Base-$\frac{p}{Q}$ structure of states in automata arising from Christol's theorem

Eric Rowland
Hofstra University

Joint work with Reem Yassawi

Numeration
Utrecht, 2024–6–4

Numeration system

Let p be a prime, $D = \{0, 1, \ldots, p-1\}$, and $\alpha \geq 1$.

Let $Q \in \mathbb{Z}[x, y]$ such that $Q(0, 0) \not\equiv 0 \mod p$.
$\frac{1}{Q}$ has a series expansion modulo p^α.

Base-$\frac{p}{Q}$ representation with digits $T_k \in D[x, y]$:

$$\left(T_0 + T_1 \tfrac{p}{Q} + T_2 \left(\tfrac{p}{Q}\right)^2 + \cdots + T_{\alpha-1} \left(\tfrac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1} \mod p^\alpha$$

Example

Let $p = 2$, $\alpha = 2$, $Q = 1 + x + xy^2$,

$$S = 1 + 3x^2 + (3 + 2x + 3x^2)y + 2xy^2 + x^2y^4 + x^2y^5 \in (\mathbb{Z}/4\mathbb{Z})[x, y].$$

Then $S \equiv \left((1 + y) + (x + y + xy)\frac{2}{Q}\right) Q \mod 4$.
Its digits are $1 + y$ and $x + y + xy$.

To obtain a finite set of kth digits, require $\deg T_k \leq c\,(k + 1)$.

$$\boxed{S \equiv \left(T_0 + T_1 \, \tfrac{p}{Q} + \cdots + T_{\alpha-1} \left(\tfrac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1} \mod p^\alpha}$$

Not every polynomial S has a representation.
Necessary condition: $S \equiv T_0 Q^{p^{\alpha-1}-1} \mod p$

Proposition

If S has a representation, then this representation is unique.

Proof: Assume

$$\left(T_0 + T_1 \, \tfrac{p}{Q} + \cdots + T_{\alpha-1} \left(\tfrac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1}$$
$$\equiv \left(U_0 + U_1 \, \tfrac{p}{Q} + \cdots + U_{\alpha-1} \left(\tfrac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1} \mod p^\alpha.$$

Then $T_0 Q^{p^{\alpha-1}-1} \equiv U_0 Q^{p^{\alpha-1}-1} \mod p$, so $T_0 = U_0$.
Also

$$T_0 Q^{p^{\alpha-1}-1} + T_1 p Q^{p^{\alpha-1}-2} \equiv U_0 Q^{p^{\alpha-1}-1} + U_1 p Q^{p^{\alpha-1}-2} \mod p^2,$$

which implies $T_1 = U_1$. And so on.

$$\boxed{S \equiv \left(T_0 + T_1 \tfrac{p}{Q} + \cdots + T_{\alpha-1} \left(\tfrac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1} \mod p^\alpha}$$

Perform carries if a coefficient doesn't belong to $D = \{0, 1, \ldots, p-1\}$.

Suppose $T_k \notin D[x, y]$. Quotient by p:

$$\begin{aligned} S &\equiv \left(\cdots + T_k \left(\tfrac{p}{Q}\right)^k + T_{k+1} \left(\tfrac{p}{Q}\right)^{k+1} + \cdots\right) Q^{p^{\alpha-1}-1} \\ &\equiv \left(\cdots + (R_k + pU_k) \left(\tfrac{p}{Q}\right)^k + T_{k+1} \left(\tfrac{p}{Q}\right)^{k+1} + \cdots\right) Q^{p^{\alpha-1}-1} \\ &\equiv \left(\cdots + R_k \left(\tfrac{p}{Q}\right)^k + (U_k Q + T_{k+1}) \left(\tfrac{p}{Q}\right)^{k+1} + \cdots\right) Q^{p^{\alpha-1}-1}. \end{aligned}$$

Corollary

The set of representable polynomials in $(\mathbb{Z}/p^\alpha\mathbb{Z})[x, y]$ is closed under addition and scalar multiplication.

Why is this numeration system natural?

Theorem (Christol 1979/1980)

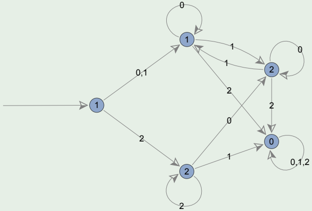
A sequence $s(n)_{n\geq 0}$ of elements in $\mathbb{F}_q$ is algebraic if and only if it is q-automatic.

Example

$q = 3, \quad s(n)_{n\geq 0} = 1, 1, 2, 2, 2, 0, 0, 0, 2, 2, 2, 1, 1, 1, 0, 0, \ldots$

The generating series $F = \sum_{n\geq 0} s(n)x^n$ satisfies $xF^2 + 2F + 1 = 0$.

This automaton outputs $s(n)$ when fed the base-3 representation of n:



$s(11) = s(102_3) = 1$

Catalan numbers $C(n)_{n\geq 0} = 1, 1, 2, 5, 14, 42, \ldots \qquad xy^2 - y + 1 = 0$

Catalan numbers modulo 3: $1, 1, 2, 2, 2, 0, \ldots$

What about $C(n) \bmod p^\alpha$?

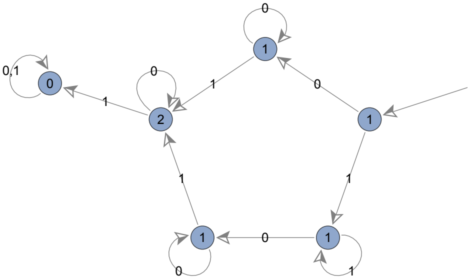
Catalan numbers modulo 4: $1, 1, 2, 1, 2, 2, 0, 1, 2, 2, 0, 2, 0, 0, 0, 1, \ldots$

Theorem (Eu–Liu–Yeh 2008)

For all $n \geq 0$,

$$C(n) \bmod 4 = \begin{cases} 1 & \text{if } n = 2^a - 1 \text{ for some } a \geq 0 \\ 2 & \text{if } n = 2^b + 2^a - 1 \text{ for some } b > a \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

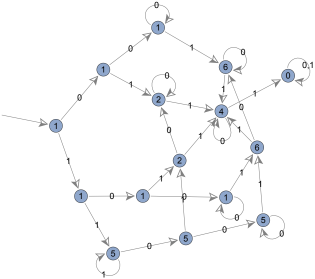
In particular, $C(n) \not\equiv 3 \mod 4$.



Catalan numbers modulo 8: $1, 1, 2, 5, 6, 2, 4, 5, 6, 6, 4, 2, 4, 4, 0, 5, \ldots$

Theorem 4.2. *Let C_n be the nth Catalan number. First of all, $C_n \not\equiv_8 3$ and $C_n \not\equiv_8 7$ for any n. As for other congruences, we have*

$$C_n \equiv_8 \begin{cases} 1 & \text{if } n = 0 \text{ or } 1; \\ 2 & \text{if } n = 2^a + 2^{a+1} - 1 \text{ for some } a \geq 0; \\ 4 & \text{if } n = 2^a + 2^b + 2^c - 1 \text{ for some } c > b > a \geq 0; \\ 5 & \text{if } n = 2^a - 1 \text{ for some } a \geq 2; \\ 6 & \text{if } n = 2^a + 2^b - 1 \text{ for some } b - 2 \geq a \geq 0; \\ 0 & \text{otherwise.} \end{cases}$$



Why are these sequences 2-automatic?

Theorem (Denef–Lipshitz 1987)

A sequence $s(n)_{n\geq 0}$ of elements in $\mathbb{Z}/p^\alpha\mathbb{Z}$ is p-automatic if and only if $\sum_{n\geq 0} s(n)x^n \equiv F \mod p^\alpha$ for some algebraic series $F \in \mathbb{Z}_p[\![x]\!]$.

$\mathbb{Z}_p$ is the set of p-adic integers.

How big is the automaton for $(C(n) \bmod 2^\alpha)_{n\geq 0}$?

α	1	2	3	4	5	6	7	8	9
# states	5	6	15	37	83	194	445	1034	2403

Suggested asymptotics: $p^{\text{polynomial function of } \alpha}$

Upper bound from the construction: $p^{p^{2(\alpha-1)}\alpha h d}$

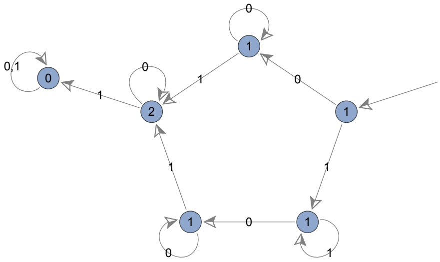
height $h = \deg_x P$

degree $d = \deg_y P$

$P = xy^2 - y + 1$

Why is the bound so large?

$C(n) \bmod 4$:



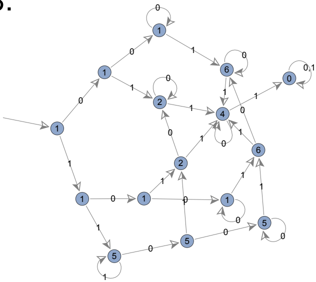
Each state is represented by a polynomial:

$$\begin{aligned}
S_0 &= 1 + 2x + x^2 + (1 + 3x)y + 2xy^2 + (x + 2x^2)y^3 + 3x^2y^4 + 2x^2y^5 \\
\lambda_0(S_0) &= 1 + 2x + x^2 + (2x + 2x^2)y + 2x^2y^3 + 3x^2y^4 \\
\lambda_1(S_0) &= 1 + 3x + (3 + 3x)y + xy^2 + xy^3 \\
\lambda_1(\lambda_0(S_0)) &= 2 + 2x + 2xy^2 \\
\lambda_0(\lambda_1(S_0)) &= 1 + 3x + 2xy + 3xy^2 \\
\lambda_1(\lambda_1(\lambda_0(S_0))) &= 0
\end{aligned}$$

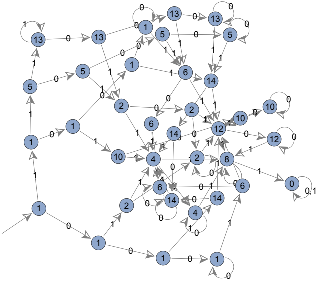
What's special about these polynomials?

$(C(n) \bmod 4)_{n \geq 0}$ projects to $(C(n) \bmod 2)_{n \geq 0}$.
The corresponding automata project to each other...

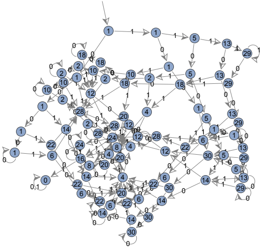
mod 8:



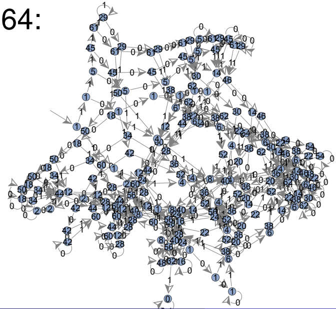
mod 16:



mod 32:



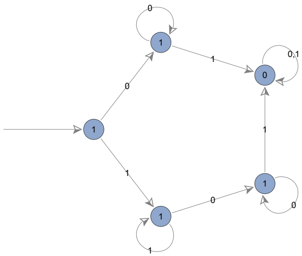
mod 64:



The states should project to each other too…

$C(n)$ mod 2:

$$
\begin{aligned}
S_0 &= 1 + x + y + xy^2 \\
\lambda_0(S_0) &= 1 + x + xy^2 \\
\lambda_1(S_0) &= 1 + y \\
\lambda_1(\lambda_0(S_0)) &= 0 \\
\lambda_0(\lambda_1(S_0)) &= 1
\end{aligned}
$$



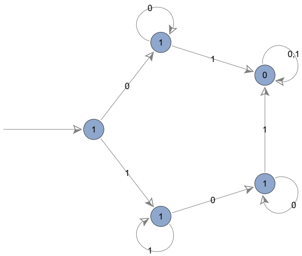
$C(n)$ mod 4:

$$
\begin{aligned}
S_0 &= 1 + 2x + x^2 + (1 + 3x)y + 2xy^2 + (x + 2x^2)y^3 + 3x^2y^4 + 2x^2y^5 \\
\lambda_0(S_0) &= 1 + 2x + x^2 + (2x + 2x^2)y + 2x^2y^3 + 3x^2y^4 \\
\lambda_1(S_0) &= 1 + 3x + (3 + 3x)y + xy^2 + xy^3 \\
\lambda_1(\lambda_0(S_0)) &= 2 + 2x + 2xy^2 \\
\lambda_0(\lambda_1(S_0)) &= 1 + 3x + 2xy + 3xy^2 \\
\lambda_1(\lambda_1(\lambda_0(S_0))) &= 0
\end{aligned}
$$

The states should project to each other too...

$C(n)$ mod 2:

$$S_0 = 1 + x + y + xy^2$$
$$\lambda_0(S_0) = 1 + x + xy^2$$
$$\lambda_1(S_0) = 1 + y$$
$$\lambda_1(\lambda_0(S_0)) = 0$$
$$\lambda_0(\lambda_1(S_0)) = 1$$



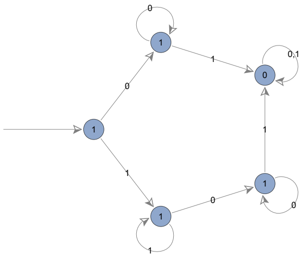
$C(n)$ mod 4: Reduce modulo 2...

$$S_0 \equiv (1 + x + y + xy^2)(1 + x + xy^2) \mod 2$$
$$\lambda_0(S_0) \equiv (1 + x + xy^2)^2 \mod 2$$
$$\lambda_1(S_0) \equiv (1 + y)(1 + x + xy^2) \mod 2$$
$$\lambda_1(\lambda_0(S_0)) \equiv 0 \mod 2$$
$$\lambda_0(\lambda_1(S_0)) \equiv 1 + x + xy^2 \mod 2$$
$$\lambda_1(\lambda_1(\lambda_0(S_0))) \equiv 0 \mod 2$$

They're all divisible by $1 + x + xy^2$ modulo 2!

The states should project to each other too...

$C(n)$ mod 2:



$$S_0 = 1 + x + y + xy^2$$
$$\lambda_0(S_0) = 1 + x + xy^2$$
$$\lambda_1(S_0) = 1 + y$$
$$\lambda_1(\lambda_0(S_0)) = 0$$
$$\lambda_0(\lambda_1(S_0)) = 1$$

$C(n)$ mod 4: Reduce modulo 2... $Q = 1 + x + xy^2$

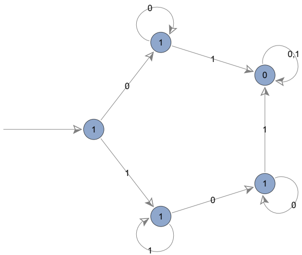
$$S_0 \equiv (1 + x + y + xy^2)Q \mod 2$$
$$\lambda_0(S_0) \equiv (1 + x + xy^2)Q \mod 2$$
$$\lambda_1(S_0) \equiv (1 + y)Q \mod 2$$
$$\lambda_1(\lambda_0(S_0)) \equiv 0 \cdot Q \mod 2$$
$$\lambda_0(\lambda_1(S_0)) \equiv 1 \cdot Q \mod 2$$
$$\lambda_1(\lambda_1(\lambda_0(S_0))) \equiv 0 \cdot Q \mod 2$$

They're all divisible by $1 + x + xy^2$ modulo 2!

The states should project to each other too...

$C(n)$ mod 2:

$$S_0 = 1 + x + y + xy^2$$
$$\lambda_0(S_0) = 1 + x + xy^2$$
$$\lambda_1(S_0) = 1 + y$$
$$\lambda_1(\lambda_0(S_0)) = 0$$
$$\lambda_0(\lambda_1(S_0)) = 1$$



$C(n)$ mod 4: Base-$\frac{p}{Q}$ representations... $Q = 1 + x + xy^2$

$$S_0 \equiv \left(\left(1 + x + y + xy^2\right) + \left(xy + x^2y^2 + x^2y^3 + x^2y^4 + x^2y^5\right)\tfrac{2}{Q}\right)Q$$
$$\lambda_0(S_0) \equiv \left(\left(1 + x + xy^2\right) + \left((x + x^2)y + (x + x^2)y^2 + x^2y^3 + x^2y^4\right)\tfrac{2}{Q}\right)Q$$
$$\lambda_1(S_0) \equiv \left(\left(1 + y\right) + \left(x + (1 + x)y\right)\tfrac{2}{Q}\right)Q \mod 4$$
$$\lambda_1(\lambda_0(S_0)) \equiv \left(0 + \left(1 + x + xy^2\right)\tfrac{2}{Q}\right)Q \mod 4$$
$$\lambda_0(\lambda_1(S_0)) \equiv \left(1 + \left(x + xy + xy^2\right)\tfrac{2}{Q}\right)Q \mod 4$$
$$\lambda_1(\lambda_1(\lambda_0(S_0))) \equiv \left(0 + 0\tfrac{2}{Q}\right)Q \mod 4$$

The 0th digit gives the projected state modulo 2.

Where does Q come from?
If $F = \sum_{n\geq 1} s(n)x^n$ satisfies $P(x, F) = 0$, let $Q = P(xy, y)/y$.

Catalan: $P = x(y+1)^2 - (y+1) + 1$, so $Q = xy^2 + 2xy + x - 1$.

Theorem

If $s(n)_{n\geq 0}$ is an algebraic sequence of integers, then every state in the automaton for $(s(n) \bmod p^\alpha)_{n\geq 0}$ has a unique base-$\frac{p}{Q}$ representation

$$\left(T_0 + T_1\,\frac{p}{Q} + T_2\left(\frac{p}{Q}\right)^2 + \cdots + T_{\alpha-1}\left(\frac{p}{Q}\right)^{\alpha-1}\right) Q^{p^{\alpha-1}-1}$$

where $T_k \in D[x, y]$ for each $k \in \{0, 1, \ldots, \alpha - 1\}$.

We have bounds on $\deg_x T_k$ and $\deg_y T_k$. $D = \{0, 1, \ldots, p - 1\}$

Much better upper bound:

$$(1 + o(1))\, p^{\frac{1}{6}\alpha(\alpha+1)((2hd-1)\alpha+hd+1)}$$

References

- Gilles Christol, Teturo Kamae, Michel Mendès France, and Gérard Rauzy, Suites algébriques, automates et substitutions, *Bulletin de la Société Mathématique de France* **108** (1980) 401–419.
- Jan Denef and Leonard Lipshitz, Algebraic power series and diagonals, *Journal of Number Theory* **26** (1987) 46–67.
- Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh, Catalan and Motzkin numbers modulo 4 and 8, *European Journal of Combinatorics* **29** (2008) 1449–1466.
- Eric Rowland and Reem Yassawi, Automatic congruences for diagonals of rational functions, *Journal de Théorie des Nombres de Bordeaux* **27** (2015) 245–288.
- Eric Rowland, Manon Stipulanti, and Reem Yassawi, Algebraic power series and their automatic complexity I: finite fields, `https://arxiv.org/abs/2308.10977` (29 pages).