MATHEMATICAL PROBLEMS IN COMPUTATION THEORY
BANACH CENTER PUBLICATIONS, VOLUME 21
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1988

A NOTE ON THE MINIMALIZATION OF TREE-AUTOMATA

BRUNO COURCELLE

Bordeaux University, Department of Mathematics and Computer Science, Talence, France

Introduction

Courcelle has shown in [3] that many classical constructions in language theory can be expressed in terms of a set of basic transformations of systems of polynomial equations, which are correct in the sense that they preserve the least solutions of the systems to which they apply. This note shows that this is the case for the *minimalization* of tree-automata (reduction and determinization have already been treated in [3]).

This paper assumes the knowledge of Gecseg–Steinby [4]. The necessary definitions are recalled from [3].

1. Polynomial systems

Let F be a finite ranked alphabet with rank function $\varrho\colon F \to N$. Let $M(F)$ be the set of well-formed terms over F, let $M(F, U)$ be the set of well-formed terms over $F \cup U$ where U is a set of 0-ary variables also called *unknowns* in the sequel.

Let $P = P(M(F))$ be the power set of $M(F)$. We equip P with an algebraic structure defined as follows:

$$\Omega_P = \emptyset \qquad (\Omega \text{ is a constant not in } F),$$

$$T_1 +_P T_2 = T_1 \cup T_2 \qquad (+ \text{ is a binary symbol not in } F),$$

$$f_P(T_1, \ldots, T_k) = \{f(t_1, \ldots, t_k)/\ t_1 \in T_1, \ldots, t_k \in T_k\},$$

where $T_1, T_2, \ldots, T_k \subseteq M(F)$, $f \in F$, $\varrho(f) = k \geqslant 0$. It is not difficult to verify that P is an ω-complete F_+-magma with set inclusion as an ordering. By F_+

we mean $F \cup \{+, \Omega\}$ and F_+-magma is synonymous with F_+-algebra (but we keep the terminology of [2], [3]).

Furthermore, P satisfies the following set $\mathscr{S}$ of equational laws (for all $f \in F$, with $k = \varrho(f)$):

$$T_1 + (T_2 + T_3) = (T_1 + T_2) + T_3,$$

$$T_1 + T_2 = T_2 + T_1,$$

$$T_1 + \Omega = \Omega + T_1,$$

$$T_1 + T_1 = T_1,$$

$$f(T_1, \ldots, \Omega, \ldots, T_k) = \Omega,$$

$$f(T_1, \ldots, T_i + T_i', \ldots, T_k) = f(T_1, \ldots, T_k) + f(T_1, \ldots, T_i', \ldots, T_k),$$

where $T_1, \ldots, T_k, T_i'$ range over $P(M(F))$.

A *polynomial system* is a system of equations of the form $S = \langle u_1 = p_1, \ldots, u_n = p_n \rangle$ where $U = \{u_1, \ldots, u_n\}$ is the set of unknowns of S and each p_i is a *polynomial*, i.e. either Ω or a term in $M(F_+, U)$ of the form

$$p_i = t_{i,1} + t_{i,2} + \ldots + t_{i,n_i} \tag{1}$$

with $t_{i,j} \in M(F, U)$.

Every polynomial system S has a least solution in P consisting of an n-tuple $(L_1, \ldots, L_n)$ of regular sets of elements of $M(F)$ (equivalently of *trees*, see [1], [4]).

Conversely, every regular set of trees L can be defined as the first component of the least solution of a polynomial system of the special form $S = \langle u = p, u_1 = p_1, \ldots, u_n = p_n \rangle$ where p is either Ω or of the form

$$u_{i_1} + u_{i_2} + \ldots + u_{i_k}$$

for some $i_1, \ldots, i_k$ in $\{1, \ldots, n\}$ and each p_i, $i = 1, \ldots, n$, is as in (1) with $t_{i,j}$ of the form

$$f(u_{i_1}, \ldots, u_{i_l})$$

for some $f \in F$, $l = \varrho(f)$, $i_1, \ldots, i_l \in \{1, \ldots, n\}$.

These polynomial systems correspond to finite-state bottom-up tree-automata.

By a *quasi-deterministic system* we shall mean a system of the above form with $t_{i,j} \neq t_{i',j'}$ whenever $i \neq i'$ or $j \neq j'$.

These systems correspond to *deterministic* finite-state bottom-up tree-automata.

It is known from Brainerd [1], Gecseg–Steinby [4] that every deterministic finite-state bottom-up tree-automaton can be transformed into an equivalent one with a minimum number of states.

We want to prove that this transformation is a special case of one of the transformations of polynomial systems which have been defined in Courcelle [3]. We first recall a definition from [3].

2. Redefinition of polynomial systems

Let $S = \langle u_1 = p_1, \ldots, u_n = p_n \rangle$ and $S' = \langle u_1 = p'_1, \ldots, u_n = p'_n \rangle$.

We say that S *redefines into* S', written S **redef** S', if the following conditions hold for some $k \in \{1, 2, \ldots, n\}$:

(1) $$p'_i = p_i \quad \text{for all } i \in \{k+1, \ldots, n\},$$

(2) $$p'_i \in M(F_+, \{u_1, \ldots, u_k\}) \quad \text{for} \quad i \in \{1, \ldots, k\},$$

(3) $$p_i[\Omega/u_1, \ldots, \Omega/u_n] \leqslant^0 p_i'^m[\Omega/u_1, \ldots, \Omega/u_k],$$

for some $m \geqslant 1$, and this for all $i = 1, \ldots, k$,

(4) $$p_i[p_1/u_1, \ldots, p_n/u_n] \overset{*}{\underset{\mathscr{G}}{\longleftrightarrow}} p'_i[p_1/u_1, \ldots, p_k/u_k], \text{ for all } i = 1, \ldots, k.$$

By $p_i[t_1/u_1, \ldots, t_n/u_n]$ we mean the result of the simultaneous substitution of t_1 for $u_1, \ldots, t_n$ for u_n in p_i. By $p_i'^m[\Omega/u_1, \ldots, \Omega/u_n]$ we mean $p'_i[\Omega/u_1, \ldots, \Omega/u_k]$ if $m = 1$ and $p'_i[p_1'^{m'}[\Omega/u_1, \ldots, \Omega/u_k]/u_1, \ldots, p_k'^{m'}[\Omega/u_1, \ldots]/u_k]$ if $m = m'+1$, $m' \geqslant 0$. By $p \overset{*}{\underset{\mathscr{G}}{\longleftrightarrow}} p'$ we mean that p and p' are interconvertible by finitely many uses of the equational laws of $\mathscr{D}$. By $p \leqslant^0 p'$ we mean that p can be transformed into p' by $\overset{*}{\underset{\mathscr{G}}{\longleftrightarrow}}$ and by replacements of Ω by any element of $M(F_+, U)$.

We write more generally S **redef** S' if these conditions hold up to a renaming of $u_1, \ldots, u_n$ (the specific ordering that we have used was just a convenient way to state conditions (1)–(4)). And we shall say that this transformation *redefines* U' where U' is $\{u_1, \ldots, u_k\}$ in conditions (1)–(4) or the set of unknowns corresponding to them if a renaming is used. In [3] the notation $\mathbf{redef}_{\mathscr{D}}$ is used for **redef**.

3. Example

Let $S = \langle u = v+w,\ v = f(g(w))+a+b,\ w = f(g(v))+a \rangle$ and let S' be the same system except for the first equation transformed into $u = f(g(u))+a+b$.

Then S **redef** S' by a redefinition of u. Conditions (1) and (2) are obvious. Condition (3) holds since

$$\Omega+\Omega \overset{*}{\underset{\mathscr{D}}{\longleftrightarrow}} \Omega \leqslant^0 t$$

where t can be anything and in particular $f(g(\Omega))+a+b$. Condition (4) holds since

$$f(g(w))+a+b+f(g(v))+a \xleftrightarrow[\mathscr{G}]{*} f(g(v+w))+a+b$$

(both sides are $\xleftrightarrow[\mathscr{G}]{*}$-equivalent to $f(g(v))+f(g(w))+a+b$).

4. Congruences and quotient systems

Let $S = \langle u = p, u_1 = p_1, \ldots, u_n = p_n \rangle$ be a quasi-deterministic system where

$$p = u_{i_1} + u_{i_2} + \ldots + u_{i_k},$$
$$p_1 = t_{1,1} + \ldots + t_{1,n_1},$$
$$\vdots$$
$$p_n = t_{n,1} + \ldots + t_{n,n_n}.$$

A *congruence* $\equiv$ on S is an equivalence relation on $\{u_1, u_2, \ldots, u_n\}$ such that

(1) if $u_{i_j} \equiv u_{i'}$, then $i' = i_{j'}$ for some $j' = 1, \ldots, k$,

(2) if $t_{i,j} = f(u_{j_1}, \ldots, u_{j_l})$, $t_{i',j'} = f(u_{j'_1}, \ldots, u_{j'_l})$,

$u_{j_1} \equiv u_{j'_1}, \ldots, u_{j_l} \equiv u_{j'_l}$, then $u_i \equiv u_{i'}$,

(3) if $t_{i,j} = f(u_{j_1}, \ldots, u_{j_l})$, $u_{j'_1} \equiv u_{j_1}$, $u_{j'_2} \equiv u_{j_2}, \ldots, u_{j'_l} \equiv u_{j_l}$,

then $t_{i',j'} = f(u_{j'_1}, \ldots, u_{j'_l})$ for some i' and j'.

If $\equiv$ is a congruence on S, a *quotient system* $S' = S/{\equiv}$ can be defined as follows. We let $\{v_1, \ldots, v_m\}$ be a new set of unknowns in bijection with $U/{\equiv}$ (we denote by $[u]$ the equivalence class of u in U and with a slight abuse of notation $v = [u]$ if v corresponds to $[u]$).

Then we take $S' = \langle u = p', v_1 = q_1, \ldots, v_m = q_m \rangle$ with

$$p' = [u_{i_1}] + \ldots + [u_{i_k}],$$
$$q_i = \sum \{[t_{j,j'}] / u_j \in v_i,\ j' = 1, \ldots, n_j\},$$

where $[f(u_{j_1}, \ldots, u_{j_l})]$ denotes $f([u_{j_1}], \ldots, [u_{j_l}])$.

Of course, the idempotency law for $+$ can be used to reduce the polynomials $p', q_1, \ldots, q_m$.

5. Example

Let S be the system

$$u = u_1 + u_2 + u_3,$$
$$u_1 = f(u_2) + g(u_1, u_2) + g(u_2, u_1) + g(u_2, u_2) + f(u_3),$$

$$u_2 = f(u_1)+g(u_1, u_1)+a,$$

$$u_3 = g(u_3, u_1)+g(u_3, u_2)+b.$$

This system is quasi-deterministic and $(\{u_1, u_2\}, \{u_3\})$ is a congruence on S. The quotient system is the following system S' below with v_1 representing $\{u_1, u_2\}$ and v_2 representing $\{u_3\}$:

$$u = v_1 + v_2,$$

$$v_1 = f(v_1)+g(v_1, v_1)+f(v_2)+a,$$

$$v_2 = g(v_2, v_1)+b. \quad ■$$

6. Main result

THEOREM. *Let S be a quasi-deterministic system, let $\equiv$ be a congruence on S, let $S' = S/\equiv$. Then there exist two quasi-deterministic systems S_1 and S'_1 such that*

$$S \subseteq S_1 \textbf{ redef } S'_1 \supseteq S'.$$

Proof. Let $S = \langle u = p, u_1 = p_1, \ldots, u_n = p_n\rangle$, let $S' = \langle u = p', v_1 = q_1, \ldots, v_m = q_m\rangle$ with the notation of Definition 4.

For every $i = 1, \ldots, m$, let

$$r_i = \sum \{p_j/u_j \in v_i\},$$

and finally

$$S_1 = \langle u = p, v_1 = r_1, \ldots, v_m = r_m, u_1 = p_1, \ldots, u_n = p_n\rangle,$$

$$S'_1 = \langle u = p', v_1 = q_1, \ldots, v_m = q_m, u_1 = p_1, \ldots, u_n = p_n\rangle.$$

We now prove that S_1 **redef** S'_1 by a redefinition of $u, v_1, \ldots, v_m$.

In the following claim we denote by $\sum v$ the polynomial $u_{j_1} + \ldots + u_{j_l}$ if $v = \{u_{j_1}, \ldots, u_{j_l}\}$.

CLAIM. *For all $i = 1, \ldots, m$, $q_i[\sum v_1/v_1, \ldots, \sum v_m/v_m] \overset{*}{\underset{\mathcal{D}}{\longleftrightarrow}} r_i$.*

Proof of the claim. From the definitions, $r_i = \sum \{t_{j,j'}/u_j \in v_i,\ j' = 1, \ldots, n_j\}$ and $f(v_{j_1}, \ldots, v_{j_l})$ is a subterm of q_i if and only if there exist $j, j', j'_1, \ldots, j'_l$ such that $u_j \in v_i$, $j' \in \{1, \ldots, n_j\}$ and $t_{j,j'} = f(u_{j'_1}, \ldots, u_{j'_l})$ with $u_{j'_i} \in v_{j_i}$ for all $i = 1, \ldots, l$.

Hence

$$q_i[\sum v_1/v_1, \ldots] \overset{*}{\underset{\mathcal{D}}{\longleftrightarrow}} r_i. \quad ■$$

To establish the theorem we now verify conditions (1)–(4) of Definition 2.

Conditions (1) and (2) are clear from the definitions.

Condition (3) holds since clearly

$$p[\Omega/u_1, \ldots, \Omega/u_n] \overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} \Omega,$$

and by the claim

$$r_i[\Omega/u_1, \ldots, \Omega/u_n] \overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} q_i[\Omega/v_1, \ldots, \Omega/v_m].$$

For condition (4), we have to prove

$$p[p_1/u_1, \ldots, p_n/u_n] \overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} p'[r_1/v_1, \ldots, r_m/v_m].$$

But $p \overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} p'\ [\sum v_1/v_1, \ldots, \sum v_m/v_m]$ by the definition of p'. On the other hand, $r_j \overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} (\sum v_j)[p_1/u_1, \ldots, p_n/u_n]$ by the definition of r_j. Hence

$$\begin{aligned} p[p_1/u_1, \ldots, p_n/u_n] &\overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} p'[\textstyle\sum v_1/v_1, \ldots][p_1/u_1, \ldots] \\ &= p'[(\textstyle\sum v_1)[p_1/u_1, \ldots]/v_1, \ldots] \\ &\overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} p'[r_1/v_1, \ldots, r_m/v_m]. \end{aligned}$$

We must also prove that, for $i = 1, \ldots, m$,

$$r_i[p_1/u_1, \ldots, p_n/u_n] \overset{*}{\underset{\mathcal{G}}{\leftrightarrow}} q_i[r_1/v_1, \ldots, r_m/v_m]$$

but this follows immediately from the claim with the help of similar computations.

This completes the proof of the theorem. ∎

COROLLARY 1. *Let M be any F-magma. The components corresponding to u of the least solutions of S and S' in $P(M)$ are equal.*

Proof. It follows, from [3], Proposition (5.12), that the least solutions of S_1 and S'_1 are the same. Since S is a subsystem of S_1, its least solution is the appropriate projection of the least solution of S_1 and similarly for S'_1. It follows that the components corresponding to u of the least solutions of S and S' are the same. ∎

Brainerd [1] has characterized the minimal tree-automaton defining a given regular tree language L as some quotient of any deterministic tree-automaton defining L. Since the polynomial system associated with a quotient-automaton is a quotient of the quasi-deterministic system associated with the given automaton, we have

COROLLARY 2. *The minimalization of tree-automata can be expressed in terms of inclusion and redefinition of polynomial systems.*

7. Example (continuation of Example 5)

Let S'' be the set of equations of S with left-hand side u_1, u_2, u_3; we take for S_1 the system S'' together with the equations

$$u = u_1 + u_2 + u_3,$$

$$v_1 = f(u_1) + f(u_2) + g(u_1, u_1) + g(u_1, u_2) + g(u_2, u_1) + g(u_2, u_2) + f(u_3) + a,$$

$$v_2 = g(u_3, u_1) + g(u_3, u_2) + b.$$

Note that these equations can be factorized as follows:

$$u = (u_1 + u_2) + u_3,$$

$$v_1 = f(u_1 + u_2) + g((u_1 + u_2), (u_1 + u_2)) + f(u_3) + a,$$

$$v_2 = g(u_3, u_1 + u_2) + b.$$

We take for S'_1 the system S'' together with the equations

$$u = v_1 + v_2,$$

$$v_1 = f(v_1) + g(v_1, v_1) + f(v_2) + a,$$

$$v_2 = g(v_2, v_1) + b.$$

To establish that S_1 **redef** S'_1, one must verify in particular condition (4) of Definition 2.

Let us consider in particular the verification that

$$q_1[r_1/v_1, r_2/v_2] \overset{*}{\underset{\mathcal{G}}{\longleftrightarrow}} r_1[p_1/u_1, p_2/u_2, p_3/u_3]$$

where p_i, q_i, r_i are as in the proof of the theorem.

Our factorization of r_1 shows that

$$q_1[u_1 + u_2/v_1, u_3/v_2] \overset{*}{\underset{\mathcal{G}}{\longleftrightarrow}} r_1$$

which in the proof of the theorem follows from the claim. Since the general proof is easier to follow than the example, we do not detail the computation. ∎

References

[1] W. Brainerd, *The minimalization of tree-automata*, Information and Control 13 (1968), 484–491.

[2] B. Courcelle, *Fundamental properties of infinite trees*, Theor. Comput. Sci. 25 (1983), 95–169.

[3] —, *Equivalence and transformations of regular systems; applications to recursive program schemes and grammars*, Theor. Comput. Sci. 42 (1986), 1–122.

[4] F. Gecseg, M. Steinby, *Tree-automata*, Akademiai Kiadó, Budapest 1984.

Presented to the semester
Mathematical Problems in Computation Theory
September 16 – December 14, 1985