Finite Automata

over

Conway Semirings

Werner Kuich

Technische Universität Wien

Conway semirings, matrices and formal power series.

Advantages:

(i) Constructions needed in the proofs are mainly the usual ones.

(ii) Proofs are separated from the constructions and do not need the intuitive content of the constructions.

(iii) Proofs are more satisfactory from the mathematical point of view.

(iv) Results are more general than the usual ones.

Conway semirings: defined by

sum – star – equation and product – star – equation.

Proofs can be separated into two parts:

(i) establish the needed results of the theory of Conway semirings,

(ii) simple equational reasoning.

Leads to a transparent structure of the proofs.

Semiring: <S,+,·,0,1> or simply S.

(i) <S,+,0> is a commutative monoid,

(ii) <S,·,1> is a monoid,

(iii) the distribution laws $a\cdot(b+c) = a\cdot b+a\cdot c$ and $(a+b)\cdot c = a\cdot c+b\cdot c$ hold for every a,b,c,

(iv) $0\cdot a = a\cdot 0 = 0$ for every a.

In the sequel, S denotes a semiring and A a finite alphabet.

Commutative: if ab = ba for every a and b.

Starsemiring: additional unary operation *.

Complete semiring: infinite sums are defined.

Complete starsemiring: complete as a semiring and, for each element a

$$a^* = \sum_{k\geq 0} a^k$$

Examples of complete starsemirings are:

Boolean semiring **B** = <{0,1},+,·,0,1> with
$1 + 1 = 1$

Nonnegative numbers with ∞:
$N^{\infty} = \langle N \cup \{\infty\}, +, \cdot, 0, 1 \rangle$ with $0 \cdot \infty = \infty \cdot 0 = 0$, $0^* = 1$ and $a^* = \infty$ for all $a \neq 0$.

The semiring of **Formal Languages** over A.

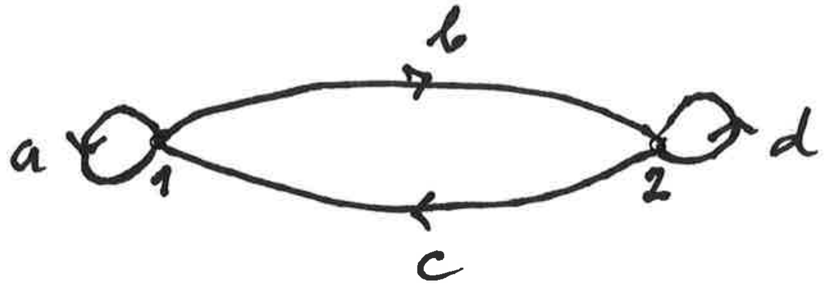
The semiring of **Formal Power Series** over a commutative semiring S and A: S<<A*>>;
B<<A*>> is isomorph to the semiring of Formal Languages over A.

The **tropical** semirings.

The semiring of **binary relations**.

In the sequel, if graphs are considered, we assume that the basic semiring of the inscriptions of the edges is a complete starsemiring.

Example.



with adjacency matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$(M^k)_{ij}$ is the language of inscriptions of paths of length k from i to j.

M^* defined by $(M^*)_{ij} = \sum_{k \geq 0} (M^k)_{ij}$
is the language
of inscriptions of all paths from i to j.

Inscriptions of paths from 1 to 1 **not** passing through 1:
a, bd^nc, $n \geq 0$.

Language of these inscriptions: $a + bd^*c$.

Language of inscriptions from 1 to 1:
$(M^*)_{11} = (a + bd^*c)^*$.

Language of inscriptions from 1 to 2:
$(M^*)_{12} = (a + bd^*c)^*bd^*$.

The (2,1) and (2,2) entries of M* are given by symmetry.

This yields

$$M^* = \begin{pmatrix} (a + bd^*c)^* & (a + bd^*c)^*bd^* \\ (d + ca^*b)^*ca^* & (d + ca^*b)^* \end{pmatrix}$$

The semiring of **square matrices of dimension n:**
$<S^{nxn}, +, \cdot, 0, E>$.

For a matrix M of dimension n, M* is inductively defined as follows:

(i) For n = 1 and M = (a),

$M^* = (a^*)$.

(ii) For n >1 and

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$M^* = \begin{pmatrix} (a + bd^*c)^* & (a + bd^*c)^*bd^* \\ (d + ca^*b)^*ca^* & (d + ca^*b)^* \end{pmatrix}$$

where

a	1x1	b	1x(n – 1)
c	(n – 1)x1	d	(n – 1)x(n – 1)

Given a starsemiring, the star of a square matrix is always defined in this manner.

Three equations for starsemirings, important in automata theory:

(1) The **sum – star – equation** is valid in S if

$$(a + b)^* = (a^*b)^*a^* \quad \text{for all } a,b;$$

(2) The **product – star – equation** is valid in S if

$$(ab)^* = 1 + a(ba)^*b \quad \text{for all } a,b;$$

(3) Let M and M* be given as in the definition of the star of M, but with

a	$n_1 \times n_1$	b	$n_1 \times n_2$
c	$n_2 \times n_1$	d	$n_2 \times n_2$

where $n_1 + n_2 = n$.

The **matrix – star – equation** is valid in S if the computation of M* is independent of the partition of n into summands n_1, n_2.

Conway semiring: starsemiring satisfying the sum – star equation and the product – star – equation.

Theorem (Conway). If S is a Conway semiring then the matrix semirings $S^{n\times n}$ are Conway semirings. Moreover, the matrix – star – equation is valid for Conway semirings.

The matrix – star equation implies

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}^* = \begin{pmatrix} a^* & a^*bd^* \\ 0 & d^* \end{pmatrix}$$

One can remember the (1,2) – block by

A complete starsemiring is a Conway semiring.

If the Conway semiring S is a complete semiring, it can be proven that

$$M^* = \sum_{k \geq 0} M^k$$

i.e., the starsemiring of nxn – matrices over S is a complete starsemiring.

In the sequel, S denotes a Conway semiring, and S′ denotes a subset of S containing 0 and 1.

A **finite S′ – automaton**

$$\mathfrak{A} = (n,M,I,P)$$

is given by

(i) the set of states $\{1, \ldots, n\}$, $n \geq 1$,
(ii) a **transition matrix** $M \in S'^{nxn}$,
(iii) an **initial state vector** $I \in S'^{1xn}$,
(iv) a **final state vector** $P \in S'^{nx1}$.

Behavior $||\mathfrak{A}||$ of $\mathfrak{A}$:

$$||\mathfrak{A}|| = \sum_{1\le i,j\le n} I_i(M^*)_{ij}P_j = IM^*P.$$

Directed labeled graph of $\mathfrak{A}$:

nodes: 1, … , n,
edges: from i to j if $M_{ij} \neq 0$ and labeled by M_{ij}
initial nodes: i if $I_i \neq 0$ with weight I_i
final nodes: j if $P_j \neq 0$ with weight P_j

Path

$M_{i,j}$ M_{jk} M_{st}
o----->o----->o ....o----->o
i j k s t

has **weight**

$$M_{ij}\, M_{jk}\, \ldots\ldots\, M_{st}$$

$(M^k)_{ij}$ sum of the weights of paths of lenght k from i to j

If S is a complete semiring, $(M^*)_{ij} = \sum_{k\ge 0} (M^k)_{ij}$ is the sum of the weights of paths from i to j.

Normalized finite S′ - automaton $\mathfrak{A} = (n,M,I,P)$, $n \geq 2$,

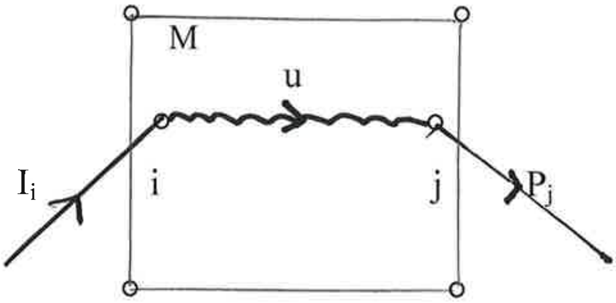
(i) $I_1 = 1$, $I_i = 0$, $2 \leq i \leq n$;
(ii) $P_n = 1$, $P_j = 0$, $1 \leq j \leq n - 1$;
(iii) $M_{i1} = M_{nj} = 0$, $1 \leq i,j \leq n$.

The finite automata $\mathfrak{A}$ and $\mathfrak{A}'$ are **equivalent** if $||\mathfrak{A}|| = ||\mathfrak{A}'||$.

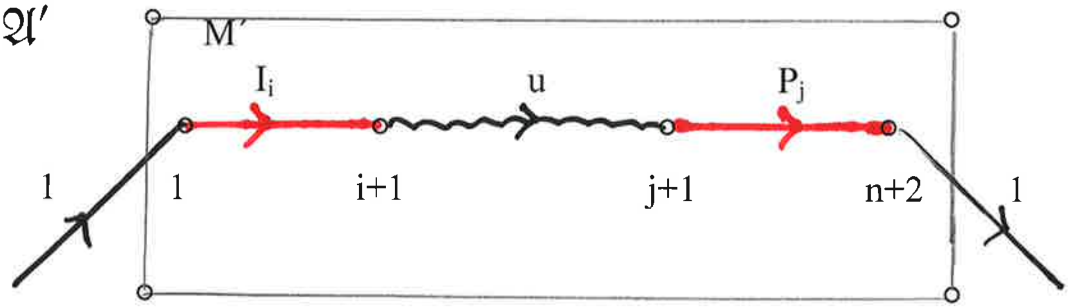
Theorem. Each finite S′ - automaton is equivalent to a normalized finite S′ - automaton.

Proof. Let $\mathfrak{A} = (n,M,I,P)$, $\mathfrak{A}' = (1 + n + 1, M', I', P')$.

$\mathfrak{A}'$



$\mathfrak{A}'$



$$M' = \begin{pmatrix} 0 & I & 0 \\ 0 & M & P \\ 0 & 0 & 0 \end{pmatrix}, \quad I' = (1\ 0\ 0), \quad P' = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\|\mathfrak{A}'\| = I'M'^*P' = (M'^*)_{1,n+2}$$

$$\left(\begin{array}{cc|c} 0 & I & 0 \\ 0 & M & P \\ \hline 0 & 0 & 0 \end{array}\right)^* [1,2]_1 = \left(\begin{pmatrix} 0 & I \\ 0 & M \end{pmatrix}^* \begin{pmatrix} 0 \\ P \end{pmatrix} 0^*\right)_1 =$$

$$\left(\begin{pmatrix} E & EIM^* \\ 0 & M^* \end{pmatrix} \begin{pmatrix} 0 \\ P \end{pmatrix}\right)_1 = \begin{pmatrix} IM^*P \\ M^*P \end{pmatrix}_1 = IM^*P = \|\mathfrak{A}\|.$$

Rat(S′)	substarsemiring generated by S′, i.e., smallest starsemiring containing S′.
Rec(S′)	collection of all behaviors of finite S′ - automata.

Theorem. Let S be a Conway semiring and S′ be a subset of S containing 0,1. Then

$$\mathrm{Rat}(S') = \mathrm{Rec}(S')$$

Proof.

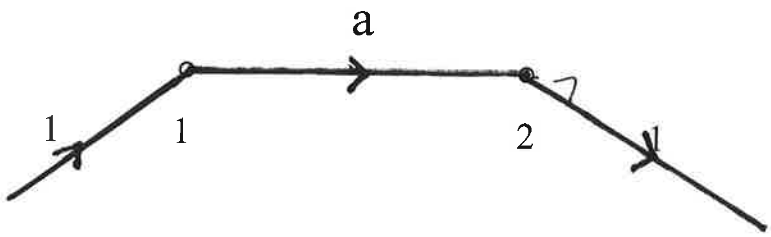
(i) $\mathrm{Rec}(S') \subseteq \mathrm{Rat}(S')$.

By induction $M^* \in (\mathrm{Rat}(S'))^{n\mathrm{x}n}$, for $\mathfrak{A} = (n,M,I,P)$,

$$\|\mathfrak{A}\| = IM^*P \in \mathrm{Rat}(S').$$

(ii) Rat(S′) ⊆ Rec(S′).

For a ϵ S′,



$\mathfrak{A}$ = (2,(a),(1),(1)) with $||\mathfrak{A}||$ = a, proving a ϵ Rec(S′) and S′ ⊆ Rec(S′).

Given finite S′ - automata

$$\mathfrak{A} = (n,M,I,P) \text{ and } \mathfrak{A}' = (n',M',I',P'),$$

we define S′- finite automata

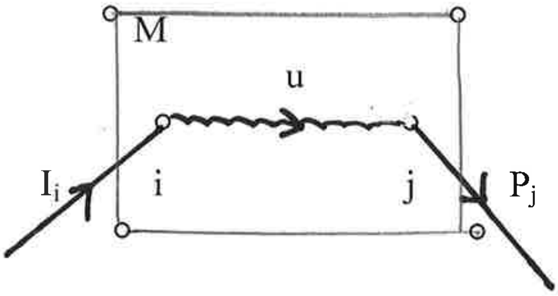
$$\mathfrak{A} + \mathfrak{A}', \mathfrak{A}\,\mathfrak{A}' \text{ and } \mathfrak{A}^*$$

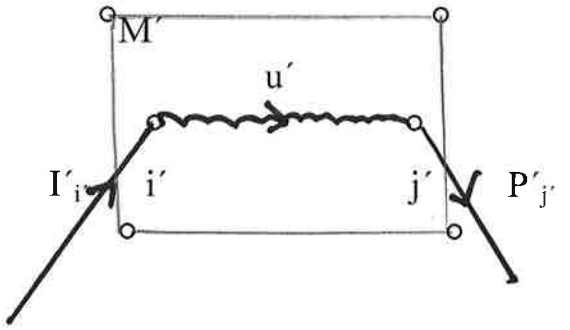
such that

$$||\mathfrak{A} + \mathfrak{A}'|| = ||\mathfrak{A}|| + ||\mathfrak{A}'||, ||\mathfrak{A}\mathfrak{A}'|| = ||\mathfrak{A}||||\mathfrak{A}'||, ||\mathfrak{A}^*|| = ||\mathfrak{A}||^*.$$

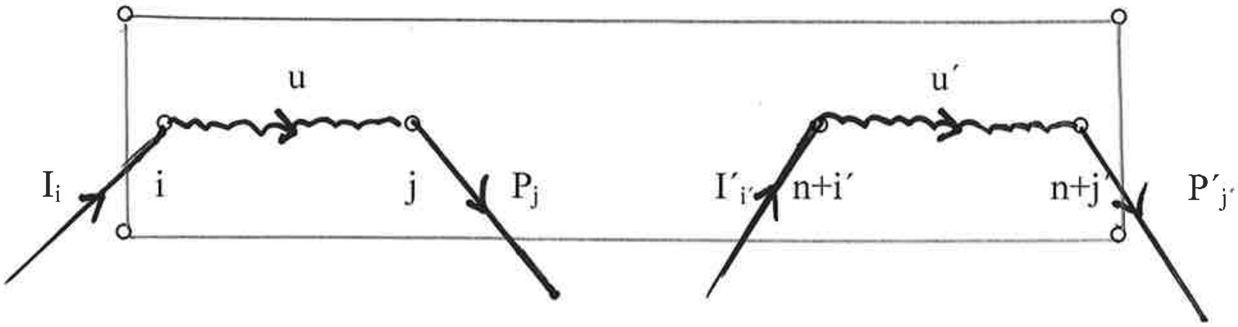
Construction of $\mathfrak{A} + \mathfrak{A}'$.

$\mathfrak{A}$ $\qquad\qquad\qquad\qquad$ $\mathfrak{A}'$





$\mathfrak{A} + \mathfrak{A}'$



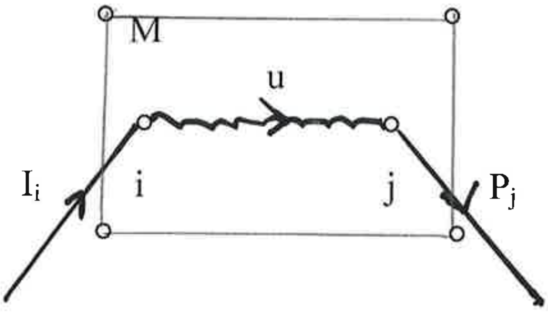
$$\mathfrak{A} + \mathfrak{A}' = \left(n + n', \begin{pmatrix} M & 0 \\ 0 & M' \end{pmatrix}, (I \;\; I'), \begin{pmatrix} P \\ P' \end{pmatrix}\right)$$

$$\|\mathfrak{A} + \mathfrak{A}'\| = (I \;\; I')\begin{pmatrix} M & 0 \\ 0 & M' \end{pmatrix}^{*}\begin{pmatrix} P \\ P' \end{pmatrix}) = (I \;\; I')\begin{pmatrix} M^{*} & 0 \\ 0 & M'^{*} \end{pmatrix}\begin{pmatrix} P \\ P' \end{pmatrix}) =$$

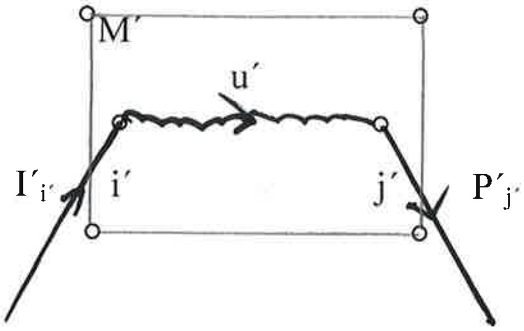
$$IM^{*}P + I'M'^{*}P' = \|\mathfrak{A}\| + \|\mathfrak{A}'\|.$$

Construction of $\mathfrak{A}\mathfrak{A}'$.

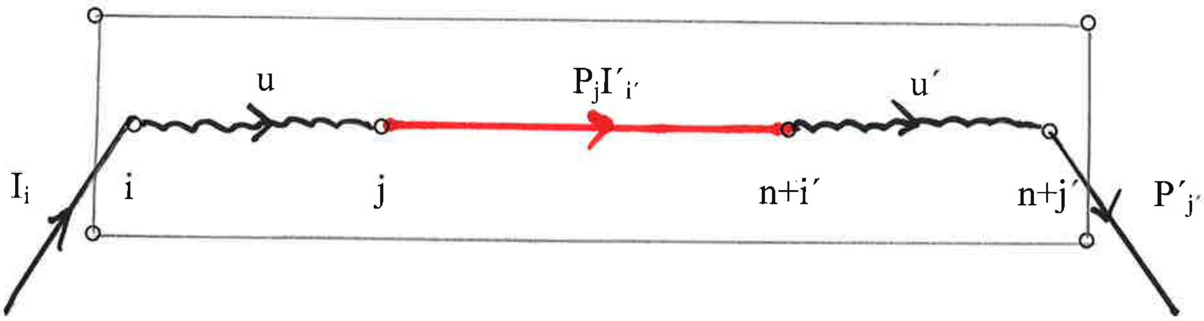
$\mathfrak{A}$



$\mathfrak{A}'$



$\mathfrak{A}\,\mathfrak{A}'$



$$\mathfrak{A}\mathfrak{A}' = \left(n + n',\ \begin{pmatrix} M & PI' \\ 0 & M' \end{pmatrix},\ (I\ \ 0),\ \begin{pmatrix} 0 \\ P' \end{pmatrix}\right)$$

Assume that $\mathfrak{A}$ or $\mathfrak{A}'$ are normalized. Then the entries of PI' are in S'. Hence, $\mathfrak{A}\mathfrak{A}'$ is a finite S' - automaton.

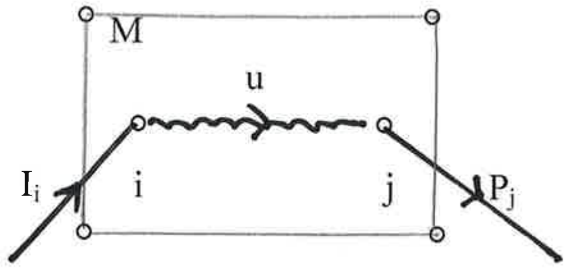
$$\|\mathfrak{A}\mathfrak{A}'\| = (I\ \ 0)\begin{pmatrix} M & PI' \\ 0 & M' \end{pmatrix}^{*}\begin{pmatrix} 0 \\ P' \end{pmatrix} =$$

$$(I\ \ 0)\begin{pmatrix} M^{*} & M^{*}PI'M'^{*} \\ 0 & M'^{*} \end{pmatrix}\begin{pmatrix} 0 \\ P' \end{pmatrix} = (IM^{*}, IM^{*}PI'M'^{*})\begin{pmatrix} 0 \\ P' \end{pmatrix} =$$

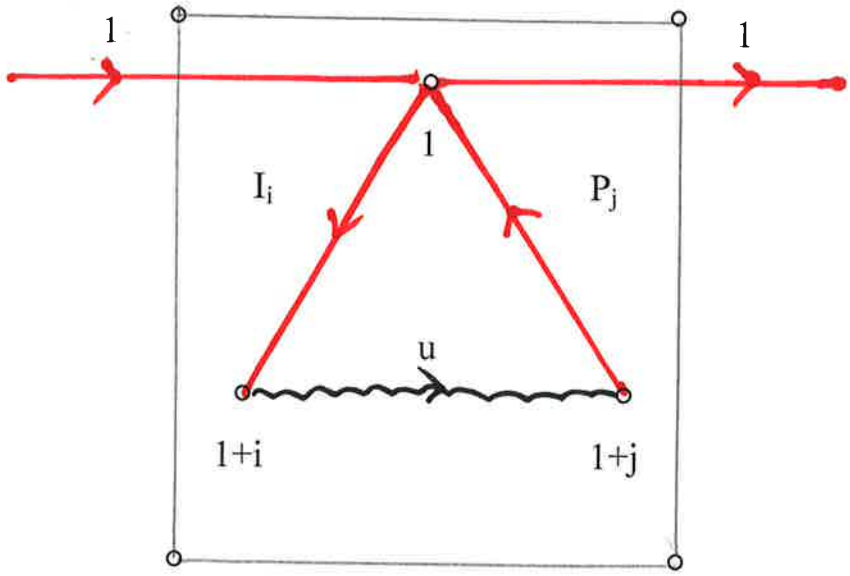
$$IM^{*}PI'M'^{*}P' = \|\mathfrak{A}\|\|\mathfrak{A}'\|.$$

Construction of $\mathfrak{A}^*$.

$\mathfrak{A}$



$\mathfrak{A}^*$



$$\mathfrak{A}^* = \left(1 + n, \begin{pmatrix} 0 & I \\ P & M \end{pmatrix}, (1 \quad 0), \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$$

$$\|\mathfrak{A}^*\| = (1 \quad 0) \begin{pmatrix} 0 & I \\ P & M \end{pmatrix}^* \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \left(\begin{pmatrix} 0 & I \\ P & M \end{pmatrix}^* \right)_{11} =$$

$$(0 + IM^*P)^* = \|\mathfrak{A}\|^*$$

$S' \in Rec(S')$ and $Rec(S')$ is closed under the operations $+, \cdot, *$, i.e., $Rec(S')$ is a starsemiring containing S'.

Since $Rat(S')$ is the smallest starsemiring containing S'

$$Rat(S') \subseteq Rec(S').$$

Formal power series over a finite alphabet A:

r: $A^* \to S$, $r(w) = (r,w)$ coefficient of w,
written as formal sum $r = \sum_{w \in A^*} (r,w)w$.

$r_1, r_2, r \in S\langle\langle A^* \rangle\rangle$:

$r_1 + r_2$ with $(r_1 + r_2, w) = (r_1, w) + (r_2, w)$

$r_1 \circ r_2$ with $(r_1 \circ r_2, w) = \sum_{uv=w} (r_1, u)(r_2, v)$

r^* with $(r^*, \varepsilon) = (r, \varepsilon)^*$,

$(r^*, w) = (r, \varepsilon)^* \sum_{uv=w, u \neq \varepsilon} (r,u)(r^*, v)$, $w \neq \varepsilon$.

Given a starsemiring, the star of a formal power series is always defined in this manner.

Theorem (Bloom, Esik). If S is a Conway semiring then the semiring of formal power series $S\langle\langle A^* \rangle\rangle$ is a Conway semiring.

If the Conway semiring S is a complete semiring, it can be proven that

$$r^* = \sum_{k \geq 0} r^k$$

i.e., the starsemiring of formal power series over A is a complete starsemiring.

Notation:

$S\langle A\cup\varepsilon\rangle$ ..... $(r,\varepsilon)\varepsilon + \sum_{x\in A}(r,x)x.$

$S\langle A\rangle$ ....$\sum_{x\in A}(r,x)x.$

$S\langle\varepsilon\rangle$ .... $(r,\varepsilon)\varepsilon.$

A finite $S\langle A\cup\varepsilon\rangle$ - automaton $\mathfrak{A} = (n,M,I,P)$ is called **standard** finite $S\langle A\cup\varepsilon\rangle$ - automaton if

(i) $M \in (S\langle A\rangle)^{nxn}$,
(ii) $I_1 = \varepsilon$, $I_i = 0$, $2 \leq i \leq n$,
(iii) $P_j \in S\langle\varepsilon\rangle$, $1 \leq j \leq n$.

$Rec_{st}(S\langle A\cup\varepsilon\rangle) =$
$\{ \|\mathfrak{A}\| /\ \mathfrak{A}$ is a standard finite $S\langle A\cup\varepsilon\rangle$ - automaton$\}$.

Theorem. Let S be a Conway semiring.

Then $\mathrm{Rec}(S\langle A\cup\varepsilon\rangle) = \mathrm{Rec}_{st}(S\langle A\cup\varepsilon\rangle)$.

Proof (Esik, K.). Let $\mathfrak{A} = (n,M,I,P)$ with

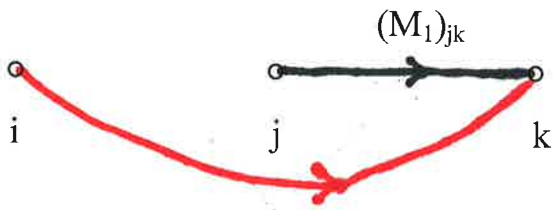
(i) $M \in (S\langle A\cup\varepsilon\rangle)^{nxn}$,
(ii) $I_1 = \varepsilon$, $I_i = 0$, $2 \leq i \leq n$,
(iii) $P_j \in S\langle\varepsilon\rangle$, $1 \leq j \leq n$.

Partition M into ε – transitions and non ε – transitions:

$$M = M_0 + M_1,\ M_0 = (M,\varepsilon)\varepsilon,\ M_1 = \sum_{x\in A}(M,x)x.$$

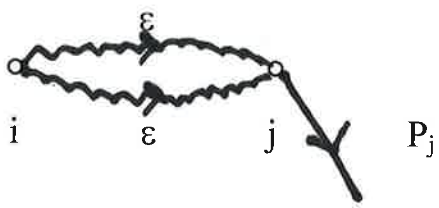
$\mathfrak{A}$

$\mathfrak{A}'$

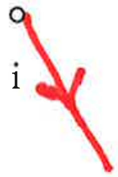


$$\sum_{1\leq j\leq n} (M_0^*)_{ij}\,(M_1)_{jk} = (M_0^*M_1)_{ik}$$

$\mathfrak{A}$



$\mathfrak{A}'$



$$\sum_{1\leq j\leq n} (M_0^*)_{ij}\,P_j = (M_0^*P)_i$$

Let $\mathfrak{A}' = (n, M_0^*M_1, I, M_0^*P)$. Then

$$||\mathfrak{A}'|| = [I][(M_0^*M_1)^*][M_0^*P] = [I][(M_0^*M_1)^*M_0^*][P] =$$

$$I(M_0 + M_1)^*P = IM^*P = ||\mathfrak{A}||.$$

Corollary (Schützenberger).
$\mathrm{Rat}(S\langle A\cup\varepsilon\rangle) = \mathrm{Rec}_{st}(S\langle A\cup\varepsilon\rangle)$.

Corollary (Kleene).
$\mathrm{Rec}(\mathbf{B}\langle A\rangle) = \mathrm{Rec}_{st}(\mathbf{B}\langle A\cup\varepsilon\rangle)$.