Arden's Rule and the Kleene-Schützenberger Theorem

Joost Winter

Centrum Wiskunde & Informatica

December 3, 2013

The Kleene-Schützenberger theorem

- Rational power series (or languages, streams): power series characterizable by rational expressions (over arbitrary semirings S).
- Recognizable power series (or languages, streams): power series that can be recognized by a weighted automaton.
- Kleene-Schützenberger theorem: S-rational $=$ S-recognizable.
- Proven by Kleene for $\mathbb{B}$ (Kleene's theorem), by Schützenberger for $\mathbb{Z}$ and by Eilenberg for arbitrary semirings S.
- Coalgebraic proof by Rutten for $\mathbb{B}$ in both directions, and for arbitrary semirings in the rational $\to$ recognizable direction.

Formal power series

Given a semiring S and a finite alphabet A, let $S\langle\langle A\rangle\rangle$ denote the function space:

$$\{\sigma \mid \sigma \in A^* \to S\}$$

We assign a semiring structure to $S\langle\langle A\rangle\rangle$ (we use 1 to denote the empty word):

$$\begin{array}{rcl} 0(w) & = & 0 \\ 1(w) & = & \textbf{if } w = 1 \textbf{ then } 1 \textbf{ else } 0 \\ (\sigma + \tau)(w) & = & \sigma(w) + \tau(w) \\ (\sigma\tau)(w) & = & \displaystyle\sum_{uv=w} \sigma(u)\tau(v) \end{array}$$

Also: alphabet injections $A \to S\langle\langle A\rangle\rangle$:

$$a(w) \quad = \quad \textbf{if } w = a \textbf{ then } 1 \textbf{ else } 0$$

Formal power series (2)

We can also assign *output* and *derivative* operators O and Δ on $S\langle\langle A\rangle\rangle$

$$\begin{array}{rcl} O(\sigma) & = & \sigma(1) \\ \Delta(\sigma)(a)(w) & = & \sigma(aw) \end{array}$$

and will simply write σ_a for $\Delta(\sigma)(a)$.
The semiring structure on $S\langle\langle A\rangle\rangle$ now can be characterized using the following *behavioural differential equations*:

$$\begin{array}{rclrcl} O(0) & = & 0 & 0_a & = & 0 \\ O(1) & = & 1 & 1_a & = & 0 \\ O(b) & = & 0 & b_a & = & \textbf{if } b = a \textbf{ then } 1 \textbf{ else } 0 \\ O(\sigma + \tau) & = & O(\sigma) + O(\tau) & (\sigma + \tau)_a & = & \sigma_a + \tau_a \\ O(\sigma\tau) & = & O(\sigma)O(\tau) & (\sigma\tau)_a & = & \sigma_a\tau + O(\sigma)\tau_a \end{array}$$

Polynomials and proper series

- We call a power series $\sigma \in S\langle\langle A\rangle\rangle$ a *polynomial* iff for only finitely many $w \in A^*$, $\sigma(w) \neq 0$.
  The set of polynomials in $S\langle\langle A\rangle\rangle$ is denoted by $S\langle A\rangle$.
- We call a power series $\sigma \in S\langle\langle A\rangle\rangle$ *proper* iff $O(\sigma) = 0$.

Recognizable series

Some equivalent characterizations:

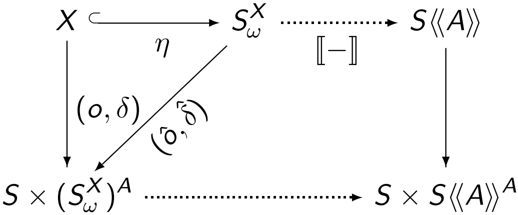
- A power series is S-recognizable iff it occers as the solution to a linear system of behavioural differential equations.
- A power series σ_0 is S-recognizable iff there is a finite set $\Sigma = \{\sigma_0, \ldots, \sigma_k\}$ s.t. for each $\sigma \in \Sigma$ and each $a \in A$, σ_a is a linear combinations of elements from Σ.
- A power series σ is S-recognizable iff there is a $k \in \mathbb{N}$, and there are $c_{ij}, b_i \in S$, such that σ occurs as a component of the unique solution in $S\langle\langle A \rangle\rangle$ to the system of equations

$$x_i = b_i + \sum_{a \in A} a \sum_{j \leq n} c_{ij} x_j$$

- A power series σ is S-recognizable iff it is contained in a stable finitely generated submodule of $S\langle\langle A \rangle\rangle$.

Recognizable series (2)

- A power series σ is S-recognizable iff it occurs in the final coalgebra mapping of the determinization of a $S \times (S_\omega^X)^A$-coalgebra, as follows:



- A power series σ is S-recognizable iff σ is accepted by a finite S-weighted automaton.

The star operator

The star operator can be defined in several ways:

- If we assume a topological structure on S (i.e. S is a topological semiring), we can define σ^* as the limit

$$\sigma^* = \lim_{n \to \infty} \sum_{i=0}^{n} \sigma^i$$

  (wherever this limit exists).
- Simple coinductive definition: σ^* is defined iff σ is proper, and in this case σ^* is defined as:

$$O(\sigma^*) = 1 \qquad (\sigma^*)_a = \sigma_a(\sigma^*)$$

For *any* semiring, we can obtain a topological semiring by assuming the discrete topology on S. The coinductive definition of the star is always compatible with this definition.

Rational power series

Given a set $X \subseteq S\langle\langle X \rangle\rangle$, the class of S-rational power series in X $\mathbf{Rat}_S[X]$ can be defined as the smallest subset of $S\langle\langle A \rangle\rangle$ such that

1. $X \subseteq \mathbf{Rat}_S[X]$
2. $S\langle X \rangle \subseteq \mathbf{Rat}_S[X]$
3. $\mathbf{Rat}_S[X]$ is closed under the operators $+$ and $\cdot$
4. If $\sigma \in \mathbf{Rat}_S[X]$ and σ is proper, then σ^* in $\mathbf{Rat}_S[X]$

We call a power series simply S-rational if it is S-rational in the empty set.
Any element of $\mathbf{Rat}_S[X]$ can be described using a rational (regular) expression with variables in X.

Rational to recognizable

Induction on size of regular expressions. Base cases trivial.
If σ_0 and τ_0 are recognizable, there are Σ and T with $\sigma_0 \in \Sigma$, $\tau_0 \in T$, s.t. for each $\sigma \in \Sigma$ and $\tau \in T$ and $a \in A$, σ_a and τ_a can be written as a linear combination of elements of Σ and T, respectively.

- $(\sigma + \tau)_a = \sigma_a + \tau_a$ so $\Sigma \cup T \cup \{\sigma + \tau\}$ again has the required property (i.e. 'is a stable finitely generated S-submodule of $S\langle\langle A \rangle\rangle$').
- For $(\sigma\tau)_a = \sigma_a\tau + o(\sigma)\tau_a$ so $\{\sigma\tau \mid \sigma \in \Sigma, \tau \in T\} \cup T$ has the required property.
- If σ is proper, $(\sigma^*)_a = \sigma_a\sigma^*$, and $(\upsilon\sigma^*)_a = \upsilon_a\sigma^* + o(\upsilon)\sigma_a\sigma^*$, so $\{\upsilon\sigma^* \mid \upsilon \in \Sigma\} \cup \{\sigma^*\}$ has the required property.

Arden's Rule (left version)

Lemma

Given any $\sigma, \tau \in S\langle\langle A\rangle\rangle$ *with* τ *proper, the unique solution to the equation*

$$x = \sigma + \tau x$$

is given by:

$$x = \tau^* \sigma$$

General unique solution lemma (left version)

Lemma

Given a $k \in \mathbb{N}$ and a family of r_{ij} $(i, j \leq k)$ that are proper and S-rational in X for all i, j, as well as a family of p_i $(i \leq k)$ that are S-rational in X for all i, the system of equations with components

$$x_i = p_i + \sum_{j=0}^{k} r_{ij} x_j$$

for all $i \leq k$ has a unique solution, and each x_i is S-rational in X.

Proof: Natural induction on k.

Base case, if $k = 0$, there is a single equation

$$x_0 = p_0 + r_{00} x_0$$

and Arden's rule now gives a unique solution

$$x_0 = (r_{00})^* p_0$$

which is K-rational in X again.

General unique solution lemma (left version) (2)

Inductive case: if $k = n + 1$, write the last equation in the system as

$$x_k = p_k + \sum_{j=0}^{n} r_{kj} x_j + r_{kk} x_k,$$

apply Arden's rule:

$$x_k = (r_{kk})^* \left(p_k + \sum_{j=0}^{n} r_{xj} x_j \right)$$

General unique solution lemma (left version) (3)

Substituting this equation for x_k into the equations x_i for $i \leq n$ gives

$$x_i = p_i + r_{ik}(r_{kk})^* p_k + \sum_{j=0}^{n} (r_{ij} + r_{ik}(r_{kk})^* r_{kj}) x_j$$

Now set

$$q_i := p_i + r_{ik}(r_{kk})^* p_k \qquad \text{and} \qquad s_{ij} := r_{ij} + r_{ik}(r_{kk})^* r_{kj}$$

and we get a system in n variables:

$$x_i = q_i + \sum_{j=0}^{n} s_{ij} x_j$$

By IH, this system has a unique solution (with each component rational in X), and it follows that the original system has a unique solution, too (again, with each component rational in X).

From recognizable to rational

If σ is S-recognizable, it occurs as a solution to a system of $n+1$ equations

$$x_i = b_i + \sum_{a \in A} a \sum_{j \leq n} c_{ij} x_j$$

or equivalently

$$x_i = b_i + \sum_{j \leq n} \left(\sum_{a \in A} a c_{ij} \right) x_j$$

Because each b_i is rational, and all $\sum_{a \in A} a c_{ij}$ are rational and proper, it follows from the preceding lemma that the system has a unique solution and all x_i are rational.

(Constructively) algebraic series

Two equivalent characterizations:

- A power series τ is S-algebraic iff there is a finite set Σ with $\tau \in \Sigma$, s.t. for each $\sigma \in \Sigma$ and each $a \in A$, σ_a can be written as a polynomial over Σ.
- A power series τ is S-algebraic iff there is a finite set Σ with $\tau \in \Sigma$, s.t. for each $\sigma \in \Sigma$ and each $a \in A$, σ_a is S-rational in Σ.

Algebraic power series generalize context-free languages, in the sense that a language is context-free iff it is $\mathbb{B}$-algebraic.

Proper systems and their solutions

The traditional way of obtaining (constructively) algebraic series is as solutions to proper systems of equations, generalizing CF grammars. The systems of equations consist of a finite X and a mapping:

$$p : X \to S\langle X + A\rangle$$

A system is called *proper* iff for all $x \in X$, $p(x)(1) = 0$, and for all $x, y \in X$, $p(x)(y) = 0$.
A solution is a mapping $[-] : X \to S\langle\langle A\rangle\rangle$ such that for all x

$$[x] = [p(x)]^\sharp$$

where $[-]^\sharp$ is the inductive extension of $[-]$.
A solution $[-]$ is *strong* iff for all $x \in X$, $O[x] = 0$.

Proper systems to GNF

Proper systems can be represented as

$$x_i = \sum_{j=0}^{k} x_j q_{ij} + \sum_{a \in A} a r_{ia}$$

with q_{ij} rational in X and proper, and r_{ia} rational in X.
Assuming that we have a strong solution, we take the derivative to obtain:

$$(x_i)_a = r_{ia} + \sum_{j=0}^{k} (x_j)_a q_{ij}$$

Now apply the (right version of the) unique solution lemma to conclude that all $(x_i)_a$ are rational in X.

Conclusions and future work

- A uniform way of presenting two different results via a sufficiently generally formulated lemma: the Kleene-Schützenberger theorem and the construction of the Greibach Normal form from proper systems.
- The construction of the GNF does, unlike traditional presentations, not require a detour via the Chomsky Normal form.
- The construction of the GNF transforms a proper system in n nonterminals into a GNF-system in $2n + |A|$ nonterminals, less than the $n^2 + n$ nonterminals yielded by Rosenkrantz' procedure.
- Future work: investigate the connections with other limit notions/topologies, unique solutions vs. least solutions, ϵ-transitions and construction of proper systems from arbitrary systems.