Computability and Enumeration

Igor Pak, UCLA

Joint work with Scott Garrabrant

FPSAC, KAIST, Daejeon, South Korea

July 8, 2015

Let's count...

The number of ways to approach mathematics:

(1) **Adventure**

(2) **Beauty**

(3) **Order**

...

(4) **Complexity** $\leftarrow$ clue to this story!

Questions about combinatorial sequences

Let $\{a_n\}$ be a combinatorial sequence, e.g.

- $a_n =$ number of graphs on n vertices, s.t. $(*)$
- $a_n =$ number of permutations in S_n, s.t. $(**)$
- $a_n =$ number of Dyck paths of length $2n$, s.t. $(***)$

Question 1: Does $\mathcal{A}(t) = \sum_n a_n t^n$ have a *formula*?

Question 2: Can a_n be computed efficiently?

Classes of combinatorial sequences

(1) ***rational*** g.f. $\mathcal{A}(t) = P(t)/Q(t), \quad P, Q \in \mathbb{Z}[t]$

e.g. $a_n = \mathrm{Fib}(n),\ \mathcal{A}(t) = 1/(1 - t - t^2)$

(2) ***algebraic*** g.f. $c_0\mathcal{A}^k + c_1\mathcal{A}^{k-1} + \ldots + c_k = 0, \quad c_i(t) \in \mathbb{Z}[t]$

e.g. $a_n = \mathrm{Cat}(n),\ \mathcal{A}(t) = (1 - \sqrt{1 - 4t})/2t$

(3) ***D-finite*** g.f. $c_0\mathcal{A} + c_1\mathcal{A}' + \ldots + c_k\mathcal{A}^{(k)} = 0, \quad c_i(t) \in \mathbb{Z}[t]$

e.g. $a_n = \#\text{involutions in } S_n,\ a_n = a_{n-1} + (n-1) \cdot a_{n-2}$

Sequences $\{a_n\}$ are called ***P-recursive***.

(4) ***ADE*** g.f. $Q(t, \mathcal{A}, \mathcal{A}', \ldots, \mathcal{A}^{(k)}) = 0,\ \ Q \in \mathbb{Z}[t, x_0, x_1, \ldots, x_k]$

e.g. $a_n = \#\{\sigma(1) < \sigma(2) > \sigma(3) < \ldots \in S_n\},\ \mathcal{A}' = \mathcal{A}^2 + 1$

Permutation classes

Permutation $\sigma \in S_n$ *contains* $\omega \in S_k$ if M_ω is a submatrix of M_σ. Otherwise, σ *avoids* ω. Such ω are called *patterns*.

For example, (5674123) contains (321) but avoids (4321).

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \qquad \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

For a set of patterns $\mathcal{F} \subset S_k$, denote $\mathcal{C}_n(\mathcal{F})$ the set of $\sigma \in S_n$ avoiding each $\omega \in \mathcal{F}$. Let $C_n(\mathcal{F}) = |\mathcal{C}_n(\mathcal{F})|$.

Notable examples:

(1) $C_n(123) = C_n(213) = \mathrm{Cat}(n)$ [MacMahon, 1915] and [Knuth, 1973].

(2) $C_n(123, 132, 213) = \mathrm{Fib}(n+1)$ [Simion, Schmidt, 1985]

(3) $C_n(2413, 3142) = \mathrm{Schröder}(n)$ [Shapiro, Stephens, 1991]

(4) $C_n(1234) = C_n(2143)$ has D-finite g.f. [Gessel, 1990]

(5) $C_n(1342)$ has algebraic g.f. [Bona, 1997]

(6) $C_n(3412, 4231)$ has algebraic g.f. [Bousquet-Mélou, Butler, 2007]
counts the number of smooth Schubert varieties X_σ, $\sigma \in S_n$
[Lakshmibai, Sandhya, 1990], see also [Abe, Billey, 2014].

Main result

Noonan–Zeilberger Conjecture:

For every $\mathcal{F} \subset S_k$, the sequence $\{C_n(\mathcal{F})\}$ is P-recursive.

(Equivalently, the g.f. for $\{C_n(\mathcal{F})\}$ is D-finite).

Theorem 1. [Garrabrant, P., 2015+]

NZ Conjecture is false. To be precise, there is a set $\mathcal{F} \subset S_{80}$, $|\mathcal{F}| < 31000$, s.t. $\{C_n(\mathcal{F})\}$ is **not** P-recursive.

A bit of history

- First stated as an open problem by Gessel (1990)
- Upgraded to a conjecture and extended to count copies contained of each pattern, by Noonan and Zeilberger (1996)
- Atkinson reduced the extended version to the original (1999)
- In 2005, Zeilberger changes his mind, conjectures that $\{C_n(1324)\}$ is not P-recursive [this is still open]
- In 2014, Zeilberger changes his mind half-way back, writes: "if I had to bet on it now I would give only a 50% chance".

As bad as it gets!

Main Lemma [here $\mathbb{X}$ is LARGE, to be clarified below]

Let $\xi : \mathbb{N} \to \mathbb{N}$ be a function in $\mathbb{X}$. Then there exist $k, a, b \in \mathbb{N}$ and sets of patterns $\mathcal{F}, \mathcal{F}' \subset S_k$, s.t.

$\xi(n) = C_{an+b}(\mathcal{F}) - C_{an+b}(\mathcal{F}')$ mod 2 for all n.

Note: Here mod 2 can me changed to any mod p but cannot be completely removed. For example, $C_n(\mathcal{F}) = 0$ implies $C_{n+1}(\mathcal{F}) = 0$, which does not hold for functions $\xi \in \mathbb{X}$.

Theorem 2. [Garrabrant, P., 2015+]

The problem whether $C_n(\mathcal{F}) - C_n(\mathcal{F}') = 0$ mod 2 for all $n \geq 1$, is undecidable.

Not convinced yet?

Corollary 1. For all k large enough, there exists $\mathcal{F}, \mathcal{F}' \subset S_k$ such that the smallest n for which $C_n(\mathcal{F}) \neq C_n(\mathcal{F}')$ mod 2 satisfies

$$n > 2^{2^{2^{2^{2^{2^k}}}}}.$$

Corollary 2. There exist two finite sets of patterns $\mathcal{F}$ and $\mathcal{F}'$, such that the problem of whether $C_n(\mathcal{F}) = C_n(\mathcal{F}')$ mod 2 for all $n \in \mathbb{N}$, is independent of ZFC.

Computational Complexity Classes

$\oplus\mathsf{P}$ = parity version of the class of counting problem $\#\mathsf{P}$

e.g. $\oplus$Hamiltonian cycles in $G \in \oplus\mathsf{P}$

$\mathsf{P} \neq \oplus\mathsf{P}$ is similar to $\mathsf{P} \neq \mathsf{NP}$

In fact, $\mathsf{P} = \oplus\mathsf{P}$ implies $\mathsf{PH} = \mathsf{NP} = \mathsf{BPP}$ [by Toda's theorem]

EXP = exponential time

$\oplus\mathsf{EXP}$ = exponential time version of $\oplus\mathsf{P}$

e.g. $\oplus$Hamiltonian 3-connected graphs on n vertices $\in \oplus\mathsf{EXP}$

$\mathsf{EXP} \neq \oplus\mathsf{EXP}$ is similar to $\mathsf{P} \neq \oplus\mathsf{P}$

believed to be correct for more technical CC reasons,

Complexity Implications

Theorem 3. [Garrabrant, P., 2015+]
If $\mathsf{EXP} \neq \oplus\mathsf{EXP}$, then there exists a finite set of patterns $\mathcal{F}$, such that the sequence $\{C_n(\mathcal{F})\}$ cannot be computed in time polynomial in n.

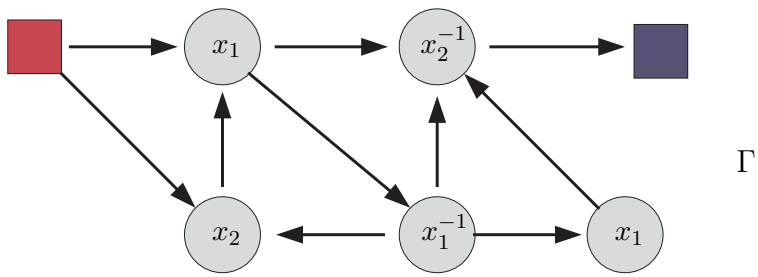
Remark 1: All sequences with D-finite g.f. can be computed in time polynomial in n.

Remark 2: This also answers to Wilf's question (1982):

Can one describe a reasonable and natural family of combinatorial enumeration problems for which there is provably no polynomial-in-n time formula or algorithm to compute $f(n)$?

Two-stack Automata

In the Main Lemma, $\mathbb{X} = \{\xi_\Gamma\}$, where $\xi_\Gamma(n)$ = number of balanced paths of some two-stack automaton Γ.



Here $\xi(1) = \xi(2) = \xi(3) = 0$, $\xi(4) = 1$, $\xi(5) = 0$, $\xi(6) = 1$.

Note: Two-stack automata are as powerful as Turing machines.

How not to be P-recursive

Lemma 1. Let $\{a_n\}$ be a P-recursive sequence, and let $\overline{\alpha} = (\alpha_1, \alpha_2, \ldots) \in \{0,1\}^\infty$, $\alpha_i = a_i \bmod 2$. Then there is a finite binary word $w \in \{0,1\}^*$ which is NOT a subword of $\overline{\alpha}$.

Lemma 2. There is a two-stack automaton Γ s.t. the number of balanced paths $\xi_\Gamma(n)$ is given by the sequence

$$0,\ 1,\ 0,0,\ 0,1,\ 1,0,\ 1,1,\ 0,0,0,\ 0,0,1,\ 0,1,0,\ \ldots$$

Now Lemma 1, Lemma 2 and Main Lemma imply Theorem 1.

Main Lemma: outline

(0) Allow general **partial patterns** (rectangular 0-1 matrices with no two 1's in the same row or column).

(1) Fix a sufficiently large "alphabet" of "incomparable" matrices Specifically, we take all simple 10-permutations which contain (5674123). Arbitrarily name them $P, Q, B, B', E, T_1, \ldots, T_v, Z_1, \ldots, Z_m$.

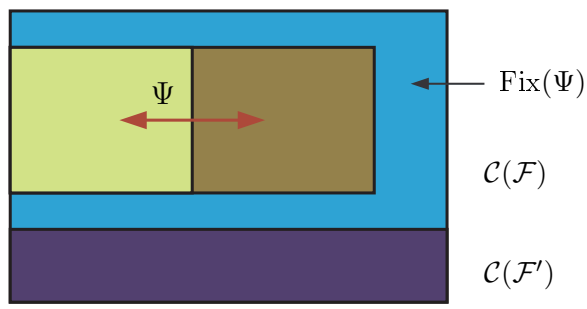
(2) Thinking of T_i's as vertices of Γ and Z_j as variables x_p, y_q, select block matrices $\mathcal{F}$ to simulate Γ. Let $\mathcal{F}' = \mathcal{F} \cup \{B, B'\}$.

(3) Define involution Ψ on $\mathcal{C}_n(\mathcal{F}) \setminus \mathcal{C}_n(\mathcal{F}')$ by $B \leftrightarrow B'$. Check that fixed points of Ψ are in bijection with balanced paths in Γ.

Sample of forbidden matrices in $\mathcal{F}$:

$$\begin{pmatrix} \circ & \circ & T_i & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & T_j & \circ & \circ \\ L & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & \circ & \circ & Z_p \\ \circ & \circ & \circ & \circ & \circ & \circ & T_k & \circ \\ \circ & B' & \circ & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & R & \circ & \circ & \circ \\ \circ & \circ & \circ & Z_p & \circ & \circ & \circ & \circ \end{pmatrix} \qquad \begin{pmatrix} \circ & \circ & T_i & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ & \circ & T_2 \\ L & \circ & \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & E & \circ & \circ \\ \circ & \circ & \circ & \circ & Q & \circ \\ \circ & B' & \circ & \circ & \circ & \circ \end{pmatrix} \qquad \begin{pmatrix} \circ & \circ & T_j \\ Z_p & \circ & \circ \\ \circ & Z_q & \circ \end{pmatrix}$$

Final count:



Example

$$
M = \begin{pmatrix}
\circ & \circ & \circ & \circ & T_1 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & P & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & E & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & T_3 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
B & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & Z_1 & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & T_5 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & B & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & Z_2 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & T_3 & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & B & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \bullet & \circ & \circ & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & Z_1 & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \bullet & \circ & \circ & \bullet & \circ & \circ & \bullet & \circ & \circ & T_6 & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \bullet & B & \circ & \bullet & \circ & \circ & \bullet & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & Z_2 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & T_4 & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & B & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & Z_1 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & T_2 & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & B & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & E & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \bullet & \circ & T_4 & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & B & \circ & \circ & \circ & \circ & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & Z_1 & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & T_2 \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & B & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & E & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & Q & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & B & \circ & \circ & \circ & \circ
\end{pmatrix}.
$$

Notes on the proof

(i) We use exactly 6854 partial patterns.

(ii) Automaton Γ in Lemma 2 uses 31 vertices, which is why the alphabet has size 10×10 only.

(iii) The largest matrix in $\mathcal{F}$ has 8×8 blocks, which is why Theorem 1 has permutations in S_{80}.

(iv) Proof of Lemma 1 has 2 paragraphs, but it took over a year of hard work to state. Natural extensions remain open.

Conjecture 0. [Garrabrant, P.] Let $\overline{\alpha}$ be as in Lemma 1. Then $\overline{\alpha}$ has $O(n)$ subwords of length n.

The non-ADE extension

Theorem 1′. [Garrabrant, P., in preparation]

There is a set $\mathcal{F} \subset S_{80}$, s.t. the g.f. for $\{C_n(\mathcal{F})\}$ is **not** ADE.

Lemma 1′. Let $\{a_n\}$ be an integer sequence, and let $\{n_i\}$ be the sequences of indices with **odd** a_n. Suppose

1) for all $b, c \in \mathbb{N}$, there exists k such that $n_k = b \mod 2^c$,
2) $n_k/n_{k+1} \to 0$ as $k \to \infty$.

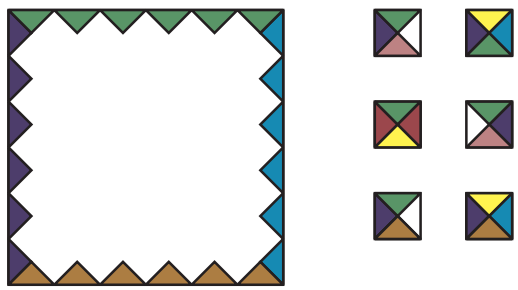
Then the g.f. for $\{a_n\}$ is **not** ADE.

Corollary. Let $\{a_n\}$ be an integer sequence, s.t. a_n is odd if only if $n = k! + k$, for some k. Then the g.f. for $\{a_n\}$ is **not** ADE.

Note: cf. EC2, Exc. 6.63c.

First prequel: Wang tilings

Long and classical story going back to 1960s (Wang, Berger, Robinson, etc.) Key result: tileability of the plane with fixed set of Wang tiles is undecidable. Delicate part: ensuring that the "seed tile" must be present in a tiling. This is what we do by introducing $\mathcal{F}'$.



Second prequel: Kontsevich's problem

Let G be a group and $\mathbb{Z}[G]$ denote its group ring. Fix $u \in \mathbb{Z}[G]$.
Let $a_n = [1]u^n$, where $[g]u$ denote the value of u on $g \in G$.
In 2014, Maxim Kontsevich asked whether $\{a_n\}$ is always P-recursive when $G \subseteq \mathrm{GL}(k, \mathbb{Z})$.

Theorem 4. [Garrabrant, P., 2015+]
There exists an element $u \in \mathbb{Z}[\mathrm{SL}(4, \mathbb{Z})]$, such that the sequence $\{[1]u^n\}$ is not P-recursive.

Note: Proof uses the same Lemma 1(!)
When $G = \mathbb{Z}^k$ or $G = F_k$, the sequence $\{a_n\}$ is known to be P-recursive for all $u \in \mathbb{Z}[G]$ (Haiman, 1993).

Open problems:

Conjecture 1. The *Wilf-equivalence* problem of whether $C_n(\mathcal{F}_1) = C_n(\mathcal{F}_2)$ for all $n \in \mathbb{N}$ is undecidable.

Conjecture 2. For forbidden sets with a single permutation $|\mathcal{F}| = |\mathcal{F}'| = 1$, the Wilf-equivalence problem is decidable.

Conjecture 3. Sequence $\{C_n(1324)\}$ is not P-recursive.

Conjecture 4. There exists a finite set of patterns $\mathcal{F}$, s.t. computing $\{C_n(\mathcal{F})\}$ is $\#\mathsf{EXP}$-complete, and computing $\{C_n(\mathcal{F}) \bmod 2\}$ is $\oplus\mathsf{EXP}$-complete.

Grand Finale:

A story how Doron Zeilberger lost $100.

Thank you!