Finite Automata and Curves

Andrew Bridy

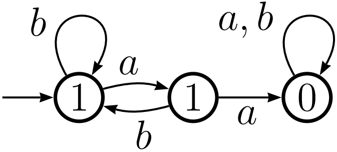
Yale University

November 6, 2018

Finite Automata

A finite automaton is a machine that processes words. It has finitely many states, one of which is marked as initial, and reads a word one letter at a time to determine how to move between states. Each state is marked with an output that is produced once the entire word is read.

The automaton below outputs 1 when a word does not contain two consecutive a's, and 0 otherwise.



$abab \mapsto 1$
$aabb \mapsto 0$

An automaton is a computer (Turing machine) with *no memory*.

Finite Automata

Formally, an automaton $\mathcal{A}$ consists of a finite input alphabet Σ, a finite output alphabet Δ, and

- A finite set of states Q
- An initial state $q_0 \in Q$
- A transition function $\delta : Q \times \Sigma \to Q$
- An output function $\tau : Q \to \Delta$

$\mathcal{A}$ repeatedly applies δ to the current state and current letter of the word, then applies τ to the last state reached. (We should also fix a convention of reading left-to-right or right-to-left.)

If $\Delta = \{0, 1\}$, then $\mathcal{A}$ accepts or rejects each input word. The set of words accepted by an automaton is a regular language (Kleene).

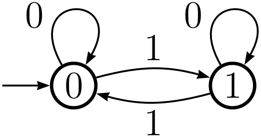
Automatic Sequences

Let $k \geq 2$. A sequence $\mathbf{a}$ is k-automatic if there exists an automaton (k-DFAO) with input alphabet $\Sigma_k := \{0, 1, \ldots, k-1\}$ that maps the base-k expansion of the integer n to the output $\mathbf{a}(n)$.

For example, the nth term of the Thue-Morse sequence

$$\mathbf{a} = 01101001\ldots$$

is the parity of the sum of the bits in the binary expansion of n. The sequence $\mathbf{a}$ is 2-automatic, as demonstrated below:



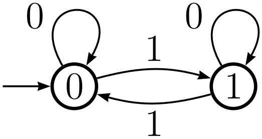
Christol's Theorem

Theorem (Christol)

The power series $\sum_{n=0}^{\infty} \mathbf{a}(n)x^n \in \mathbb{F}_p[[x]]$ *is algebraic over* $\mathbb{F}_p(x)$ *if and only if the sequence* $\mathbf{a}$ *is* p*-automatic.*

Christol's Theorem: Example

Let **a** be the Thue-Morse sequence.



$\mathbf{a}$=01101001…

Let $y = \sum_{n=0}^{\infty} \mathbf{a}(n)x^n \in \mathbb{F}_2[[x]]$. It is easy to check that $\mathbf{a}(2n) = \mathbf{a}(n)$ and $\mathbf{a}(2n+1) = \mathbf{a}(n) + 1$. Therefore

$$\begin{aligned} y &= \sum \mathbf{a}(2n)x^{2n} + \sum \mathbf{a}(2n+1)x^{2n+1} \\ &= \sum \mathbf{a}(n)x^{2n} + \sum \mathbf{a}(n)x^{2n+1} + \sum 1x^{2n+1} \\ &= \left(\sum \mathbf{a}(n)x^n\right)^2 + x\left(\sum \mathbf{a}(n)x^n\right)^2 + \frac{x}{1+x^2} \\ &= y^2 + xy^2 + \frac{x}{1+x^2}. \end{aligned}$$

Christol's Theorem: Consequences

With Christol's Theorem, certain hard problems of algebra become easy problems of automata theory.

Irrelevance of Symbols: If $\sum \mathbf{a}(n)x^n \in \mathbb{F}_p[[x]]$ is algebraic, then applying any self-mapping of $\mathbb{F}_p$ to the terms of the sequence $\mathbf{a}$ (e.g. changing each appearance of 1 to 2) preserves algebraicity.

Hadamard Product: If $\sum \mathbf{a}(n)x^n, \sum \mathbf{b}(n)x^n \in \mathbb{F}_p[[x]]$ are algebraic, then $\sum \mathbf{a}(n)\mathbf{b}(n)x^n$ is algebraic. This is because the product of automatic sequences is automatic – picture two machines running in parallel. (This fails in characteristic 0.)

The Main Question

Question

What is the relation between the complexity (number of states) of a minimal p-DFAO that outputs $\mathbf{a}$ and the complexity of the algebraic power series $y = \sum_{n=0}^{\infty} \mathbf{a}(n)x^n \in \mathbb{F}_p[[x]]$?

Let $N_p(\mathbf{a})$ (respectively $N_p^f(\mathbf{a})$) be the complexity of a minimal reverse-reading (respectively forward-reading) p-DFAO that outputs the sequence $\mathbf{a}$.

It is not clear how to define the complexity of an algebraic function. The most obvious choice is algebraic degree, but this will not be enough information for our purposes.

Algebraic Complexity

Let y be algebraic over $\mathbb{F}_p(x)$ with minimal polynomial

$$P(x,T) = T^d + f_{d-1}T^{d-1} + \cdots + f_1 T + f_0 \in \mathbb{F}_p(x)[T].$$

That is, P has minimal degree such that $P(x,y)=0$. Let

- $d = \deg(y) = \deg_T(P)$,
- $h = \mathrm{height}(y)$ be the maximum of the degrees of the coefficients $f_i \in \mathbb{F}_p(x)$, and
- $g = \mathrm{genus}(y)$ be the genus of the normalization of the projective closure of the plane curve defined by $P = 0$.

It can be deduced from the usual proof of Christol's theorem that $N_p(\mathbf{a}) \leq p^{pd^4h^2}$. Examples suggest this is very far from sharp.

The Main Theorem

Theorem (B.)

Let $y = \sum_{n=0}^{\infty} \mathbf{a}(n)x^n \in \mathbb{F}_p[[x]]$ *be algebraic over* $\mathbb{F}_p(x)$*. Then*

$$N_p(\mathbf{a}) \leq (1 + o(1))p^{h+d+g-1},$$

where $o(1)$ *tends to* 0 *as any of* $p, h, d, g \to \infty$*.*

This bound is qualitatively sharp for rational functions (sharp neglecting the $o(1)$ term). For forward-reading complexity,

Theorem (B.)

$$N_p^f(\mathbf{a}) \leq p^{h+2d+g-1}.$$

It is possible to eliminate g from the bounds. For example, by Castelnuovo's Inequality, $N_p(\mathbf{a}) \leq (1 + o(1))p^{hd}$.

The p-kernel

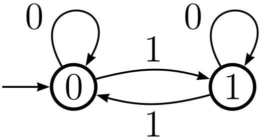
The p-kernel of a sequence $\mathbf{a}$ is defined as the set of subsequences $n \mapsto \mathbf{a}(p^i n + j)$ such that $i \geq 0$ and $0 \leq j < p^i$.

Theorem (Eilenberg)

A sequence $\mathbf{a}$ *is p-automatic iff its p-kernel is finite.*

In fact, there is a bijection between the p-kernel of $\mathbf{a}$ and the states of a minimal reverse-reading p-DFAO that outputs $\mathbf{a}$.

For the Thue-Morse sequence:



$$\mathbf{a}(n) = \mathbf{a}(2n) = \mathbf{a}(4n) = \mathbf{a}(4n+3) = \dots$$
$$\mathbf{a}(2n+1) = \mathbf{a}(4n+1) = \mathbf{a}(4n+2) = \dots$$

$$\#2\text{-kernel}(\mathbf{a}) = N_2(\mathbf{a}) = 2$$

Cartier Operators on Power Series

For $i \in \{0, \dots, p-1\}$, define the $\mathbb{F}_p$-linear operator

$$\Lambda_i : \mathbb{F}_p[[x]] \to \mathbb{F}_p[[x]]$$

by

$$\Lambda_i \left(\sum_{n=0}^{\infty} \mathbf{a}(n) x^n \right) = \sum_{n=0}^{\infty} \mathbf{a}(pn+i) x^n.$$

Let S be the semigroup $\langle \Lambda_0, \Lambda_1, \dots, \Lambda_{p-1} \rangle$. The orbit $S(\sum_{n=0}^{\infty} \mathbf{a}(n) x^n)$ is in bijection with the p-kernel of $\mathbf{a}$ (thus with a minimal p-DFAO).

Observe that for any $y, z \in \mathbb{F}_p[[x]]$,

- $\Lambda_i(z^p y) = z \Lambda_i(y)$ and
- $y = (\Lambda_0(y))^p + x(\Lambda_1(y))^p + \cdots + x^{p-1}(\Lambda_{p-1}(y))^p$.

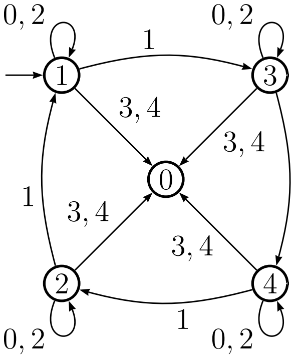
Automata from Cartier operators

Take $y \in \mathbb{F}_5[[x]]$ given by

$$y = (1-x)^{-1/2} = 1 + 3x + x^2 + 3x^5 + 4x^6 + \dots$$

We compute $y = y^5(1 + 3x + x^2)$, and so $\Lambda_i(y) = y\Lambda_i(1 + 3x + x^2)$. This leads to a 5-DFAO that outputs the coefficients of y:

$$\Lambda_0(y) = y$$
$$\Lambda_1(y) = 3y$$
$$\Lambda_2(y) = y$$
$$\Lambda_3(y) = 0$$
$$\Lambda_4(y) = 0$$



Differentials

Let X be a curve over $\mathbb{F}_p$, and let K be its function field. The $\mathbb{F}_p$-vector space Ω of (Kähler) differentials of X is generated by symbols of the form df for $f \in K$, subject to the following relations.

- $d(f+g) = df + dg$
- $d(fg) = f\,dg + g\,df$
- $da = 0$ for $a \in \mathbb{F}_p$

It can be shown that $\dim_K(\Omega) = 1$.

Let $v_P(f)$ denote the order of vanishing of $f \in K$ at the point P of X. At any P, there exists a local coordinate $t \in K$ such that $v_P(t) = 1$. Any $\omega \in \Omega$ can be written as $\omega = f\,dt$. Define $v_P(\omega) = v_P(f)$.

The Cartier Operator on Ω

Fix $x \in K$ such that $x \notin K^p$ (equivalently, $dx \neq 0$). Any $\omega \in \Omega$ can be written uniquely as

$$\omega = \left(u_0^p + u_1^p x + u_2^p x^2 + \cdots + u_{p-1}^p x^{p-1} \right) dx$$

for some $u_0, \ldots, u_{p-1} \in K$. Define $\mathcal{C} : \Omega \to \Omega$ by $\mathcal{C}(\omega) = u_{p-1}\, dx$. Amazingly, this construction does not depend on the choice of the coordinate x!

The Cartier operator $\mathcal{C}$ is an $\mathbb{F}_p$-linear operator on Ω with many nice properties. In particular, if $v_P(\omega) < 0$, then $v_P(\mathcal{C}(\omega)) \geq v_P(\omega)$, so $\mathcal{C}$ "improves" the poles of differentials.

Christol Revisited

Define the twisted Cartier operator $\sigma_i : \Omega \to \Omega$ by

$$\sigma_i(\omega) = \mathcal{C}(x^{p-i-1}\omega).$$

For any $y = \sum_{n=0}^{\infty} \mathbf{a}(n) x^n \in \mathbb{F}_p[[x]] \cap K$, we have

$$\sigma_i(y\,dx) = \Lambda_i(y)\,dx.$$

Let $\mathcal{S}$ be the semigroup $\langle \sigma_0, \dots, \sigma_{p-1} \rangle$. For $s \in \mathcal{S}$ and a point P of X, the tendency of the Cartier operator to improve poles yields

$$v_P(s(y\,dx)) \geq \min\{0, v_P(y\,dx)\} + \min\{0, v_P(x)\}.$$

So $\mathcal{S}(y\,dx)$ consists of differentials with poles at finitely many points of bounded orders. By the Riemann-Roch theorem, $\mathcal{S}(y\,dx)$ is finite. By Eilenberg's theorem, $\mathbf{a}$ is p-automatic (Speyer).

Proof Sketch of Main Theorem

If X is the curve defined by the minimal polynomial of y, then $\mathcal{S}(y\,dx)$ is contained in the Riemann-Roch space

$$V = \Omega((y\,dx)_\infty + (x)_\infty).$$

A straightforward computation in the algebraic geometry of curves gives $\dim V \leq h + 3d + g - 1$ and therefore $N_p(\mathbf{a}) \leq p^{h+3d+g-1}$. Lowering the bound requires more finesse: $\mathcal{S}$ eventually moves most of the differentials in V into smaller Riemann-Roch spaces.

To pass from $N_p(\mathbf{a})$ to $N_p^f(\mathbf{a})$, we replace V by its dual space. The output function of the automaton can be identified with an adele (repartition), and the Cartier operator on differentials is replaced by the Frobenius operator on adeles.

Example: $y = (1+x^3)^{-1/2} \in \mathbb{F}_5[[x]]$

Let $y = \sum \mathbf{a}(n)x^n \in \mathbb{F}_5[[x]]$ be a root of $y^2(x^3+1) = 1$. This equation defines a curve X that contains y in its function field. We compute

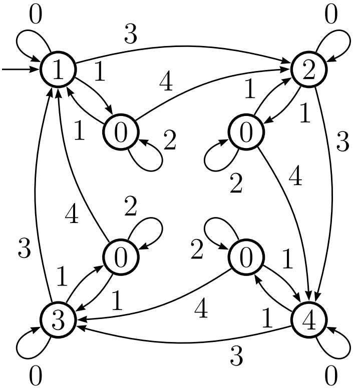
$$(y\,dx)_\infty = 0 \quad \text{and} \quad (x)_\infty = 2P_\infty.$$

By the proof of the main theorem, $\mathcal{S}(y\,dx) \subseteq \Omega(2P_\infty)$.

A Riemann-Roch calculation gives $\dim \Omega(2P_\infty) = 2$, with a basis given by $\{y\,dx, xy\,dx\}$. Therefore $N_5(\mathbf{a}) \leq 25$. In fact, $N_5(\mathbf{a}) = 9$, and we can explicitly compute a 5-DFAO that outputs $\mathbf{a}$.

Example: $y = (1 + x^3)^{-1/2} \in \mathbb{F}_5[[x]]$

Below is the desired 5-DFAO. All missing transitions go to a trap state (not pictured) that outputs 0. The σ_i operators are written explicitly as endomorphisms of $\Omega(2P_\infty)$ in our chosen basis.



$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\sigma_3 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\sigma_4 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$$

Questions for Further Study

- Is the bound sharp in general?
- What about multidimensional sequences and the multivariate analogue of Christol's Theorem?
- How does state complexity behave under algebraic operations on power series (+, ×, composition)?
- Is there a canonical correspondence (functor) between automata and curves? Could this provide a geometric approach to semigroup representation theory?

Finite Automata and Curves

Andrew Bridy

Yale University

November 6, 2018