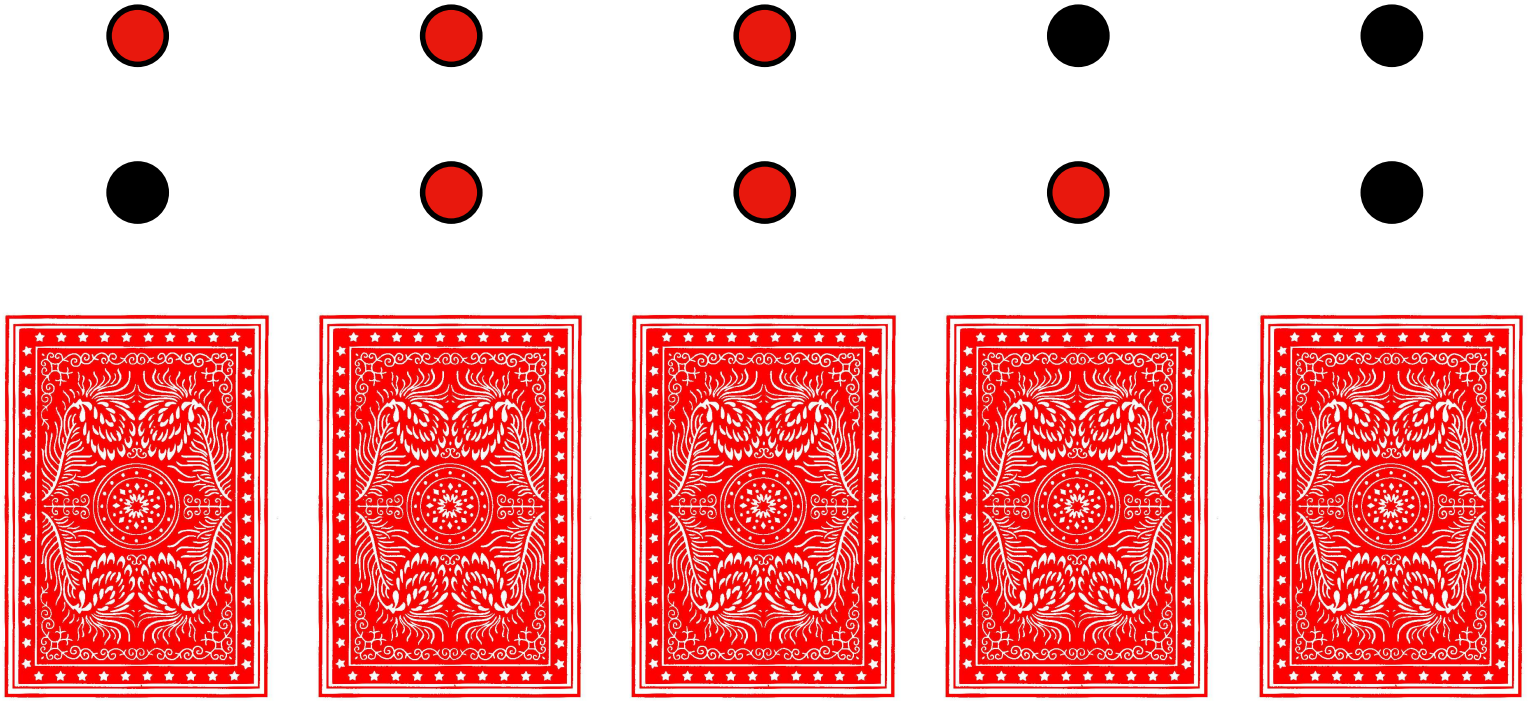
Pick



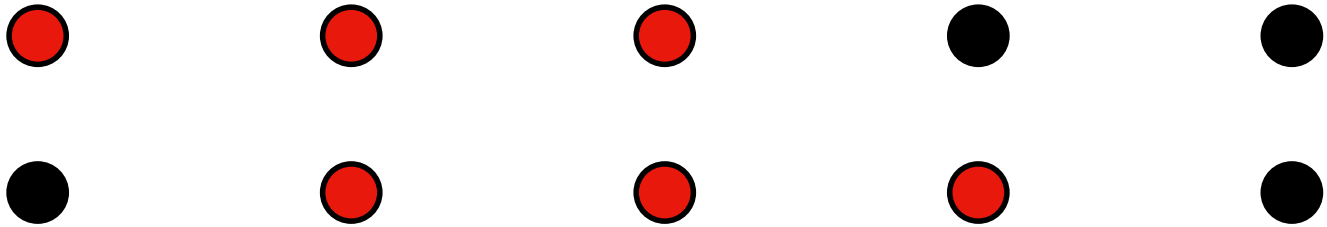
Test

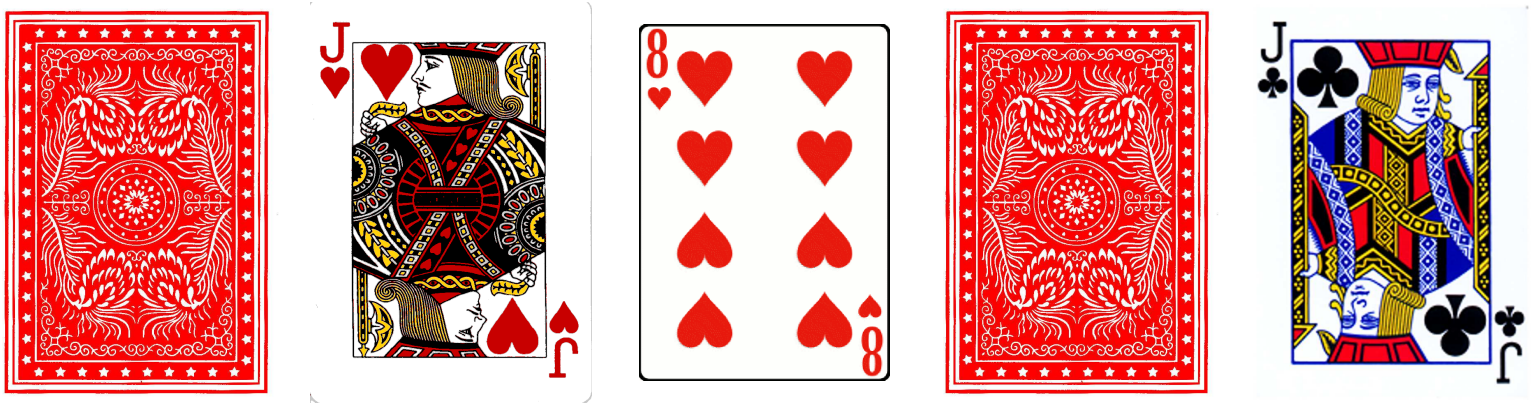


Test



Guess

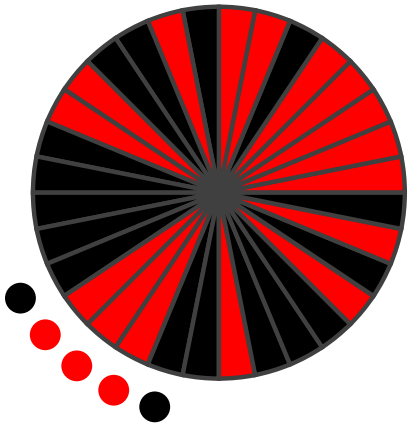




Magic Trick

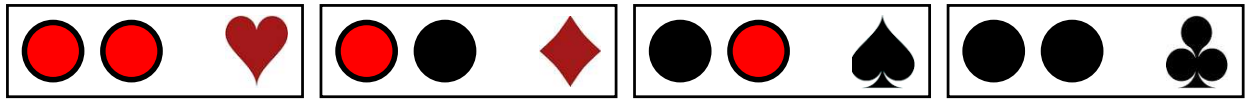
Test : $2^5 = 32$

Shuffle

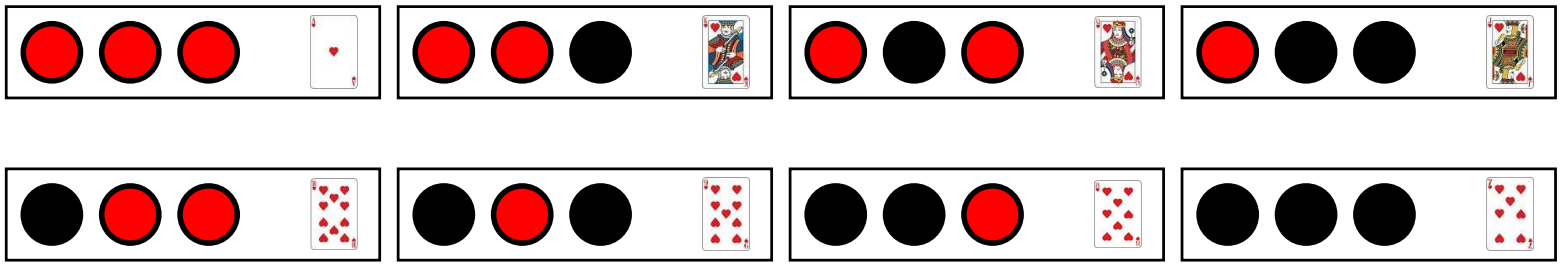


Encoding

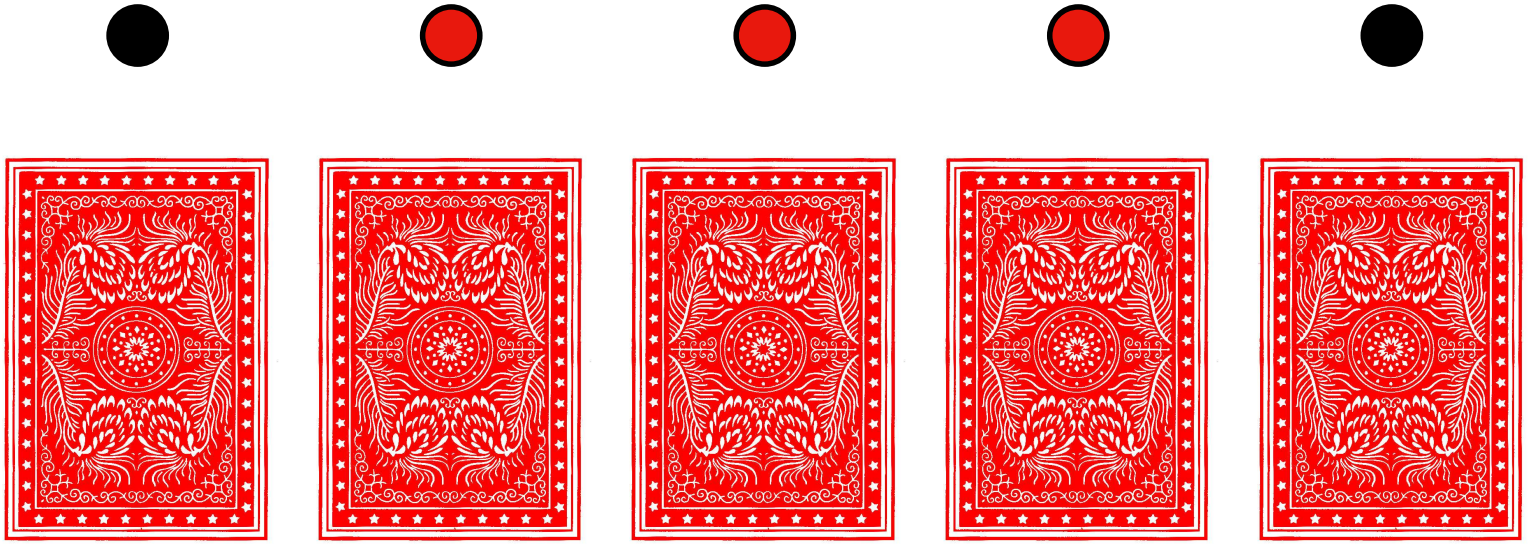
Suits



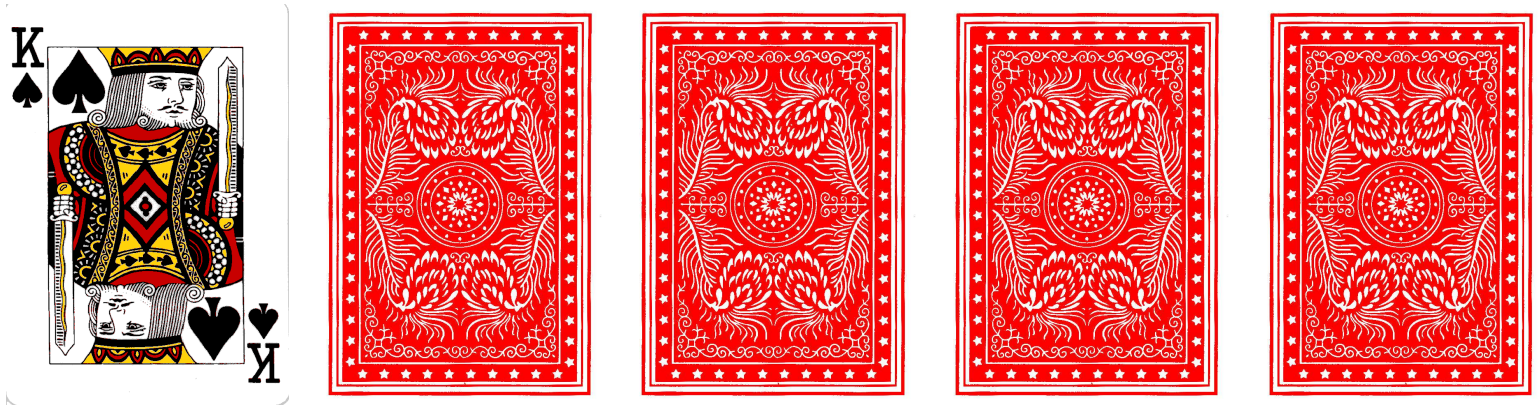
Ranks



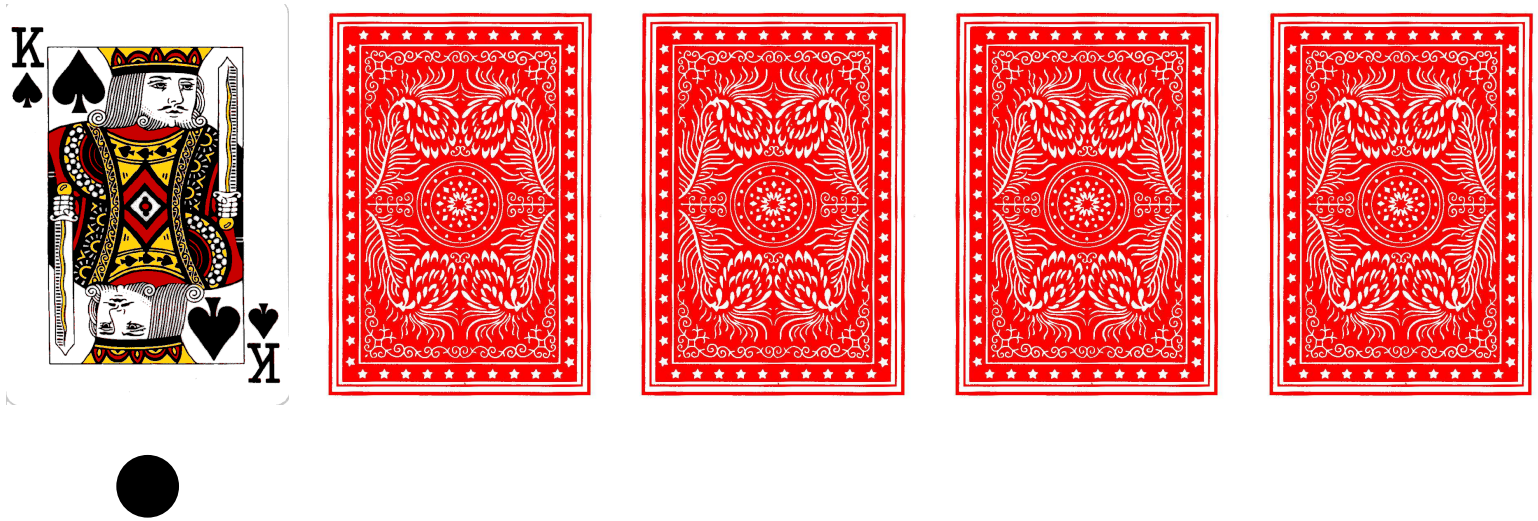
Decoding



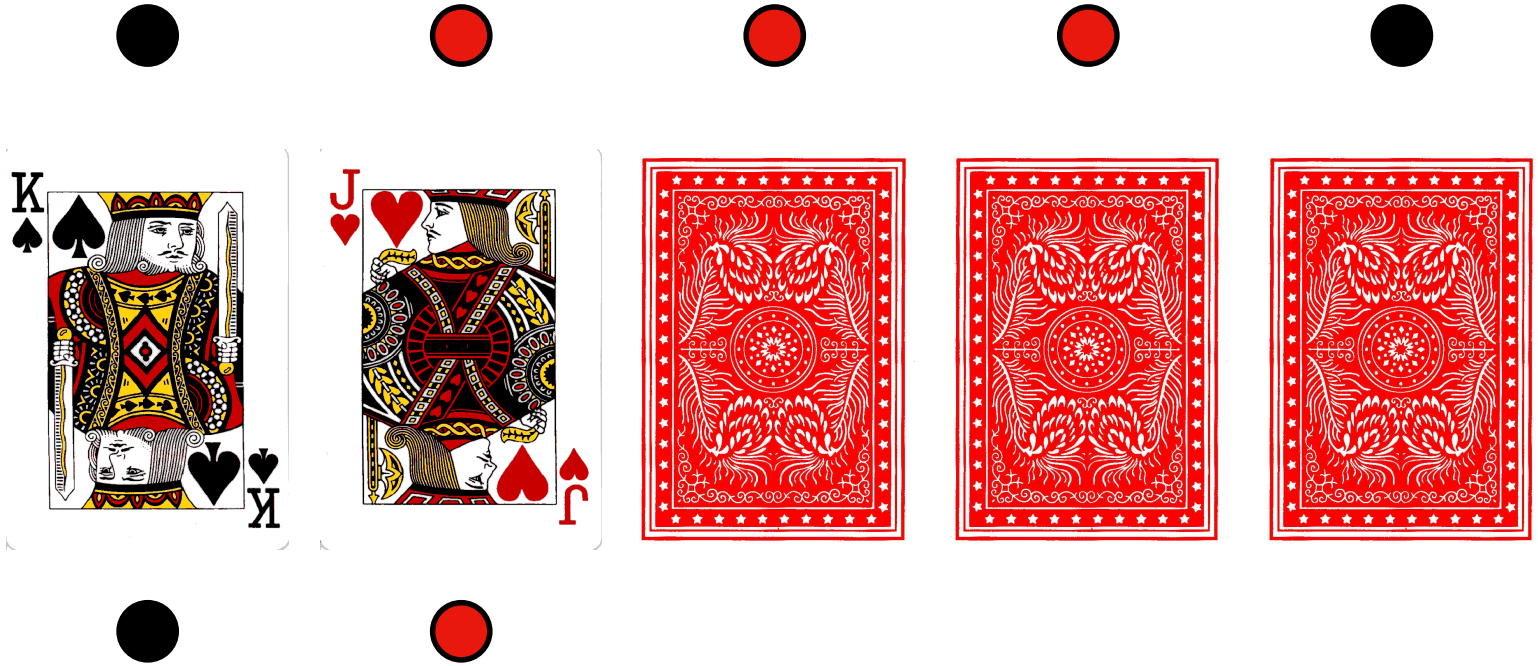
Decoding



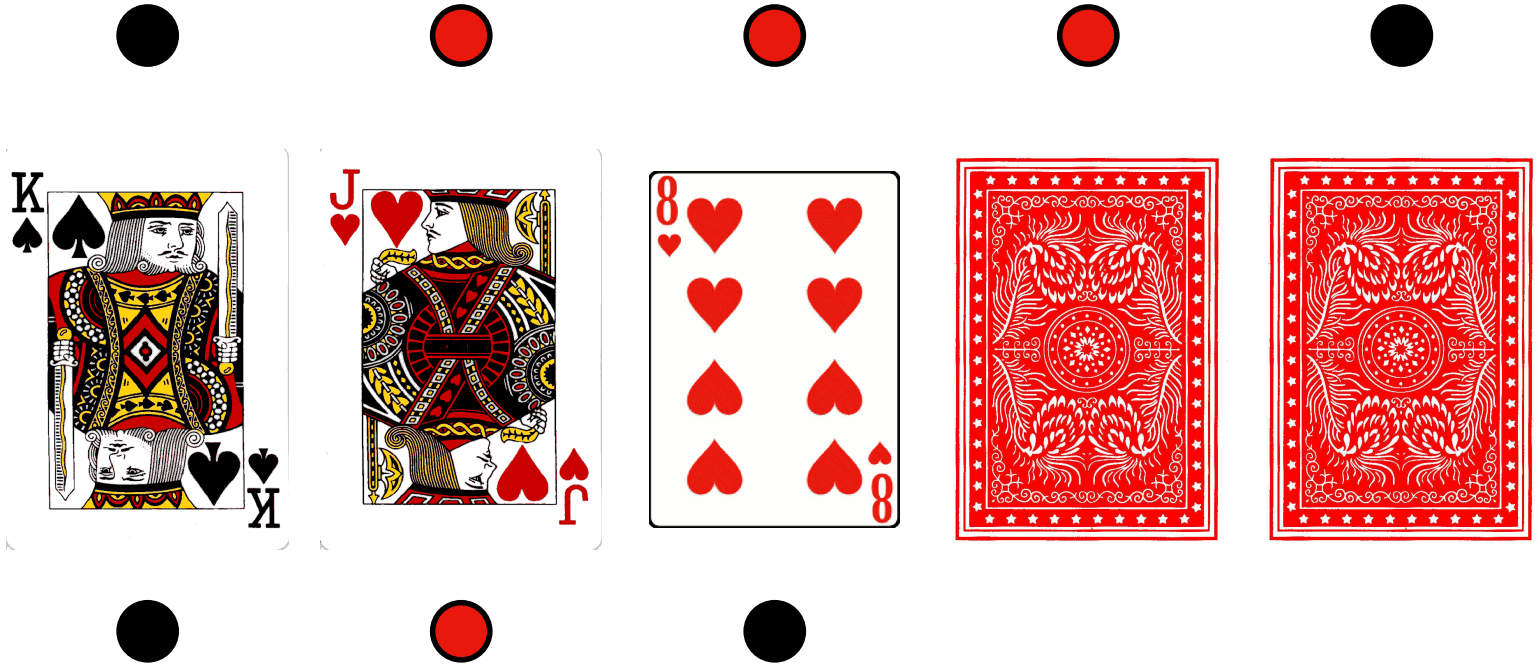
Decoding



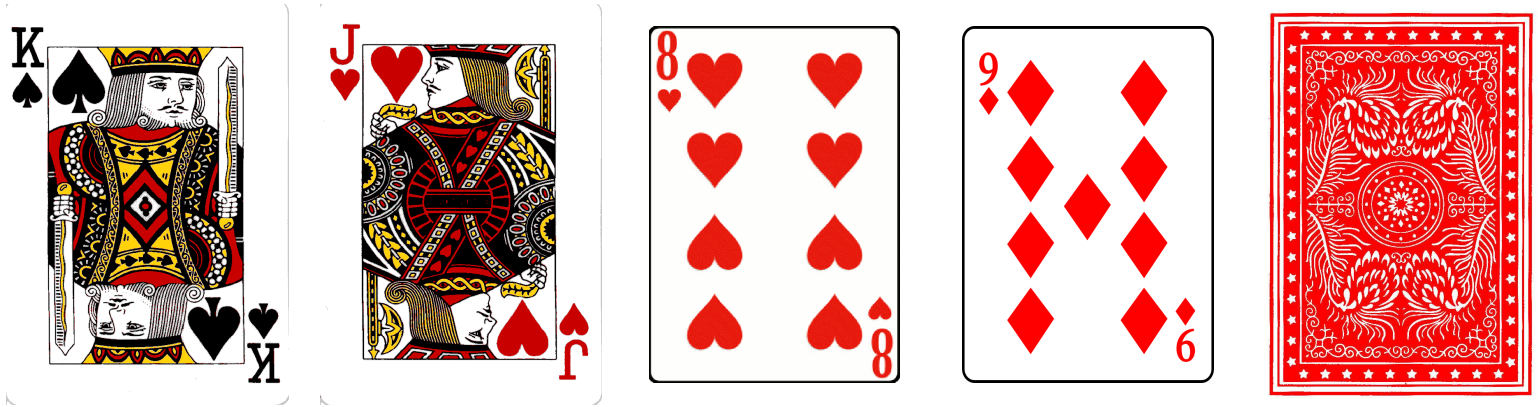
Decoding



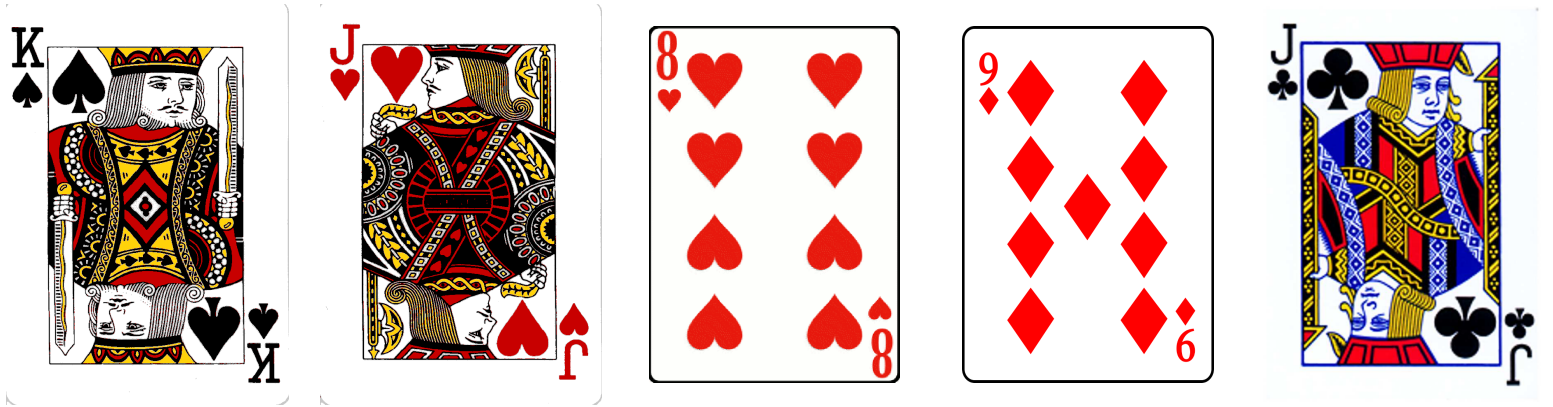
Decoding



Decoding

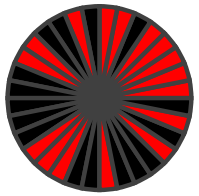


Decoding



Direct Verification

Recursive definition



$$
\begin{aligned}
s_0 &= [0,0,0,0,0] \\
s_1 &= [0,0,0,0,1] \\
&\dots \\
s_{n+1} &= [s_{n,2}, s_{n,3}, s_{n,4}, s_{n,5}, s_{n,1} + s_{n,3}]
\end{aligned}
$$

Property

$$\texttt{uniq}\,[\,s_i \mid i < 32\,]$$

Indirect Verification

$$
\begin{aligned}
s_n &= [a_{n-5}, a_{n-4}, a_{n-3}, a_{n-2}, a_{n-1}] \\
\\
a_{-n} &= 0 && 0 < n \\
a_0 &= 1 \\
a_n &= a_{n-5} + a_{n-3} && 0 < n
\end{aligned}
$$

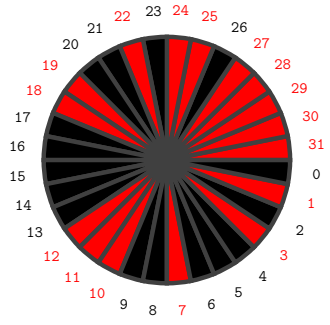
$$P = X^5 + X^2 + 1 \quad (\mathbb{F}_2[X])$$

$$1/P = \sum_{i \geq 0} a_i X^i$$

Indirect Verification

$$(1+X^{31})/P = X^{26}+X^{23}+X^{21}+X^{20}+X^{17}+X^{16}+X^{15}+X^{14}+X^{13}$$

$$+X^{9}+X^{8}+X^{6}+X^{5}+X^{4}+X^{2}+1$$

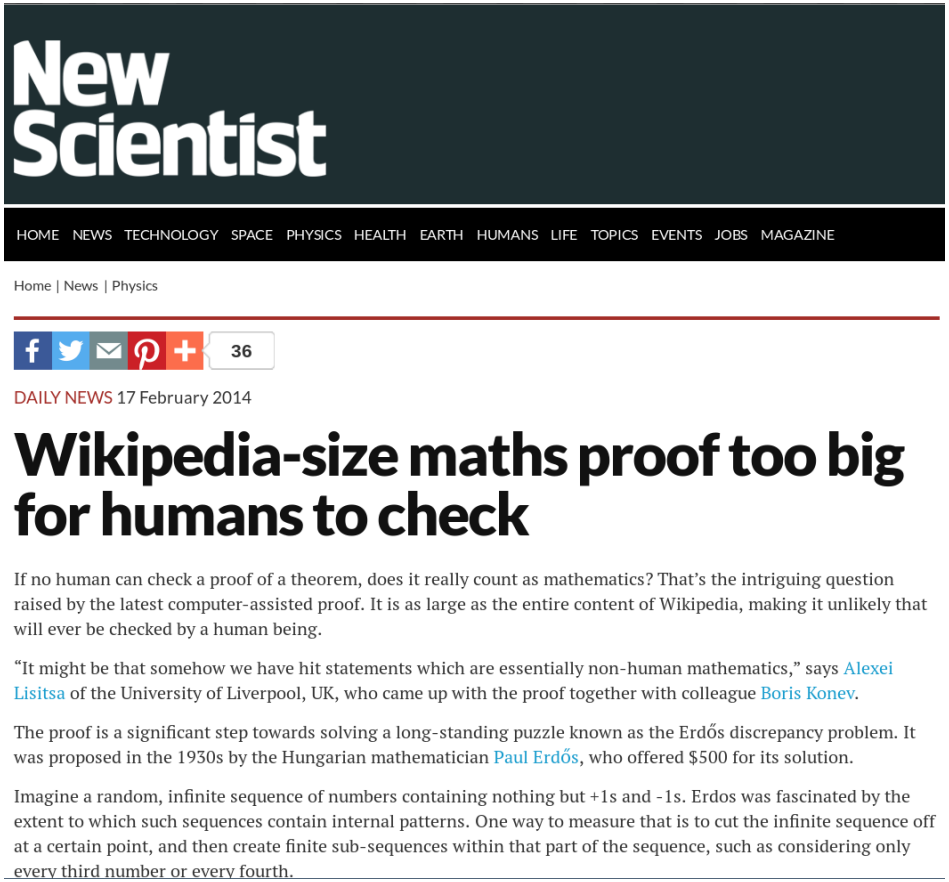


Property

irreducible P

Erdös Discrepancy

Erdös Discrepancy



New Scientist

HOME NEWS TECHNOLOGY SPACE PHYSICS HEALTH EARTH HUMANS LIFE TOPICS EVENTS JOBS MAGAZINE

Home | News | Physics

36

DAILY NEWS 17 February 2014

Wikipedia-size maths proof too big for humans to check

If no human can check a proof of a theorem, does it really count as mathematics? That's the intriguing question raised by the latest computer-assisted proof. It is as large as the entire content of Wikipedia, making it unlikely that will ever be checked by a human being.

"It might be that somehow we have hit statements which are essentially non-human mathematics," says Alexei Lisitsa of the University of Liverpool, UK, who came up with the proof together with colleague Boris Konev.

The proof is a significant step towards solving a long-standing puzzle known as the Erdős discrepancy problem. It was proposed in the 1930s by the Hungarian mathematician Paul Erdős, who offered $500 for its solution.

Imagine a random, infinite sequence of numbers containing nothing but +1s and -1s. Erdos was fascinated by the extent to which such sequences contain internal patterns. One way to measure that is to cut the infinite sequence off at a certain point, and then create finite sub-sequences within that part of the sequence, such as considering only every third number or every fourth.

Sat Solver

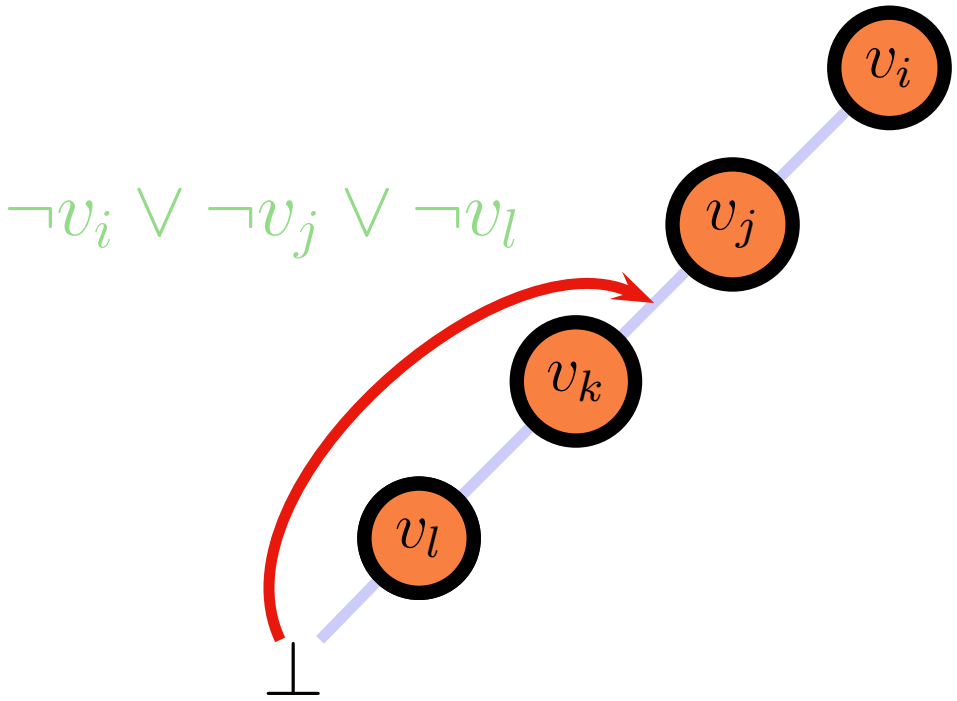
Problem :

Is it possible to put n pigeons in m holes in such a way that each hole contains at most one pigeon?

Encoding :

$v_{i,p}$	pigeon i is in hole p
$v_{i,1} \vee v_{i,2} \vee \cdots \vee v_{i,m}$	n clauses
$\neg v_{i,p} \vee \neg v_{i,q}$	$n(m^2 - m)/2$ clauses

Sat Solver



VS + UP + BJ + L

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

a

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

$a \quad b$

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

$a \quad b \quad c$

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

$a \quad b \quad c$

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

$a \quad b \quad c \quad d \quad e$

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

$a \quad b \quad c \quad d \quad e$

$$\neg a \vee \neg c$$

Sat Solver

$$b \vee d$$

$$\neg a \vee \neg c \vee d$$

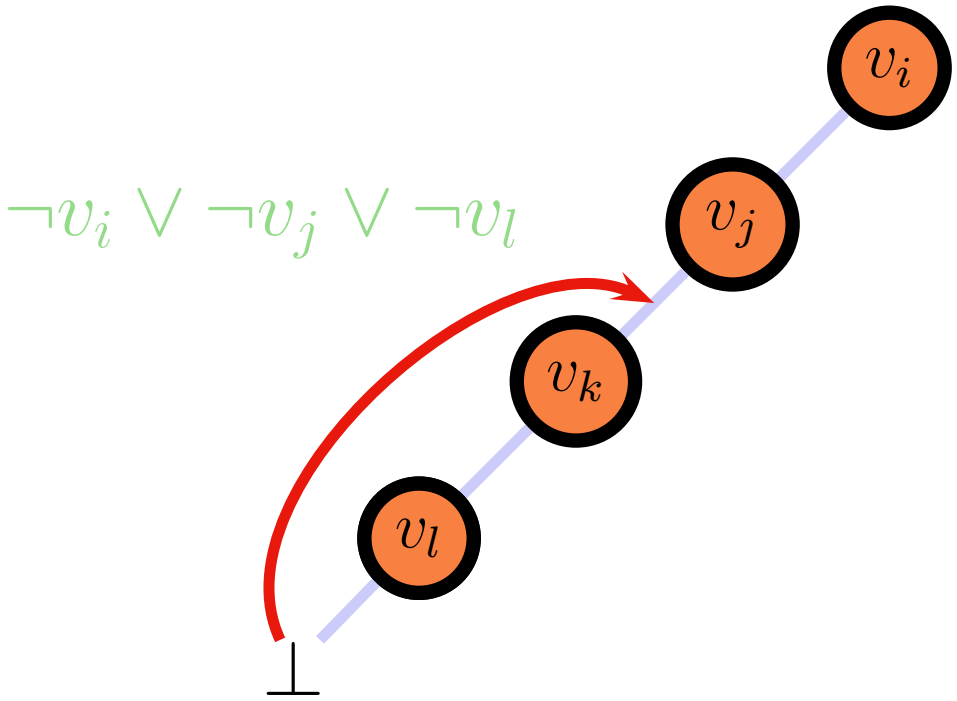
$$\neg c \vee e$$

$$\neg a \vee \neg d \vee \neg e$$

$$\neg a \vee \neg c$$

a

Sat Solver



VS + UP + BJ + L

Sat Proof

Resolution

$$\frac{v \vee C \qquad \neg v \vee C'}{C \vee C'}$$

Example

$$\dfrac{\dfrac{\neg a \vee \neg c \vee d \qquad \dfrac{\neg c \vee e \qquad \neg a \vee \neg d \vee \neg e}{\neg a \vee \neg c \vee \neg d}}{\neg a \vee \neg c}}{}$$

Sat Proof

```
 3  2
 1  0
14 13
21 23
20 19 79
12 11
18 16 15 81 78 80
 8  7
 5  6  9 83  4 76
77 84
80 85
81 85
82 85
78 86
17 88 89
15 87 89
18 91 89 90
```

Erdös Discrepancy

For every infinite sequence (x_n) $(x_i \in \{-1, 1\})$, for every C, there exists (k, d) such that

$$|\sum_{i=1}^{d} x_{ki}| > C$$

Erdös Discrepancy for $C = 1$

$$
\begin{array}{ll}
\mathrm{k}=1 & (+1,?,?,?,?,?,?,?,?,?,?,?) \\
\mathrm{k}=2 & (+1,-1,?,?,?,?,?,?,?,?,?,?) \\
\mathrm{k}=1 & (+1,-1,?,+1,?,?,?,?,?,?,?,?) \\
\mathrm{k}=3 & (+1,-1,-1,+1,?,?,?,?,?,?,?,?) \\
\mathrm{k}=1 & (+1,-1,-1,+1,?,+1,?,?,?,?,?,?) \\
\mathrm{k}=4 & (+1,-1,-1,+1,-1,+1,?,?,?,?,?,?) \\
\mathrm{k}=1 & (+1,-1,-1,+1,-1,+1,?,-1,?,?,?,?) \\
\mathrm{k}=5 & (+1,-1,-1,+1,-1,+1,+1,-1,?,?,?,?) \\
\mathrm{k}=1 & (+1,-1,-1,+1,-1,+1,+1,-1,?,+1,?,?) \\
\mathrm{k}=6 & (+1,-1,-1,+1,-1,+1,+1,-1,-1,+1,?,?) \\
\mathrm{k}=1 & (+1,-1,-1,+1,-1,+1,+1,-1,-1,+1,?,-1) \\
\mathrm{k}=3 & (+1,-1,-1,+1,-1,+1,+1,-1,-1,+1,+1,-1)
\end{array}
$$

Erdös Discrepancy for $C = 2$

Arbitrary sequence (x_n) of 1161 elements

Express the constraints using clauses

Boolean encoding

$x_i = \top$ $\quad$ x_i has value $+1$

$x_i = \bot$ $\quad$ x_i has value -1

Every partial sum of odd length can be encoded by a boolean.

Erdös Discrepancy for $C = 2$

$y_{i_3} = x_{i_1} + x_{i_2} + x_{i_3}$ is between -2 and 2

$$x_{i_1} \vee x_{i_2} \vee x_{i_3}$$
$$\neg x_{i_1} \vee \neg x_{i_2} \vee \neg x_{i_3}$$

$$x_{i_1} \vee x_{i_2} \vee \neg y_{i_3}$$
$$\neg x_{i_1} \vee \neg x_{i_2} \vee y_{i_3}$$

$$x_{i_1} \vee x_{i_3} \vee \neg y_{i_3}$$
$$\neg x_{i_1} \vee \neg x_{i_3} \vee y_{i_3}$$

$$x_{i_2} \vee x_{i_3} \vee \neg y_{i_3}$$
$$\neg x_{i_2} \vee \neg x_{i_3} \vee y_{i_3}$$

Erdös Discrepancy for $C = 2$

Encoding :

4206 variables, 25142 clauses

Glucose takes 30 min to prove unsatisfiability and the proof is 2 Gb.

Erdös Discrepancy in Coq

```
Lemma Erdos :
   forall x_ :  nat -> bool, exists k, exists d,
    `| \ sum_(1 <= i < d) [x_ (k * i)] | > 2.
```

Checking Time: 5 min

Process Size: 2 Gb

Compiled File: 1 Mb

Erdös Discrepancy in Coq

What has been done?

Verified proof-cheker for resolution

Verified translatation

How has it been done?

Preprocessing

Use some “impure features”

Ternary Goldbach Conjecture

THE TERNARY GOLDBACH CONJECTURE IS TRUE

H. A. HELFGOTT

ABSTRACT. The ternary Goldbach conjecture, or three-primes problem, asserts that every odd integer n greater than 5 is the sum of three primes. The present paper proves this conjecture.

Both the ternary Goldbach conjecture and the binary, or strong, Goldbach conjecture had their origin in an exchange of letters between Euler and Goldbach in 1742. We will follow an approach based on the circle method, the large sieve and exponential sums. Some ideas coming from Hardy, Littlewood and Vinogradov are reinterpreted from a modern perspective. While all work here has to be explicit, the focus is on qualitative gains.

The improved estimates on exponential sums are proven in the author's papers on major and minor arcs for Goldbach's problem. One of the highlights of the present paper is an optimized large sieve for primes. Its ideas get reapplied to the circle method to give an improved estimate for the minor-arc integral.

Ternary Goldbach Conjecture

"Every odd number greater than 5 can be expressed as the sum of three primes"

$$
\begin{aligned}
7 &= 5+2+2 \\
9 &= 3+3+3 \\
11 &= 3+3+5 \\
13 &= 3+5+5 \\
15 &= 5+5+5 \\
17 &= 7+5+5 \\
&\dots
\end{aligned}
$$

Ternary Goldbach Conjecture

NUMERICAL VERIFICATION OF THE TERNARY GOLDBACH CONJECTURE UP TO $8.875 \cdot 10^{30}$

H. A. HELFGOTT AND DAVID J. PLATT

ABSTRACT. We describe a computation that confirms the ternary Goldbach Conjecture up to $8,875,694,145,621,773,516,800,000,000,000$ ($> 8.875\text{-}10^{30}$).

Proving Prime

Particular Prime

Many Primes

Proving a Prime

Suppose we want to prove that
$n = 10000000000000066000000000000001$
is prime.

Primality Certificate

Suppose we want to prove that
$n = 10000000000000666000000000000001$
is prime.

Find an a of "big" order m

$$a^m = 1 \bmod n$$

$$\forall p, \texttt{prime}\, p \ \wedge \ p \,|\, m \ \Rightarrow \ a^{m/p} \neq 1 \bmod n$$

If $m \geq \sqrt{n}$, n is prime.

Primality Certificate

Suppose we want to prove that
$n = 10000000000000666000000000000001$
is prime.

Find an a of "big" order m

$$a^m = 1 \bmod n$$

$$\forall p, \texttt{prime}\, p \ \wedge \ p \,|\, m \ \Rightarrow \ gcd(a^{m/p} - 1, n) = 1$$

If $m \geq \sqrt{n}$, n is prime.

Primality Certificate

Suppose we want to prove that
$n = 10000000000000666000000000000001$
is prime.

Find an a of "big" order m

$$a^m = 1 \bmod n$$

$$\forall p, \texttt{prime}\, p \ \wedge \ p \,|\, m \ \Rightarrow \ gcd(a^{m/p} - 1, n) = 1$$

If $m \geq \sqrt{n}$, n is prime.

Problem: partial factorization of $n - 1$

Primality Certificate

Primality Certificate

```
Lemma primo0:
  prime  49409758827788562229->
  prime  1000000000000066600000000000001.
Proof.
intro H.
apply (Pocklington_refl
     (Ell_certif
      1000000000000066600000000000001
      202389168400
      ([49409758827788562229,1]
      100
      0
      20
      100)
      [Proof_certif _ H]).
exact (refl_equal true).
Qed.
```

Primality Certificate

```
Lemma primo0:
  prime  49409758827788562229->
  prime  10000000000000666000000000000001.
Proof.
...
Qed.

Lemma primo1:
  prime 44630042033-> prime 49409758827788562229.
Proof.
...
Qed.

Lemma primo2 : prime 44630042033.
Proof.
...
Qed.


Lemma  primo: prime 10000000000000666000000000000001.
Proof.
exact (primo0 (primo1 primo2)).
Qed.
```

Proving Many Primes

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Erathosthenes

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

2 3 5 7 9 11 13 15 17 19

2 3 5 7 11 13 17 19

2 3 5 7 11 13 17 19

Erathosthenes

What is a reasonable algorithm?

Compressing

$$C = [1;\ 7;\ 11;\ 13;\ 17;\ 19;\ 23;\ 27;\ 29]$$
$$n = 30(p/9) + C[p \mod 9]$$

Jumping

$$J[7] = [(29,\ 4,\ 0),\ (19,\ 6,\ 1),\ (23,\ 6,\ 1),\ (27,\ 2,\ 0),\ (1,\ 2,\ 1),$$
$$(17,\ 4,\ 1),\ (7,\ 2,\ 1),\ (11,\ 2,\ 1),\ (13,\ 2,\ 1)]$$

Ternary Goldbach Conjecture

Idea:

$[p_1, p_2, \ldots, p_n]$ such that $p_{i+1} \leq p_i + A$

If BGC holds under A, TGC holds under $p_n + A$.

$n - p_i = 2q = p_j + p_k$

Computation:

$A = 4\,10^{18}$

Candidate:

Fast: Proth numbers $k\,2^h + 1$
Slow: Probable prime.

Ternary Goldbach Conjecture

```
Lemma Goldbach_main :
  (forall n :  Z,
  Zeven n ->
  4 ≤ n ≤ 4 * 10 ^ 18 ->
  exists p1 p2 :  Z, prime p1 ∧ prime p2 ∧ n = p1 + p2) ->
  forall n :  Z,
  Zodd n ->
  7 ≤ n ≤ 22240000000000 * 2 ^ 52 ->
  exists p1 p2 p3 :  Z,
  prime p1 ∧ prime p2 ∧ prime p3 ∧ n = p1 + p2 + p3
```

Ternary Goldbach Conjecture

What made it possible?

556 files

Frame: $4\,10^{10}\,2^{52}$

Size : 86 Gb

227 Non Proth

RunTime: 5 days with 24 cores

Binary Goldbach Conjecture

EMPIRICAL VERIFICATION OF THE EVEN GOLDBACH CONJECTURE, AND COMPUTATION OF PRIME GAPS, UP TO $4 \cdot 10^{18}$

TOMÁS OLIVEIRA E SILVA, SIEGFRIED HERZOG, AND SILVIO PARDI

ABSTRACT. This paper describes how the even Goldbach conjecture was confirmed to be true for all even numbers not larger than $4 \cdot 10^{18}$. Using a result of Ramaré and Saouter, it follows that the odd Goldbach conjecture is true up to $8.37 \cdot 10^{26}$. The empirical data collected during this extensive verification effort, viz., counts and first occurrences of so-called minimal Goldbach partitions with a given smallest prime and of gaps between consecutive primes with a given even gap, are used to test several conjectured formulas related to prime numbers. In particular, the counts of minimal Goldbach partitions and of prime gaps are in excellent accord with the predictions made using the prime k-tuple conjecture of Hardy and Littlewood (with an error that appears to be $O(\sqrt{t \log\log t})$, where t is the true value of the quantity being estimated). Prime gap moments also show excellent agreement with a generalization of a conjecture made in 1982 by Heath-Brown.

Binary Goldbach Conjecture

$n = p_i + p_j$

$n = 33255817073339605 28$ $p_i = 9781$

Highly optimised program

770 years

Conclusions

Proving and Computing

Careful Computing

Parallelising with theorems