Representing algebraic streams using behavioural differential equations

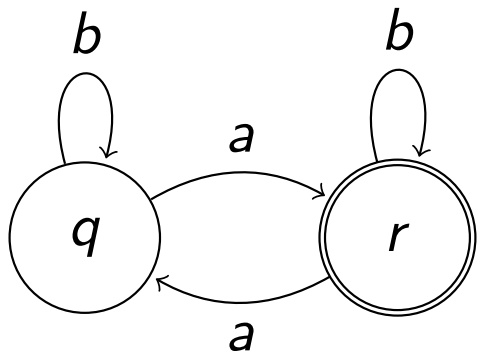
Joost Winter (jww Marcello Bonsangue and Jan Rutten)

Centrum Wiskunde & Informatica, Leiden University

December 12, 2012

Behavioural differential equations: a specification format

The automaton

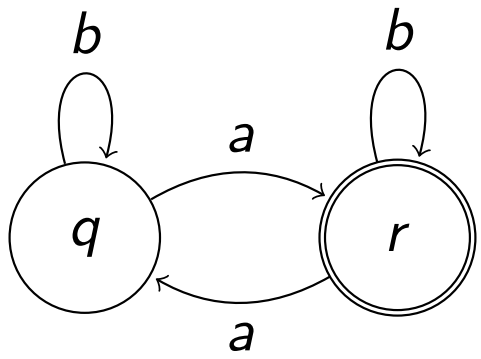


corresponds to the system of *behavioural differental equations*

$$\begin{array}{llll} o(q) = 0 & q_a = r & q_b = q \\ o(r) = 1 & r_a = q & r_b = r. \end{array}$$

Behavioural differential equations: a specification format

The automaton



corresponds to the system of *behavioural differential equations*

$$\begin{array}{llll} o(q) = 0 & q_a = r & q_b = q \\ o(r) = 1 & r_a = q & r_b = r. \end{array}$$

This defines a system $o, \delta : Q \to \mathbb{B} \times Q^A$ over state space $Q = \{q, r\}$ and alphabet $A = \{a, b\}$.
Final coalgebra mappings:

$$\begin{array}{lll} [\![q]\!] & = & \{w \in A^* \mid \#_a(w) = 1 \mod 2\} \\ [\![r]\!] & = & \{w \in A^* \mid \#_b(w) = 0 \mod 2\} \end{array}$$

Word derivatives

Note that q_a is just a shorthand for $\delta(q)(a)$.
We can extend the notion of derivatives from alphabet symbols to words inductively:

$$\begin{aligned} q_1 &= q \\ q_{a \cdot w} &= (q_a)_w. \end{aligned}$$

For streams: alphabet is $\{\mathcal{X}\}$, write q' for $q_{\mathcal{X}}$ and $q^{(n)}$ for $q_{\mathcal{X}^n}$.

Bisimulation: a proof method

Given automata P and Q, a relation $R \subseteq P \times Q$ is called a *bisimulation* iff

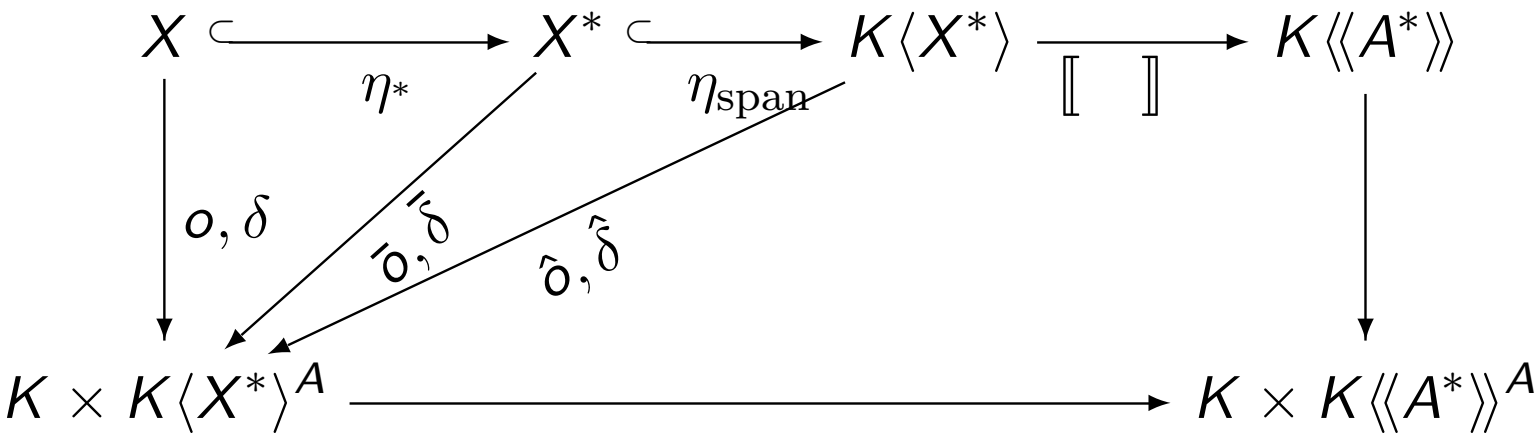
1. $o(p) = o(q)$ for all $p, q \in P, Q$, and
2. $p_a \; R \; q_a$ for all $p, q, a \in P, Q, A$.

Theorem

Coinductive proof princple: if $q \; R \; r$ for some bisimulation R, then $[\![q]\!] = [\![r]\!]$.

Polynomial systems of b.d.e.s

Let K be a semiring, X and A finite alphabets of nonterminals and terminals, respectively.



1. From o, δ to $\bar{o}, \bar{\delta}$: inductively defined product rule

$$\begin{array}{rclcrcl} o(1) & = & 1 & & 1_a & = & 0 \\ o(xw) & = & o(x)\bar{o}(w) & & (xw)_a & = & x_a w + o(x) w_a \end{array}$$

   for all $a \in A, x \in X, w \in X^*$.

2. From $\bar{o}, \bar{\delta}$ to $\hat{o}, \hat{\delta}$: determinization.

A context-free system

An example of such a system (using $\mathbb{B}$ as underlying semiring):

$$\begin{array}{lllllllll} o(x) & = & 1 & \quad x_a & = & xy & \quad x_b & = & 0 \\ o(y) & = & 0 & \quad y_a & = & 0 & \quad y_b & = & 1 \end{array}$$

A context-free system

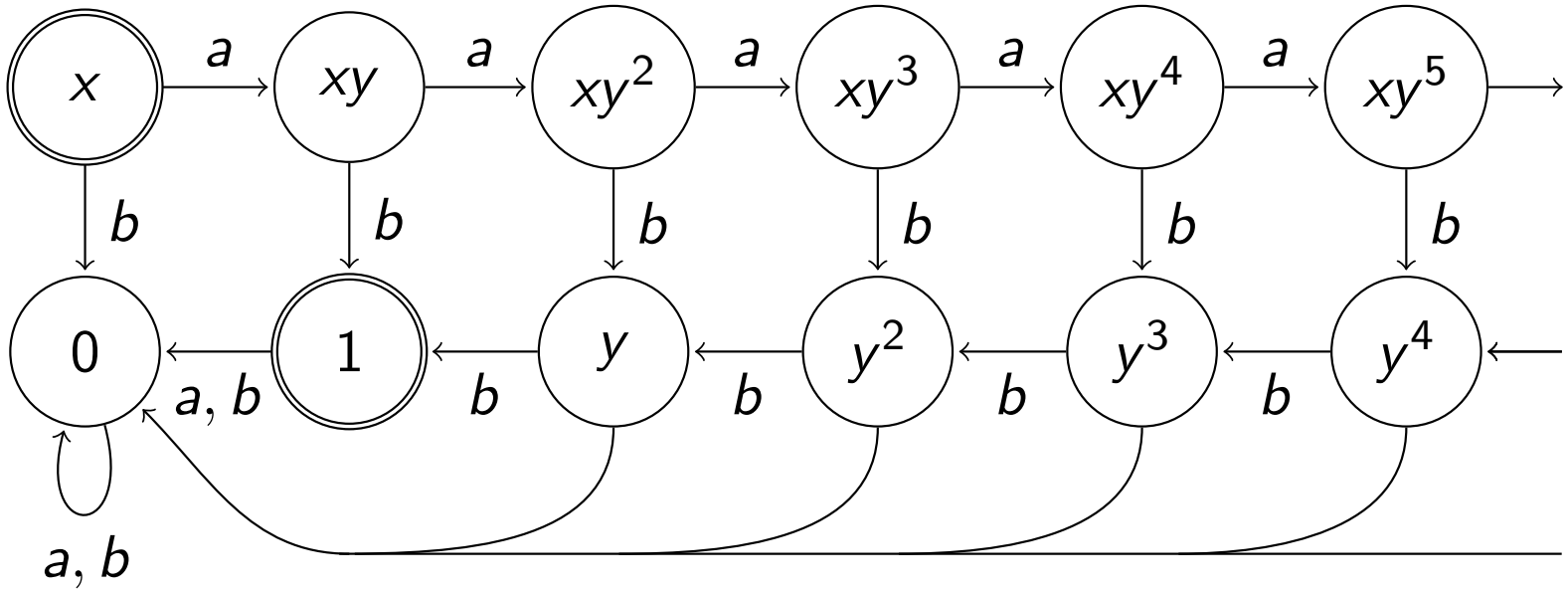
An example of such a system (using $\mathbb{B}$ as underlying semiring):

$$\begin{array}{lllllllll} o(x) & = & 1 & x_a & = & xy & x_b & = & 0 \\ o(y) & = & 0 & y_a & = & 0 & y_b & = & 1 \end{array}$$

We get

$$[\![x]\!] = \{a^n b^n \mid n \in \mathbb{N}\}$$

. . . and the resulting (infinite) automaton



Correspondences

Theorem

Using $\mathbb{B}$ as underlying semiring, the following are equivalent:

1. *$L \in \mathbb{B}\langle\langle A \rangle\rangle$ is a context-free language.*
2. *There exists an alphabet X, a polynomial system of b.d.e.s on X, and a $x \in X$, such that $[\![x]\!] = L$.*
3. *There exists an alphabet X, a polynomial system of b.d.e.s on X, and a $t \in \mathbb{B}\langle X^* \rangle$, such that $[\![t]\!] = L$.*

Correspondence between 1 and 2: through Greibach normal form.

A more general picture

Theorem

Given a semiring K, the following are equivalent:

1. $\sigma \in K\langle\langle A \rangle\rangle$ *is (constructively) K-algebraic.*
2. *There exists an alphabet X, a polynomial system of b.d.e.s on X, and a $x \in X$, such that $[\![x]\!] = \sigma$.*
3. *There exists an alphabet X, a polynomial system of b.d.e.s on X, and a $t \in K\langle X^* \rangle$, such that $[\![t]\!] = \sigma$.*

Correspondence between 1 and 2: through Greibach normal form.

Some systems of b.d.e.s for streams

Over $\mathbb{N}$:

Fibonacci numbers (1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ...)

$$\begin{array}{rclcrcl} o(x) & = & 1 & & x' & = & y \\ o(y) & = & 1 & & y' & = & x + y \end{array}$$

Catalan numbers (1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, ...)

$$o(x) = 1 \qquad x' = x^2$$

Large Schroeder numbers (1, 2, 6, 22, 90, 394, 1806, 8558, ...)

$$o(x) = 1 \qquad x' = x^2 + x$$

(as a g.f.: $\frac{1-\mathcal{X}-(1-6\mathcal{X}+\mathcal{X}^2)^{\frac{1}{2}}}{2\mathcal{X}}$)

Motzkin numbers (1, 1, 2, 4, 9, 21, 51, 127, 323, 835, ...)

$$\begin{array}{rclcrcl} o(x) & = & 1 & & x' & = & y \\ o(y) & = & 1 & & y' & = & x^2 + y \end{array}$$

Some systems of b.d.e.s for streams

Over $\mathbb{B}$:

Prouhet-Thue-Morse sequence $(0, 1, 1, 0, 1, 0, 0, 1, 1, 0, \dots)$

$$\begin{aligned} o(x) &= 0 & x' &= y \\ o(y) &= 1 & y' &= y^2 + \mathcal{X}z^2 \\ o(z) &= 0 & z' &= z^2 + \mathcal{X}y^2 \\ o(\mathcal{X}) &= 0 & \mathcal{X}' &= 1 \end{aligned}$$

Paperfolding sequence $(1, 1, 0, 1, 1, 0, 0, 1, 1, 1, \dots)$

$$\begin{aligned} o(u) &= 1 & u' &= x \\ o(x) &= 1 & x' &= y^2 + \mathcal{X}x^2 \\ o(y) &= 0 & y' &= z^2 + \mathcal{X}w^2 \\ o(z) &= 1 & z' &= (1 + \mathcal{X})z^2 \\ o(w) &= 0 & w' &= (1 + \mathcal{X})w^2 \\ o(\mathcal{X}) &= 0 & \mathcal{X}' &= 1 \end{aligned}$$

Fliess' theorem

Theorem (Fliess 1971)

Given a perfect field K and a $\sigma \in K^{\mathbb{N}}$, σ is K-algebraic iff it is constructively K-algebraic.

More correspondences

1. A $\mathbb{N}$-stream is constructively $\mathbb{N}$-algebraic iff it is the growth function of an unambiguous context-free language. (Chomsky-Schützenberger)
2. A $\mathbb{F}_q$-stream is constructively $\mathbb{F}_q$-algebraic iff it is q-automatic (Christol + Fliess) ...
3. ... iff it is the growth function modulo q of an unambiguous context-free language.
4. A $\mathbb{B}$-stream is constructively $\mathbb{B}$-algebraic iff it is eventually periodic.

Bisimulation up to linear combinations

Given automata X and Y in the category of K-modules, define the function

$$\Sigma : \mathcal{P}(X \times Y) \to \mathcal{P}(X \times Y)$$

by

$$\Sigma R = \{ \left(\sum_{i<m} k_i x_i, \sum_{i<m} k_i y_i \right) \mid m \in \mathbb{N}, \forall i < m : x_i \; R \; y_i \}$$

ΣR: see it as projections of linear combinations of pairs from R. We call a relation $R \subseteq X \times Y$ a bisimulation up to linear combinations whenever

1. $x \in X, y \in Y$, $o(x) = o(y)$ for all $x \in X, y \in Y$ and
2. $x_a \; \Sigma \; R y_a$ for all $x \in X, y \in Y, a \in A$.

Some facts about bisimulation up to linear combinations

1. For all R, $R \subseteq \Sigma R$.
2. R is a bisimulation up to linear combinations iff ΣR is a bisimulation.
3. $\sim = \Sigma \sim$.

Stream functions: even, odd, zip

$$\begin{aligned} \mathsf{even}(\sigma)(n) &= \sigma(2n) \\ \mathsf{odd}(\sigma)(n) &= \sigma(2n+1) \end{aligned}$$

$$\begin{aligned} \mathsf{zip}(\sigma,\tau)(2n) &= \sigma(n) \\ \mathsf{zip}(\sigma,\tau)(2n+1) &= \tau(n) \end{aligned}$$

Facts about even, odd, and zip

1. $$\sigma = \mathsf{zip}(\mathsf{even}(\sigma), \mathsf{odd}(\sigma)),$$

2. $$o(\mathsf{zip}(\sigma, \tau)) = o(\sigma) \qquad \mathsf{zip}(\sigma, \tau)' = \mathsf{zip}(\tau, \sigma'), \qquad \text{and}$$

3. when the underlying semiring is the finite field $\mathbb{F}_2$:

$$\mathsf{zip}(\sigma, \tau) = \sigma^2 + \mathcal{X}\tau^2.$$

zip preserves (constructive) algebraicity

Theorem

If σ and τ are constructively K-algebraic streams, then $\mathrm{zip}(\sigma, \tau)$ is constructively K-algebraic.

Automatic sequences (1)

A *q-automaton* is a finite automaton over the alphabet of digits

$$A_q = \{\bar{0}, \ldots \overline{q-1}\}.$$

Words $w \in A_q^*$ represent numbers in base q (with the least significant bit first), and we let the function

$$[\,] : A_q^* \to \mathbb{N}$$

assign to each word of digits the natural number it represents.

Automatic sequences (2)

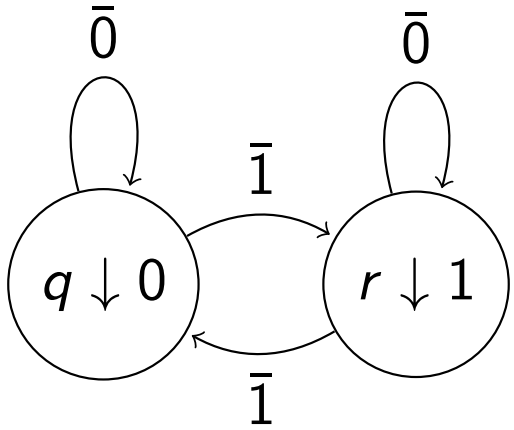
We call a stream σ over the field $\mathbb{F}_q$ q-automatic, whenever there is a q-automaton (Q, o, δ) and a state $q \in Q$, such that for each word $w \in A_q^*$,

$$o(q_w) = \sigma([w]),$$

and write $\mathrm{str}(q) = \sigma$ whenever this holds.
We call a state q consistent whenever $o(q) = o(q_{\bar{0}})$.
A state q is consistent iff $\mathrm{str}(q) = \sigma$ for some σ.

Example: the Prouhet-Thue-Morse sequence (1)

In the 2-automaton



both states are consistent. Furthermore, $\mathrm{str}(q)$ is equal to the Prouhet-Thue-Morse sequence

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

Correspondence between derivatives and even/odd

Lemma

Given a 2-automaton (Q, o, δ) and a consistent $q \in Q$, we have

$$\mathrm{str}(q_{\bar{0}}) = \mathrm{even}(\mathrm{str}(q)) \qquad \text{and} \qquad \mathrm{str}(q_{\bar{1}}) = \mathrm{odd}(\mathrm{str}(q)).$$

Correspondence between derivatives and even/odd

Lemma

Given a 2-automaton (Q, o, δ) and a consistent $q \in Q$, we have

$$\mathrm{str}(q_{\bar{0}}) = \mathrm{even}(\mathrm{str}(q)) \qquad \text{and} \qquad \mathrm{str}(q_{\bar{1}}) = \mathrm{odd}(\mathrm{str}(q)).$$

Proof.

The first equality holds because

$$\begin{aligned} \mathrm{str}(q_{\bar{0}})([w]) &= o((q_{\bar{0}})_w) \\ &= o(q_{\bar{0}\cdot w}) \\ &= \mathrm{str}(q)([\bar{0} \cdot w]) \\ &= \mathrm{str}(q)(2 \cdot [w]) \\ &= \mathrm{even}(\mathrm{str}(q))([w]) \end{aligned}$$

and the second equality can be proven analogously. □

Systems of equations from 2-automatic sequences

Given a consistent 2-automaton (Q, o, δ), consider the following system of equations, with the following set of variables

$$X = \{\bar{x} \mid x \in Q\} \cup \{\mathcal{X}\}$$

and set, for all $x \in Q$:

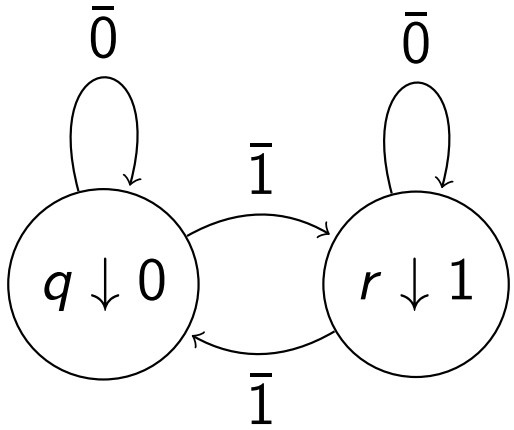
$$\begin{array}{cc} o(\bar{x}) = o(x_{\bar{1}}) & \bar{x}' = \overline{x_{\bar{0}}}^2 + \mathcal{X}\overline{x_{\bar{1}}}^2 \\ o(\mathcal{X}) = 0 & \mathcal{X}' = 1 \end{array}$$

Theorem

For all $x \in Q$, $[\![\bar{x}]\!] = \mathrm{str}(x)'$. *Hence* $\mathrm{str}(x)$ *is* $\mathbb{F}_2$*-algebraic.*

Example: the Prouhet-Thue-Morse sequence (2)

Applying this construction to the earlier example,



we get:

$$o(q) = 1 \qquad q = q^2 + \mathcal{X} r^2$$
$$o(r) = 0 \qquad r = r^2 + \mathcal{X} q^2$$
$$o(\mathcal{X}) = 0 \qquad \mathcal{X}' = 1$$

Generalizations

This construction can easily be generalized from 2-automata to q-automata. Here q must be of the form p^n for some prime p and natural number n.

- Instead of $\mathrm{zip}(\sigma, \tau)$, we now have $\mathrm{zip}_q(\sigma_0, \ldots, \sigma_{q-1})$.
- Instead of even and odd, we now have $\mathrm{unzip}_{i,q}(\sigma)$ for all $i \in \mathbb{N}$ with $i < q$.
- Instead of $\mathrm{zip}(\sigma, \tau) = \sigma^2 + \mathcal{X}\tau^2$, we now have

$$\mathrm{zip}_q(\sigma_0, \ldots, \sigma_q) = \sum_{i<q} \mathcal{X}^i \sigma_i^q.$$

Related work

- Work on automatic sequences as regular coalgebras with even/odd-specifications by Kupke/Rutten, including a nice 'relatively final' coalgebra theorem.
- Grabmayer/Endrullis/Hendriks/Klop/Moss also obtained similar results in a paper about automatic sequences and zip-specifiations.