Saturated Automata applied to the Star Height Problem

Lill Kristiansen
University of Oslo, Norway

Abstract:

The saturated automaton, Sat(R), contains homomorphic images of all automata accepting R.
We study the behavior of homomorphisms into Sat(R), in particular we prove the following theorem:

The star height of R is greater than or equal to the rank of Core(R); where Core(R) is the intersection of the minimal forward and backward deterministic automata accepting R.

This result gives the exact star height of regular events with the finite intersection property.
Sat(R) also gives an upper bound on the star height of R:
The star height of R is less than or equal to the minimum of the rank of subautomata of Sat(R) accepting R.

We end the paper by giving several examples where this upper bound is exact.

It is an open question whether this bound always is exact.

1. Introduction

One of the main unsolved problems in the theory of regular events is the star height problem.

The star height of a regular expression is the depth/height of the nesting of the $\star$-operator. The star height $h(R)$ of a regular event R, is then the minimum of the star heights of expressions denoting R.

The problem is: Given R, can the star height of R always be found?

In Eggans classical paper [7], only restricted regular expressions (with operations $\cdot, \vee, \star$) were allowed. Since the star is the most powerful of these operators, the star height is a good measure of the complexity.

Later also general star height has been considered, where we in addition allow $\neg$ (negation) and $\wedge$ (intersection) as operators [16], [17].

We will consentrate on the restricted star height. Eggan showed that with a suitable definition of the rank ("loop complexity") of an automaton $\mathcal{A}$ we get:

$$h(R) = \min\{\text{rank } (\mathcal{A}) \mid \mathcal{A} \text{ is a nondeterministic automaton accepting } R\}$$

Many authors have therefore studied the rank of an automaton. (McNaughton [13], [15], Cohen [5], [6], Hashigushi and Honda [9]) (We will also mention that Hashiguchi [8] with quite different methods has shown that it is possible given an event of star height ≤ 2, to determine if the event has star height < 2. It is not known how to generalize this to an arbitrary star height.)

A central result in the determination of the star height of various special classes of regular events is the

McNaughton's pathwise homomorphism theorem [13]: Given a homomorphism f from an automaton $\mathcal{A}'$ onto an automaton $\mathcal{A}$. If f is pathwise (i.e. onto the paths in $\mathcal{A}$), then rank ($\mathcal{A}'$) $\geqslant$ rank ($\mathcal{A}$).

Thus if $\mathcal{A}'$ and $\mathcal{A}$ are automata accepting R, $f: \mathcal{A}' \to \mathcal{A}$ is pathwise and $\mathcal{A}'$ is of minimal rank (i.e. h(R) = rank($\mathcal{A}'$)), then $\mathcal{A}$ is also of minimal rank.

This is one of the reasons for introducing the notion of a saturated automaton. We say that $\mathcal{A}$ is saturated with respect to R if

1) $\mathcal{A}$ accepts R
2) For all automata $\mathcal{A}'$ accepting R, there exists a homomorphism from $\mathcal{A}'$ into $\mathcal{A}$.

The basic properties of saturated automata (existence and uniqueness of a minimal saturated automaton Sat(R) with respect to R) was established in [12]. In this paper we consentrate on applications to the star height problem.

We obtain the following main result:

Theorem: h(R) $\geqslant$ rank(Core(R)), where Core(R) is the intersection of Det(R) and BDet(R) in Sat(R).
Since both the forward DA Det(R) and the backward DA BDet(R) can be considered as subautomata of Sat(R), this intersection makes sense.
The proof uses McNaughton's pathwise homomorphism theorem. In section 5 we also use this theorem to detemine some star heights, e.g. if R has the finite intersection property (f.i.p.), then h(R) = rank(Det(R)). (This was already obtained by Cohen [4]; the methods are, however, different.)

We end the paper by discussing the following problem: "Does it always exist a subautomaton $\mathcal{A}_m$ in Sat(R) (or in $\mathcal{A}_1$(R)) accepting R, and such that $\mathcal{A}_m$ gives the star height of R (i.e. h(R) = rank($\mathcal{A}_m$))?"
Cohen and Brzozowski [3] claims to have a solution. Our examples show that the problem is still open.

2. Preliminaries

A general non-deterministic automaton (NDA) will be written $\mathcal{A} = (Q,\Sigma,M,S,F)$ where

Σ is the (finite) input alphabet

Q is the (finite) set of states

$S \subseteq Q$ is the set of initial states

$F \subseteq Q$ is the set of final states

$M \subseteq Q \times (\Sigma \cup \{e\}) \times Q$ is the transition relation.

If $\mathcal{A}$ is deterministic, we use $\delta:Q\times\Sigma \to Q$ instead of M, and δ is extended in the usual way to a function from $Q\times\Sigma^\star \to Q$. Likewise M is regarded also as functions $M:Q\times(\Sigma\cup\{e\}) \to 2^Q$, $M:Q\times\Sigma^\star \to 2^Q$ and $M:2^Q\times 2^{\Sigma^\star} \to 2^Q$ (extended in the usual way).
The regular event accepted by $\mathcal{A}$ is written $T(\mathcal{A})$. We are also interested in the preceeding and succeeding event relative to a state q in $\mathcal{A}$:
$Pr^{\mathcal{A}}(q) = T(\mathcal{A},S,\{q\})$
$Sc^{\mathcal{A}}(q) = T(\mathcal{A},\{q\},F)$ where $T(\mathcal{A},S_0,F_0) = T((Q_{\mathcal{A}},\Sigma,M_{\mathcal{A}},S_0,F_0))$.

Definition 2.1:

q if a dead state iff $Sc^{\mathcal{A}}(q) = \emptyset$

q is an inacessible state iff $Pr^{\mathcal{A}}(q) = \emptyset$

$\mathcal{A}^-$ is the automaton $\mathcal{A}$ after removal of dead and inaccessible state and the transitions in connection with them.

Definition 2.2.

A semiautomaton (graph) $G = (Q,\Sigma,M)$ is an automaton without initial and final states. Given automata

$\mathcal{A}_i = (Q_i,\Sigma,M_i,S_i,F_i)$ $i=1,2$ and a mapping $f: Q_1 \to Q_2$.

a) f is a transition homomorphism from $G_1 = (Q_1,\Sigma,M_1)$ into $G_2 = (Q_2,\Sigma,M_2)$ iff
$(q,a,q') \in M_1 \Rightarrow$
$[(f(q),a,f(q')) \in M_2$, $a \in \Sigma\cup\{e\}$ or $(f(q) = f(q')$ and $a = e)]$.

b) f is an (automaton) homomorphism from $\mathcal{A}_1$ into $\mathcal{A}_2$ iff
1) f is a transition homomorphism from G_1 to G_2
2) $q \in S_1 \Rightarrow f(q) \in S_2$
3) $q \in F_1 \Rightarrow f(q) \in F_2$.

A transition homomorphism will then induce a mapping from paths in G_1 to paths in G_2 and V and $f(V)$ will span the same word. A homomorphism will transform an accepting path V for w in $\mathcal{A}_1$ into $f(V)$ which is an accepting path for w in $\mathcal{A}_2$.

Definition 2.3

If V is a path in $\mathcal{A}_1$ and f a homomorphism from $\mathcal{A}_1$ into $\mathcal{A}_2$ we write

$$V = (q^0,a^1,q^1,\dots,a^k,q^k) \qquad k \geq 0 \text{ where}$$

$$(q^{i-1},a^i,q^i) \in M_1 \quad \text{or} \quad (q^{i-1},a^i,q^i) = (q^i,e,q^i)$$

and $$f(V) = (f(q^0),a^1,f(q^1),\dots,a^k,f(q^k)).$$

The trivial transitions (q,e,q) may be inserted/deleted wherever q occurs in V or in $f(V)$; this does not change the paths.

We return to automata and homomorphism in section 3, and turn now to the basic notions concerning star height.

But first one more definition:

Definition 2.4

Given $\mathcal{A}_i = (Q_i \Sigma, M_i, S_i, F_i)$ and G_i, $i=1,2$.
$\mathcal{A}_1$ is a subautomaton of $\mathcal{A}_2$ (written $\mathcal{A}_1 \subseteq \mathcal{A}_2$) iff $Q_1 \subseteq Q_2$, $M_1 \subseteq M_2$, $S_1 \subseteq S_2$ and $F_1 \subseteq F_2$. Similar G_1 is a subgraph of G_2 iff $Q_1 \subseteq Q_2$, $M_1 \subseteq M_2$. We are mainly interested in the following subgraphs and subautomata for G and $\mathcal{A}$:

$f(\mathcal{A}_1)$: the image of $\mathcal{A}_1$ in $\mathcal{A}_2$ via the homomorphism f

$\mathcal{A} \upharpoonright Q_0$: the automaton $\mathcal{A}$ restricted to the states Q_0

$\mathcal{A} - (q,a,q')$: the automaton $\mathcal{A}$ without the transition (q,a,q')

$G-[Q_0]$: the graph G after removal of the states Q_0 and the corresponding transitions

The subgraph $G-[Q_0]$ will be used in the definition of the rank ("loop complexity") of a graph G.

Since we are concerned only with restricted regular expressions we have the following inductive definition of star height.

Definition 2.5:

The appearent star height h_α of an expression is defined by

$h_\alpha(\emptyset) = h_\alpha(e) = h_\alpha(a) = 0, \quad a \in \Sigma;$

$h_\alpha(E_1 \vee E_2) = h_\alpha(E_1 \cdot E_2) = \max(h_\alpha(E_1), h_\alpha(E_2));$

$h_\alpha(E^\star) = h_\alpha(E)+1.$

The star height h of a regular event R is defined as

$h(R) = \min\{h_\alpha(E) \mid E$ a regular expression denoting $R\}$.

The notion of (cycle) rank of an automaton was introduced to correspond to the star height, and we have the following theorem (from Eggan [7]):

For every regular expression E denoting R with $h_\alpha(E) = r$, there exists an automaton $\mathcal{A}$ accepting R with $\operatorname{rank}(\mathcal{A}) = r$, and vice versa.

Thus $h(R) = \min\{\operatorname{rank}(\mathcal{A}) \mid \mathcal{A}$ is a nondeterministic automaton accepting $R\}$.

We need the following notions:

Definition 2.6:

A graph $G = (Q,\Sigma,M)$ is strongly connected (s.c.) iff $\#Q \geq 1$ and for all q and q' in Q, there exists a path from q to q' (and from q' to q).

A maximal s.c. subgraph in called a section in G.

We are now able to state the inductive definition of the rank of a graph and an automaton.

Definition 2.7: Given $G = (Q,\Sigma,M)$ and $\mathcal{A} = (Q,\Sigma,M,S_0,F)$.

a) If G is s.c., then

rank(G) = 1 iff there exist a state q_0 in G such that $G - [q_0]$ is loopfree,

rank(G) = k>1 iff rank(G) is not less than k and there exists a state q_0 in G such that all sections in $G - [q_0]$ have rank at most $k-1$.

b) If G is not s.c., then

rank(G) = 0 iff G is loopfree,

$\text{rank}(G) = \max\{\text{rank}(G') \mid G' \text{ a section in } G\}$, otherwise.

c) The rank of $\mathcal{A}$ is defined as the rank of $(Q_{\mathcal{A}},\Sigma,M_{\mathcal{A}})$.

In some cases it is convenient to regard the loops (q,e,q) as transitions in $\mathcal{A}$ for all q, but this could increase the rank, and in this paper (q,e,q) is usually not allowed as a transition.

Note. McNaughton [13] defines homomorphism almost like Definition 2.2 a), but he permits $(q,a,q') \in M_1$ to be transformed to $f(q) = f(q')$ also when $a \neq e$. The V and $f(V)$ will not span the same words.

Our transition homomorphisms will be homomorphisms in his sense, and in a similar way we introduce the notion of pathwise homomorphism.

Definition 2.8:

A transition homomorphism f from G into G' is pathwise iff for all paths V' in G', there exists a path V in G such that $f(V) = V'$.

The following theorem will still be valid:

The pathwise homomorphism theorem [13]: If f is a pathwise transition homomorphism from G to G', then $\mathrm{rank}(G) \geq \mathrm{rank}(G')$.

Using this theorem Cohen has improved Eggans theorem by showing how e-transitions may be removed in an automaton without increasing the rank (and without increasing the number of states).

Theorem (Cohen [6])

$$h(R) = \min\{\mathrm{rank}(\mathcal{A}) \mid \mathcal{A} \text{ a NDA for } R \text{ without e-transitions}\}$$

3. The minimal saturated automaton Sat(R)

In [12] we introduced the notion of saturated automata:

Definition 3.1: An automaton $\mathcal{A}$ is called saturated iff

1) $\mathcal{A}$ accepts R
2) For all automata $\mathcal{A}'$ accepting (a subevent of) R, there exists a homomorphism from $\mathcal{A}'$ into $\mathcal{A}$.

We proved that a unique minimal saturated automaton, Sat(R), exists for any event R.

In this section we recall the construction and state some basic properties of Sat(R); for the proofs, see sections 2-5,7 of [12].

We assume the reader is familiar with the existence of a unique minimal deterministic automaton Det(R) with respect to R.

$$\mathrm{Det}(R) = (\hat{P},\Sigma,\delta_D,\{p_e\},F_D) \quad \text{where} \quad \hat{P} = \{p_e,p_2,\ldots,p_n\}$$

By the Nyhill-Nerode theorem (see e.g. [10]), the states corresponds to the equivalence classes ([v]) of the following relation ($\overset{D}{\sim}$):

$$v \overset{D}{\sim} w \quad \text{iff} \quad \forall u \in \Sigma^* \; (vu\in R \iff wu\in R).$$

Then $p_e = [e]$, $p_2 = [w_2], \ldots, p_n = [w_n]$, $F_D = \{p_j | [w_j] \subseteq R\}$ and $\delta_D(p_i,a) = p_j$ iff $[w_i a] = [w_j]$.

This gives $Pr^{\mathrm{Det}(R)}(p_i) = [w_i]$ and $Sc^{\mathrm{Det}(R)}(p_i) = w_i\backslash R$ where $w\backslash R$ is the derivative defined in [1] by $w\backslash R = \{u\in\Sigma^* | wu\in R\}$.

We may also define

$$R/w = \{u\in\Sigma^* | uw\in R\}$$

By regarding the dual (or transpose) event R^T we get $\mathrm{Det}(R^T) = \mathcal{A}$ and by taking the dual automaton $\overleftarrow{\mathcal{A}}$, we get the minimal backward deterministic automaton $\mathrm{BDet}(R) = (\mathrm{Det}(R^T))^{\leftarrow}$.

$\overleftarrow{\mathcal{A}}$ is defined as $(Q_{\mathcal{A}},\Sigma,\overleftarrow{M}_{\mathcal{A}},F_{\mathcal{A}},S_{\mathcal{A}})$ where $(q,a,q') \in \overleftarrow{M}_{\mathcal{A}}$ iff $(q',a,q) \in M_{\mathcal{A}}$.

BDet(R) might by duality be defined by

$$\mathrm{BDet}(R)^{\leftarrow} = (\hat{Q},\Sigma,\overleftarrow{\delta}_B,S,q_e) \quad \text{where each state} \quad q \in \hat{Q}$$

corresponds to <v>, the equivalence class of v under the relation $\overset{B}{\sim}$ defined by

$$v \overset{B}{\sim} w \quad \text{iff} \quad (\forall u \;\; uv \in R \iff uw \in R)$$

It can be shown that $w\backslash R$ is a union of equivalence classes $\langle v^1\rangle \cup \ldots \cup \langle v^k\rangle$, and similarily $R/w = [w^1] \cup \ldots \cup [w^\ell]$ (see [1]).

<u>Definition 3.2:</u> Given $P,Q \subseteq \Sigma^*$, $R \subseteq \Sigma^*$

(P,Q) is a <u>pair</u> relative to R iff $PQ \subseteq R$

(P,Q) is a <u>maximal pair</u> relative to R iff (P,Q) is a pair and neither component of (P,Q) can be extended preserving the property of (P,Q) being a pair.

It turns out that maximal pairs can be characterized by means of $[w]$ and $\langle v\rangle$.

<u>Definition 3.3:</u>

$$R::P = \{v \in \Sigma^* \mid P\{v\} \subseteq R\}$$

$$R:\,Q = \{w \in \Sigma^* \mid \{w\}q \subseteq R\}$$

$$\bar{P} = R:(R::P) \qquad \tilde{Q} = R::(R:Q)$$

<u>Note:</u> $R:Q = \bigcap_{v \in Q} R/v \qquad R::P = \bigcap_{w \in P} w\backslash R$

<u>Lemma 3.1:</u> The following are equivalent:

1) (P,Q) is a maximal pair.
2) (P,Q) is a pair, $R::P = Q$ and $R:Q = P$.
3) $P = \bar{P}$ and $Q = R::P$.
4) $P = R:Q$ and $Q = \tilde{Q}$.

Proof omitted.

<u>Proposition 3.2:</u> When R is regular there is only a finite number of maximal pairs. Whenever (P,Q) is a maximal pair, P consists

of a union of $\overset{D}{\sim}$ equivalence classes and Q consists of a union of $\overset{B}{\sim}$ equivalence classes.

Proof omitted.

In order to test whether $P = R:Q$, $Q = \tilde{Q}$ etc. when P (and Q) are unions of $\overset{D}{\sim}$ ($\overset{B}{\sim}$) equivalence classes, the following matrix is useful.

Definition 3.4: The reduced automaton matrix with respect to R (RAM(R)) contains one row for each $\overset{D}{\sim}$ equivalence class and one column for each $\overset{B}{\sim}$ equivalence class, and

$$\mathrm{RAM}([w],\langle v\rangle) = \begin{cases} + & \text{iff } [w]\langle v\rangle \subseteq R \quad (\text{iff } wv \in R) \\ - & \text{otherwise.} \end{cases}$$

Given $Q = \langle v^1\rangle \cup \ldots \langle v^k\rangle$ and $P = [w^1]\cup\ldots\cup[w^\ell]$, then

$R{:}\,Q = \{[w] \mid \forall j = 1,\ldots k\ \mathrm{RAM}([w],\langle v^j\rangle) = +\}$

$R{::}P = \{\langle v\rangle \mid \forall j = 1,\ldots \ell\ \mathrm{RAM}([w^j],\langle v\rangle) = +\}$.

Each maximal pair may then, by the one-to-one correspondence between equivalence classes and states in Det(R)/BDet(R), be identified with (P,Q) where $P \subseteq \hat{P}$, $Q \subseteq \hat{Q}$.

The rows and columns in RAM may equally well be indexed by $p_e, p_2, \ldots p_n, q_e, q_2, \ldots, q_m$ and

$$\mathrm{RAM}(p_i,q_j) = \begin{cases} + & \text{iff } \mathrm{Pr}(p_i)\mathrm{Sc}(q_j) \subseteq R \\ - & \text{otherwise.} \end{cases}$$

Instead of $R{::}[w_i]) = R/w_i = \langle v^1\rangle\cup\ldots\cup\langle v^k\rangle$ we will then write $R/p_i = \{q^1,\ldots,q^k\}$, and instead of $\{\overline{w_i}\}$ (or $[\overline{w_i}]$ or $\overline{w_i}$) we will write $\bar{p}_i$ etc.

Ex. 3.1 Consider the automaton

$$\mathcal{A}_n = (\{q_0,\ldots,q_{n-1}\},\{0,\ldots,n-1\},M_n,\{q_0\},\{q_{n-1}\})$$

where $(q_i,k,q_j) \in M_n$ iff $i+k \equiv j \bmod n$, $k = 0,\ldots n-1$.

Let $R_n = T(\mathcal{A}_n)$. It can be shown that $\mathcal{A}_n$ is both forward and backward deterministic and, in fact, $\mathcal{A}_n = \mathrm{Det}(R_n) = \mathrm{BDet}(R_n)$.

In R_3 we have the following equivalence classes:

q_0 corresponds to $[0] = [e]$ and to <2>

q_1 corresponds to $[1]$ and to <1>

q_2 corresponds to $[2]$ and to <0>

The RAM is shown in fig. 3.1 a).

The computation of maximal pairs are (by 3.1 and 3.2) done by computing $R::P$ and $R:(R::P)$ for P varying over unions of $[0]$, $[1]$, $[2]$, (or P varying over all subsets of $\hat{P}_3 = \{q_0, q_1, q_2\}$).

From RAM we get

$R/q_i = q_i \qquad q_i \backslash R = q_i \qquad i = 0,1,2.$

$R::\emptyset = \hat{Q}_3 \qquad R:\hat{Q}_3 = \emptyset$

$R::P = \emptyset \qquad R:\emptyset = \hat{P}_3 \qquad$ for any other $P \subseteq \hat{P}_3$

Thus the maximal pairs may be represented as:

(q_i, q_i) for r_i $\quad i = 0,1,2;$

$(\emptyset, \hat{Q}_3)$ for r_3

$(\hat{P}_3, \emptyset)$ for r_4

The maximal pairs are: $r_i = ([i], \langle 2-i \rangle), \quad i = 0,1,2$

$r_3 = (\emptyset , \Sigma^*);$

$r_4 = (\Sigma^* , \emptyset).$

We shall sometimes write $r_i = ([i], \langle 2-i \rangle) = (q_i, q_i)$; this should cause no confusion.

Fig. 3.1 a)

RAM(R_3)	q_0	q_1	q_2
	<2>	<1>	<0>
[0] q_0	+	-	-
[1] q_1	-	+	-
[2] q_2	-	-	+

Being now able to compute maximal pairs relative to R, we turn to the definition of the (minimal) saturated automaton Sat(R), with respect to a regular language R.

Definition 3.5. $\mathrm{Sat}(R) = (K_R, \Sigma, M_R, S_R, F_R)$ where

$$K_R = \{(P,Q) \mid (P,Q) \text{ is a maximal pair relative to } R\}$$
$$= \{(P_i,Q_i) \mid i=1,\ldots,N\},$$

$$((P_i,Q_i),a,(P_j,Q_j)) \in M_R \text{ iff } P_i\{a\} \subseteq P_j,\ a \in \Sigma \cup \{e\},$$

$$(P_i,Q_i) \in S_R \text{ iff } e \in P_i \quad (\text{iff } Q_i \subseteq R),$$
$$(P_i,Q_i) \in F_R \text{ iff } P_i \subseteq R \quad (\text{iff } e \in Q_i).$$

Theorem 3.3: Sat(R) is the (unique) minimal saturated automaton accepting R.

We will not prove this here, but only state some of the facts used in the proof.

Proposition 3.4:

1) $T(\mathrm{Sat}(R),\{(P_i,Q_i)\},\{(P_j,Q_j)\}) = \{u \in \Sigma^* \mid P_i\{u\}Q_j \subseteq R\}$.

2) $T(\mathrm{Sat}(R),\ S_R,\ \{(P_j,Q_j)\}) = \mathrm{Pr}^{\mathrm{Sat}(R)}(P_j,Q_j) = P_j$.

3) $T(\mathrm{Sat}(R),\{(P_i,Q_i)\},\ F_R) = \mathrm{Sc}^{\mathrm{Sat}(R)}(P_i,Q_i) = Q_i$.

Definition 3.6: Given $\mathcal{A} = (Q,\Sigma,M,S,F)$ define $f_i^{\mathcal{A}} : Q \to K_R$, $i=1,2$, by

$$f_1^{\mathcal{A}}(q) = (\bar{P}, R::P) \quad \text{where} \quad P = Pr^{\mathcal{A}}(q)$$

and

$$f_2^{\mathcal{A}}(q) = (R:Q,\tilde{Q}) \quad \text{where} \quad Q = Sc^{\mathcal{A}}(q)$$

In particular

$$f_1^{Det(R)}(p_i) = (\overline{[w_i]}, R::[w_i]) = (\bar{w}_i, w_i\backslash R)$$

$$f_2^{BDet(R)}(q_j) = (R:\langle v_j\rangle, \widetilde{\langle v_j\rangle}) = (R/v_j, \tilde{v}_j)$$

Instead of writing $(\bar{w}_i, w_i\backslash R)$ and $(R/v_j,\tilde{v}_j)$, we sometimes write $(\bar{p}_i, p_i\backslash R)$, $(R/q_j,\tilde{q}_j)$, respectively.

Proposition 3.5: When $T(\mathcal{A}) \subseteq R$, $f_1^{\mathcal{A}}$ and $f_2^{\mathcal{A}}$ are homomorphisms from $\mathcal{A}$ into $Sat(R)$.

Thus $f_1 = f_1^{Det(R)}$ and $f_2 = f_2^{BDet(R)}$ are homomorphisms into Sat(R), and it can be shown that since Det(R) and BDet(R) are minimal (reduced), the homomorphisms are injections.
Also since Sat(R) is minimal, it is true that Det(R) and BDet(R) have only one homomorphic image in Sat(R), so we will not destinguish between Det(R) and BDet(R) as automata on their own, and as subautomata of Sat(R). Thus we may unambiguously define their intersection Core(R).

Definition 3.7: Let f_1 and f_2 be the (uniquely determined) isomorphisms of Det(R) and BDet(R) into Sat(R). We let

$K_{Det} = f_1(\hat{P})$, $K_{BDet} = f_2(\hat{Q})$,

$M_{Det} = \{(f_1(p),a,f_1(\delta_D(p,a))) \mid p\in\hat{P},\ a\in\Sigma\} = f_1\{(p,a,\delta_D(p,a)) \mid p\in\hat{P}, a\in\Sigma\}$,

$M_{BDet} = \{(f_2(\delta_B(q,a)),a,f_2(q)) \mid q\in\hat{Q},\ a\in\Sigma\} = f_2\{(\delta_B(q,a),a,q) \mid q\in\hat{Q}, a\in\Sigma\}$.

Then $Core(R) = (K_{\mathfrak{C}},\Sigma,M_{\mathfrak{C}})$

where

$M_{\mathfrak{C}} = M_{Det} \cap M_{BDet}$, and $K_{\mathfrak{C}}$ are the states involved.

Note: In saturated automota $\mathcal{A}$ (which are not minimal) Det(R) and BDet(R) may have many images in $\mathcal{A}$, and thus the intersection Core(R) does not make sense.
Each word $w \in R$ have a unique accepting path in Det(R) and in BDet(R), and via f_1 and f_2 these paths give us two accepting paths $P_D(w)$ and $P_B(w)$ in Sat(R).

Definition 3.8: Given $w = a^1 \ldots a^k$ $a^i \in \Sigma$, write $w^i = a^1 \ldots a^i$ and $v^i = a^{i+1} \ldots a^k$ $i = 0, \ldots k$.
Then

$$P_D(w) = (f_D(e), a^1, \ldots, f_D(w^i), a^{i+1}, \ldots, f_D(w^k))$$

$$P_B(w) = (f_B(w), a^1, \ldots, f_B(v^i), a^{i+1}, \ldots, f_B(e))$$

where $f_D(u) = f_1^{Det(R)}(\delta_D(p_e, u)) = (\bar{u}, u\ R)$

$f_B(u) = f_2^{BDet(R)}(\overleftarrow{\delta}_B(q_e, \overleftarrow{u})) = (R/u, \tilde{u})$

We order the maximal pairs by

$$(P_i, Q_i) \leq (P_j, Q_j) \quad \text{iff} \quad P_i \supseteq P_j \ (\text{iff } ((P_j, Q_j), e, (P_i, Q_i)) \in M_R)$$

Proposition 3.6: Every accepting path for w in Sat(R) lies under $P_D(w)$ and over $P_B(w)$, i.e.
Given $w \in R$ where $b_1 \ldots b_\ell = w$, $b_i \in \Sigma \cup \{e\}$ and $\ell \geq k = |w|$. If $V = (t_0, b_1, \ldots, b_\ell, t_\ell)$ is an accepting path for w in Sat(R), and if we write

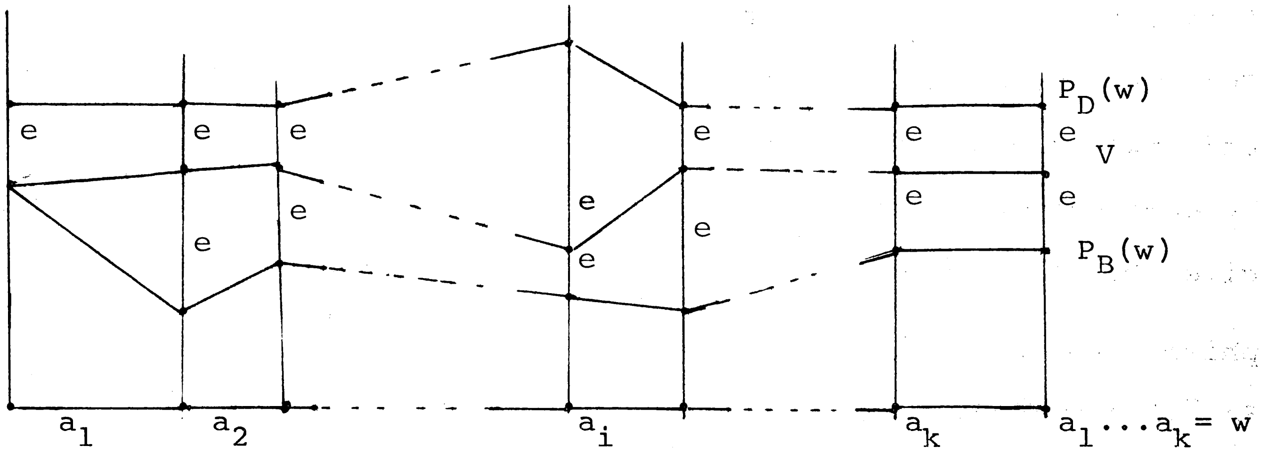
$$t_j = t_j^i \quad \text{iff} \quad b_1 \ldots b_j = w^i$$

then for all $j = 0, \ldots, \ell$ (and the correponding $i \in \{0, \ldots, k\}$) the following holds:

$f_B(v^i) \leq t_j^i \leq f_D(w^i)$.

<u>Proof</u> omitted.

Illustration of the proposition:



Along each vertical line the pairs are ordered by $\leqslant$ (down directed e-transitions).

<u>Ex. 3.1, continued</u>

We will construct $Sat(R_3)$.

We know the states $K_{R_3} = \{r_0, r_1, r_2, r_3, r_4\}$. From prop. 3.4 it follows that $r_3 = (\emptyset, \Sigma^*)$ is (the only) inaccesible state, and that $r_4 = (\Sigma^*, \emptyset)$ is (the only) dead state.

$$S_R = \{(P_i, Q_i) \mid e \in P_i\} = \{(P_i, Q_i) \mid [e] \subseteq P_i\}.$$

Since $[e] = [0]$; $S_R = \{([0], <2>), (\Sigma^*, \emptyset)\} = \{r_0, r_4\}$.

Similarily

$$F_R = \{(P_i, Q_i) \mid <e> \subseteq Q_i\}.$$

Since $<e> = <0>$;, $F_R = \{([2], <0>), (\emptyset, \Sigma^*)\} = \{r_2, r_3\}$.

The transitions are determined by

$$(r_i, k, r_j) \in M_{R_3} \quad \text{iff} \quad P_i\{k\} \subseteq P_j \ , \quad \text{for} \quad k \in \Sigma \cup \{e\}.$$

For $i = 0,1,2$, $P_i = [i]$ so

$$(r_i, k, r_j) \in M_{R_3} \quad \text{iff} \quad [i]\{k\} \subseteq [j] \quad \text{which for} \quad k = 0,1,2$$

correspond to the transitions in $Det(R_3) = \mathscr{A}_3$.

Thus $(r_i,k,r_j) \in M_{R_3}$ iff $i+k \equiv j \mod 3$, $k = 0,1,2$ (we may ignore the trivial (r_i,e,r_i) transitions).

The transitions in connection with dead and inaccessible state (r_4 and r_3) are not so interesting.

We conclude that $Sat^-(R) \simeq \mathcal{A}_3$, since the mapping $f:\mathcal{A}_3 \to Sat(R)$ given by $f(q_i) = r_i$ is easily seen to be an injective homomorphism onto $Sat^-(R)$. In fact $f = f_1^{Det(R_3)} = f_2^{BDet(R_3)}$ so $Sat^-(R_3) = f(\mathcal{A}_3) = f_1(Det(R_3)) = f_2(BDet(R_3))$ and $Sat^-(R_3) = Core(R_3) \simeq \mathcal{A}_3$.

This is a general fact:

If $\mathcal{A}$ is both forward and backward deterministic (with $\#S=\#F=1$), and $R = T(\mathcal{A})$, then $\mathcal{A} = Det^-(R) = BDet^-(R)$ and $Sat^-(R) = Core(R) \simeq \mathcal{A}$.

In this case it has long been known that $h(R) = rank(\mathcal{A})$ (see [3]), and thus $h(R) = rank(Sat^-(R)) = rank(Core(R))$ in such examples.

4. Sat(R) gives upper and lower bounds for h(R)

We know that for all automata $\mathcal{A}$ accepting R, there exists a homomorphism f from $\mathcal{A}$ into Sat(R).

If $\mathcal{A}$ is of minimal rank (i.e. $h(R) = rank(\mathcal{A})$), it could happen that $f(\mathcal{A}) \subseteq Sat(R)$ also is of minimal rank, but this need not be the case. We do, however, have an <u>upper bound</u> for h(R):

$$h(R) \leq \min\{rank(\mathcal{A}) \mid T(\mathcal{A}) = R, \mathcal{A} \subseteq Sat(R)\}$$

Despite serious efforts it is still an open question whether this upper bound is, in fact, exact (see Section 6).

We will now give a lower bound for $h(R)$ using McNaughton's pathwise homomorphism theorem and proposition 3.6.

This lower bound is known not always to be exact (e.g. $\mathrm{rank}(\mathrm{Core}(R)) = 0$ in many cases).

Theorem 4.1: $h(R) \geq \mathrm{rank}(\mathrm{Core}(R))$.

Proof. Choose $\mathcal{A}'$ of minimal rank and without e-transitions. We have a homomorphism $f: \mathcal{A}' \to \mathrm{Sat}(R)$. Write $\mathcal{A}'_0 = f^{-1}(\mathrm{Core}(R))$.

Claim. f is pathwise from $\mathcal{A}'_0$ onto $\mathrm{Core}(R)$.
Choose a path $V_0 = (r^1, a^1, \ldots, r^n, a^1, r^{n+1})$ in $\mathrm{Core}(R)$, where $a^1, \ldots, a^n = w_0$. Choose words w_1, w_2 where

$$w_1 \in Pr^{\mathrm{Det}(R)}(p^1) \quad \text{and} \quad f_1(p^1) = r^1$$

$$w_2 \in Sc^{\mathrm{BDet}(R)}(q^{n+1}) \quad \text{and} \quad f_2(q^{n+1}) = r^{n+1}.$$

Then $w_1 w_0\, w_2 \in R$ and thus accepted in $\mathcal{A}'$ by a path V.
The path $f(V')$ is by 3.3 squeezed between $P_D(w_1 w_0\, w_2)$ and $P_B(w_1 w_0\, w_2)$, and since V_0 is in $\mathrm{Core}(R)$, the two paths are identical on w_0. Thus $V' = V'_1 V'_0 V'_2$ where $f(V'_0) = V_0$ and V'_0 is a path in $f^{-1}(\mathrm{Core}(R)) = \mathcal{A}'_0$.
This shows that f is pathwise from $\mathcal{A}'_0$ onto $\mathrm{Core}(R)$, and we conclude:

$$h(R) = \mathrm{rank}(\mathcal{A}') \geq \mathrm{rank}(\mathcal{A}'_0) \geq \mathrm{rank}(\mathrm{Core}(R)).$$

5. Some exact star heights. The finite intersection property.

In which cases do we have equality in theorem 4.1? We saw already an example in 3.1, and in this section we will give some further examples and some general results to get a clearer picture of the strength and usefulness of theorem 4.1.

We need the following definition (mainly from Kameda and Weiner [11]):

Definition 5.1: Given $\mathcal{A} = (Q,\Sigma,S_0,F)$, let

$$\mathfrak{D}(\mathcal{A}) = (P,\Sigma,M',\{p_0\},F') \quad \text{where}$$

$$P = \{M(S_0,w) \mid w\in\Sigma^*\} = \{p_0,\ldots,p_m\} \subseteq 2^Q$$

$$p_0 = \{M(S_0,e)\} \qquad F' = \{p\in P \mid p\cap F \neq \emptyset\}$$

$$(p_i,a,p_j) \in M' \quad \text{iff} \quad M(p_i,a) = p_j \quad \forall i,j \quad \forall a\in\Sigma.$$

This is the subsetconstruction, and it is well known that $\mathfrak{D}(\mathcal{A})$ is a deterministic automaton accepting $T(\mathcal{A})$.

We want to reduce it to the (unique) minimal deterministic automaton for R. (See the minimization algorithm, Theorem 3.11 of Hopcroft and Ullman [10].)

Definition 5.2: Given a deterministic automaton $\mathcal{B} = (P,\Sigma,\delta,\{p_0\},F)$.

Define an equivalence relation on P by

$$p_i \overset{D}{\sim} p_j \ (\mathrm{mod}\ \mathcal{B}) \quad \text{iff} \quad Sc^{\mathcal{B}}(p_i) = Sc^{\mathcal{B}}(p_j).$$

Let $\hat{\mathcal{B}} = (\hat{P},\Sigma,\hat{M},\{\hat{p}\},\hat{F})$ be the automaton obtained from $\mathcal{B}$ by identifying equivalent states.

If no states in $\mathcal{B}$ are equivalent $(\mathcal{B} \simeq \hat{\mathcal{B}})$, we say that $\mathcal{B}$ is reduced (and $Det^-(T(\mathcal{B})) = \mathcal{B}^-$).

The following theorem will be useful:

Theorem 5.1 (Brzozowski, from [11]).

Let $\mathcal{B} = (P,\Sigma,\delta,\{p_0\},F)$ be a deterministic automaton (not necessarily reduced), with $T(\mathcal{B}) = R$. Then $\mathfrak{D}(\mathcal{B}^{\leftarrow})$ is a reduced deterministic automaton accepting R^T.

And by duality: if $\mathcal{C}$ is a backward deterministic automaton (BDA) accepting R, then $\mathfrak{D}(\mathcal{C})$ is reduced.
It is well known that the reduction in definition 5.2 gives us Det(R):

<u>Theorem 5.2</u>: For every non-deterministic automaton $\mathcal{A}$ with $T(\mathcal{A}) = R$,

$$\mathfrak{D}(\mathcal{A})^{\wedge} \simeq \mathrm{Det}(R) = (\hat{P}, \Sigma, \delta_D, p_e, F_D)$$

And since $\mathrm{BDet}(R)^{\leftarrow} = \mathrm{Det}(R^T)$, by duality we have

$$(\mathfrak{D}(\mathcal{A}^{\leftarrow}))^{\wedge} \simeq \mathrm{BDet}(R)^{\leftarrow} = (\hat{Q}, \Sigma, \overset{\leftarrow}{\delta}_B, \{q_e\}, S_B)$$

By 5.1 and 5.2, if $\mathcal{A}$ is a DA, $T(\mathcal{A}) = R$, then $\mathfrak{D}(\mathcal{A}^{\leftarrow}) \simeq \mathrm{BDet}(R)^{\leftarrow}$ and if $\mathcal{B}$ is a BDA, $T(\mathcal{B}) = R$, then $\mathfrak{D}(\mathcal{B}) \simeq \mathrm{Det}(R)$.
The reductions (the $\wedge$-operations) can easily be performed by using the following matrix:

<u>Definition 5.3</u> (from Kameda and Weiner [11])
The elementary automaton matrix (<u>EAM</u>) relative to $\mathcal{A} = (S, \Sigma, M_0, S_0, F_0)$ is defined as a $\#P \times \#Q$ matrix indexed by the states in $\mathfrak{D}(\mathcal{A})$ and the states in $\mathfrak{D}(\mathcal{A}^{\leftarrow})$ with values

$$\mathrm{EAM}(p,q) = \begin{cases} 1 & \text{iff} \quad p \cap q \neq \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Rows (and columns) with equal 1/0 patterns correspond to states in $\mathfrak{D}(\mathcal{A})$ (and $\mathfrak{D}(\mathcal{A}^{\leftarrow})$) which are $(\overset{D}{\sim})$ equivalent (in the sense of definition 5.2), and thus EAM is seen to be useful in the computation of $\mathfrak{D}(\mathcal{A})^{\wedge}$ and $\mathfrak{D}(\mathcal{A}^{\leftarrow})^{\wedge}$.

Continuing definitions 5.1-5.3, each state $[p]$ in $\mathcal{D}(\mathcal{A})^{\wedge}$ may be regarded as a union $\bigcup\{p' \mid p' \overset{D}{\sim} p\} \subseteq S$ and each state $[q]$ in $\mathcal{D}(\mathcal{A}^{\leftarrow})^{\wedge}$ may be regarded as a union $\bigcup\{q' \mid q' \overset{D}{\sim} q\} \subseteq S$.

With this notation [11] defines RAM as a $\#\hat{P} \times \#\hat{Q}$ matrix where

$$\mathrm{RAM}([p],[q]) = \begin{cases} 1 & \text{iff } [p] \cap [q] \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

By theorem 5.2 this definition of RAM is equivalent to definition 3.4.

Definition 3.4 was useful in proving the theorems of that section, but in many applications, if we are only given a NDA $\mathcal{A}$ for R, it turns out that the construction of $\mathcal{D}(\mathcal{A})$, $\mathcal{D}(\mathcal{A}^{\leftarrow})$, EAM, RAM, $\mathcal{D}(\mathcal{A})^{\wedge}$ and $(\mathcal{D}(\mathcal{A}^{\leftarrow})^{\wedge})^{\leftarrow}$ will often be more convenient for the computations.

We will now illustrate (parts of) the construction of Sat(R) by a simple example.

<u>Ex. 5.1</u>: Let $\mathcal{A}$ be the automaton shown in fig. 5.1 a). Let $R = T(\mathcal{A})$. $\mathcal{A}$ is deterministic, so 5.1 and 5.2 give $\mathrm{BDet}(R) = (\mathcal{D}(\mathcal{A}^{\leftarrow}))^{\leftarrow}$.

The transition table for $\mathcal{A}$ is shown in fig. 5.1 b), and for $\mathcal{A}^{\leftarrow}$ and $\mathcal{D}(\mathcal{A}^{\leftarrow})$ in fig. 5.1 c) d).

Fig. 5.1 a)

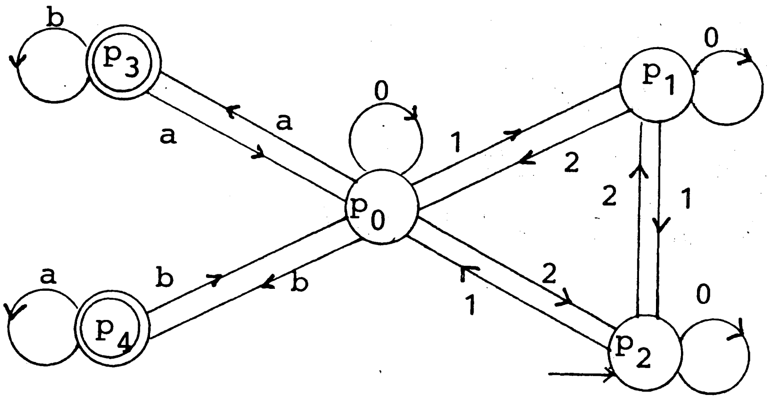


Fig. 5.1 b)

$\mathcal{A}$	0	1	2	a	b
p_0	p_0	p_1	p_2	p_3	p_4
p_1	p_2	p_0	p	-	-
p_2	p_2	p_0	p_1	-	-
p_3	-	-	-	p_0	p_3
p_4	-	-	-	p_4	p_0

Fig. 5.1 c)

0	1	2	a	b	$\mathcal{A}^{\leftarrow}$
p_0	p_2	p_1	p_3	p_4	p_0
p_1	p_0	p_2	-	-	p_1
p_2	p_1	p_0	-	-	p_1
-	-	-	p_0	p_3	p_3
-	-	-	p_4	p_0	p_4

Fig. 5.1 d)

0	1	2	a	b	$\mathfrak{D}(\mathcal{A})^{\leftarrow}$
-	-	-	p_{04}	p_{03}	$p_{34}=q_5$
p_0	p_2	p_1	p_{34}	p_{04}	$p_{04}=q_6$
p_0	p_2	p_1	p_{03}	p_{34}	$p_{03}=q_7$
p_0	p_2	p_1	p_3	p_4	$p_0=q_0$
p_1	p_0	p_2	-	-	$p_1=q_1$
p_2	p_1	p_0	-	-	$p_3=q_2$
-	-	-	p_0	p_3	$p_3=q_3$
-	-	-	p_4	p_0	$p_4=q_4$
-	-	-	-	-	- $=q_8$

Fig. 5.1 e)

EAM=RAM	q_0 $\{p_0\}$	q_1 $\{p_1\}$	q_2 $\{p_2\}$	q_3 $\{p_3\}$	q_4 $\{p_4\}$	q_5 $\{p_3,p_4\}$	q_6 $\{p_0,p_4\}$	q_7 $\{p_0,p_3\}$	q_8 $\emptyset$
p_0 $\{p_0\}$	1	0	0	0	0	0	1	1	0
p_1 $\{p_1\}$	0	1	0	0	0	0	0	0	0
p_2 $\{p_2\}$	0	0	1	0	0	0	0	0	0
p_3 $\{p_3\}$	0	0	0	1	0	1	0	1	0
p_4 $\{p_4\}$	0	0	0	0	1	1	1	0	0
p_9 $\emptyset$	0	0	0	0	0	0	0	0	0

Fig. 5.1 f)

We represent the maximal pairs in the following way:

$r_0 \ : \ (\{p_0\}, \{q_0, q_6, q_7\})$

$r_1 \ : \ (\{p_1\}, \{q_1\})$

$r_2 \ : \ (\{p_2\}, \{q_2\})$

$r_3 \ : \ (\{p_3\}, \{q_3\})$

$r_4 \ : \ (\{p_4\}, \{q_4, q_5, q_6\})$

$r_5 \ : \ (\{p_3, p_4\}, \{q_5\})$

$r_6 \ : \ (\{p_0, p_4\}, \{q_6\})$

$r_7 \ : \ (\{p_0, p_3\}, \{q_7\})$

$r_8 \ : \ (\emptyset, \hat{Q}) \qquad r_9 \ : \ (\hat{P}, \emptyset)$

From $\mathcal{A}$ (or really $\mathfrak{D}(\mathcal{A})$) and $\mathfrak{D}(\mathcal{A}^{\leftarrow})$ we get the EAM shown in fig. 5.1 e). And we see that EAM is reduced (EAM = RAM), and thus $\mathcal{A} = \mathrm{Det}^{-}(R)$ and $\mathfrak{D}(\mathcal{A}^{\leftarrow}) = \mathrm{BDet}(R)^{\leftarrow}$, as we already knew.

The maximal pairs are computed as explained in ex. 3.1. See fig. 5.1. f).

We compute $f_1^{\mathrm{Det}(R)}$ and $f_2^{\mathrm{BDet}(R)}$:

$f_1(p_i) = (\bar{p}_i, p_i\backslash R)$, $i=0,1,2,3,4,9,$

$f_2(q_j) = (R/q_j, \tilde{q}_j)$, $j=0,1,2,3,4,5,6,7,8,$

and we see that $f_1(p_i) = r_i$, $i=0,1,2,3,4,9,$

$f_2(q_j) = r_j$, $i=0,1,2,3,4,5,6,7,8.$

(it is often convenient to arrange the numbering this way).

We now turn to Core(R), the intersection of $f_1(\mathrm{Det}(R))$ and $f_2(\mathrm{BDet}(R))$, with states $\{r_0,r_1,r_2,r_3,r_4\}$.
Core(R) is naturally a subgraph in $\mathrm{Det}(R) = \mathcal{A}$ seen as a subautomaton of Sat(R).
Write $\mathcal{A}' = \mathcal{A}\restriction\{r_0,r_1,r_2,r_3,r_4\} \subseteq \mathrm{Det}^{-}(R) \subseteq \mathrm{Sat}(R)$. Since the transitions in $\mathcal{A}'$ are backward deterministic (i.e. $\#\{r' \in Q_{\mathcal{A}'} \mid (r',a,r) \in M_{\mathcal{A}'}\}$ ≤ 1), for all $r \in Q_{\mathcal{A}'}$), the transitions in $\mathcal{A}'^{\leftarrow}$ are (forward) deterministic.
And since the states r_0,r_1,r_2,r_3,r_4 are states in $\mathfrak{D}(\mathcal{A}^{\leftarrow})$, all the transitions in $\mathfrak{D}(\mathcal{A}')^{\leftarrow}$ will be transitions in $\mathfrak{D}(\mathcal{A}^{\leftarrow})$ and in $\mathrm{BDet}(R)^{\leftarrow}$.
This shows that all transition in $\mathcal{A}'$ are transitions both in BDet(R) and in Det(R), and hence $\mathcal{A}'$ is contained in Core(R).
Thus $3 = \mathrm{rank}(\mathcal{A}') \leq \mathrm{rank}(\mathrm{Core}(R)) \leq h(R)$. Since Det(R) is of rank 3, this gives $h(R) = 3$.

Definition 5.4: Given a semiautomaton $\mathcal{A} = (S,\Sigma,M)$. We say that is a permutation automaton iff

1) $\mathcal{A}$ is without e-transitions (or only (s,e,s) transitions),

2) $\forall a \in \Sigma \quad \forall s \in S \quad \#\{s' \mid (s',a,s) \in M\} \leq 1$

$$\#\{s' \mid (s,a,s') \in M\} \leq 1.$$

$\mathcal{A}$ is a complete permutation automaton if we have equalities in 2.

Thus the subsautomaton $\mathcal{A}'$ of $\mathcal{A}$ in ex. 5.1 is an incomplete permutation automaton, and we saw that the states in $\mathcal{A}'$ are states in Det(R) and BDet(R).

We can formalize this.

Lemma 5.3:

Given an incomplete permutation automaton $\mathcal{A}' \subseteq \mathcal{A}$, and states p and p' in $\mathfrak{D}(\mathcal{A})$ q and q' in $\mathfrak{D}(\mathcal{A}^{\leftarrow})$ where $p = \{s\}$, $p' = \{s'\}$, $q = \{s\}$, $q' = \{s'\}$, s,s' in $\mathcal{A}'$, then all transitions between s and s' in $\mathcal{A}'$ will give rise to corresponding transitions between p and p' in $\mathfrak{D}(\mathcal{A})$ and to transitions between q and q' in $(\mathfrak{D}(\mathcal{A}^{\leftarrow}))^{\leftarrow}$.

Combining this with theorem 5.1 we are in some cases able to find the exact star height of R.

Proposition 5.4: If there exists an incomplete permutation automaton $\mathcal{A}'$ in $\mathcal{A} = \mathrm{Det}^{-}(R)$ such that all states in $\mathcal{A}'$ correspond to states in $\mathfrak{D}(\mathcal{A}^{\leftarrow})$, then $\mathcal{A}'$ is naturally a subgraph in Core(R), and thus $h(R) \geq \mathrm{rank}(\mathcal{A}')$.

Proof. We have assumed that for all $p \in K_{\mathcal{A}'}$, there exists a $q_p \in \hat{Q}$ such that $q_p = \{p\}$.
By proposition 5.1 EAM(Det(R)) = RAM(R). Thus we have the RAM

shown in fig. 5.4. This gives $R/q_p = p$, $\tilde{q}_p = p\backslash R$ and

$$f_1(p) = (\bar{p}, p\backslash R) = (\bar{p}, \tilde{q}_p) = (R/q_p, \tilde{q}_p) = f_2(q).$$

We know from ex. 5.1 that $\mathcal{A}'$ may be regarded as subsautomaton of Det(R) and BDet(R), thus

$$f_1(\mathcal{A}') = f_2(\mathcal{A}') \subseteq f_1(\mathrm{Det}(R)) \cap f_2(\mathrm{BDet}(R)) = \mathrm{Core}(R)$$

Fig. 5.4:

		q_{p_1}	....	q_{p_i}	....	q_{p_n}	some more states
K	p_1	+	——	——	——	——	....
	p_i	——	——	+	——	——	....
	p_n	——	——	——	——	+	....
some more states	p_{n+1}	——	——	——	——	——	....
	p_m	——	——	——	——	——	....

—— means minus

.... means plus or minus (irrelevant)

We will now turn to a special class of regular events:

<u>Definition 5.5</u> (from Cohen [4])

R has the <u>finite intersection property</u> (f.i.p.) iff

$$\forall\ x,y \in \Sigma^*,\quad x\backslash R \neq y\backslash R \Rightarrow x\backslash R \cap y\backslash R \text{ is finite.}$$

In the framework of $\mathrm{Det}(R) = (\hat{P}, \Sigma, \delta_D, p_e, F_D)$, where $Sc(p_i) = w_i\backslash R$, f.i.p. is equivalent to

$$p \neq p' \Rightarrow (Sc(p) \cap Sc(p') \text{ is finite}).$$

<u>Lemma 5.5</u>: When R has the f.i.p. there exists a semiautomaton $\mathcal{A}'$ in $\mathrm{Det}^{-}(R) = \mathcal{A}$ such that

1) $\mathcal{A}'$ is an incomplete permutation automaton,
2) all states in $\mathcal{A}'$ correspond to states in $\mathcal{D}(\overleftarrow{\mathcal{A}})$,
3) $\mathrm{rank}(\mathcal{A}') = \mathrm{rank}(\mathrm{Det}^{-}(R))$.

<u>Proof</u>: (We write M_D for the transition relation in Det(R).) Write $K_f = \{p \in \hat{P} \mid Sc(p) \text{ is finite}\}$ and $K_{\mathcal{A}'} = \{p \in \hat{P} \mid Sc(p) \text{ is infinite}\}$. Let $\mathcal{A}' = \mathrm{Det}^{-}(R) \upharpoonright K_{\mathcal{A}'}$ $(= \mathrm{Det}^{-}(R) - [K_f])$.

If $(p,a,p") \in M_D$ and $(p',a,p") \in M_D$, $p \neq p'$, then $\{a\}Sc(p") \subseteq Sc(p) \cap Sc(p')$, and since R has the f.i.p., this gives $p" \in K_f$. This shows that for each $p' \in K$,, $\{p \mid (p,a,p') \in M_D\} \leqslant 1$, and thus $\mathcal{A}'$ is a permutation automaton.

Since K_f corresponds to a semiautomaton without loops, all loops in Det(R) are in $\mathcal{A}'$, showing that $\mathrm{rank}(\mathcal{A}') = \mathrm{rank}(\mathrm{Det}^{-}(R))$.

We will now prove that $\mathcal{A}'$ satisfies 2), i.e. for all $p \in K_{\mathcal{A}'}$ there exist a state q_p such that $q_p = \{p\}$.

The transitions $\overleftarrow{\delta}_B$ in the subset construction from $\overleftarrow{\mathrm{Det}^{-}(R)}$, are determined by

$$\overleftarrow{\delta}_B(q,w) = \{p \in \hat{P} \mid (p,w) \in q\}$$

Since $q_e = F_D$ we have

$$\overleftarrow{\delta}_B(q_e,w) = q \iff q = \{p \mid \delta_D(p,w) \in F_D\}$$
$$\iff q = \{p \mid w \in Sc^{\mathrm{Det}(R)}(p)\}.$$

We must show that for each $p \in K$, there exists a $w \in \Sigma^*$ such that $w \in Sc(p)$ and $(p' \neq p \Rightarrow w \notin Sc(p'))$. This follows since R has the f.i.p., so $Sc(p) \cap Sc(p')$ is finite when $p' \neq p$, while $Sc(p)$, $p \in K_{\mathcal{A}'}$ is infinite.

Choose any $w \in Sc(p) - \bigcup\{Sc(p) \cap Sc(p') \mid p' \neq p\}$. This ends the proof of 5.5 Combining this result with 5.4 we conclude,

<u>Theorem 5.6</u>: When R has the f.i.p., $h(R) = \text{rank}(\text{Det}^{-}(R))$.

This result was also proved by Cohen, see Theorem 5.1 of [4]. Cohen's definition of f.i.p. uses only the left quotients $(w\backslash R)$. We could equally well have defined "right f.i.p." by using R/w, and since R has "left f.i.p." iff R^T has "right f.i.p." and $h(R) = h(R^T)$, we can modify theorem 5.6.

<u>Theorem 5.6'</u>: If T has left or right f.i.p., then

$$h(R) = \min(\text{rank}(\text{Det}^{-}(R)), \text{rank}(\text{BDet}^{-}(R))$$
$$= \text{rank}(\text{Core}(R)).$$

Cohen [5] gives some further theorems on star heights. In particular her Theorem 4.2 reads with some slight adoption: Suppose $\text{Det}\ (R)$ is an (incomplete) permutation automaton. Let S be a section in $\text{Det}^{-}(R)$. If there exists a state q in S and a word w_0 such that $M(q,w_0) \in F_D$ and $M(q',w_0) \notin F_D$ when $q' \neq q$, then $h(R) \geq \text{rank}(S)$.

An alternative proof could be obtained by showing that the premisses corresponds to the premisses in 5.4, and that S may be regarded as a subgraph in $\text{Core}(R)$.

It seems likely that we can obtain the same results as Cohen. By the fact that $\text{Sat}(R)$ is smaller than Cohens automaton $\mathcal{A}_1(R)$, we can lose some information. But on the other hand we used the fact that $\text{Sat}(R)$ is minimal in an essential way in order to define $\text{Core}(R)$, and to prove Theorem 4.1.

6. An open question

We say that $\mathcal{A}$ is of minimal rank with respect to R iff $T(\mathcal{A}) = R$ and $\mathrm{rank}(\mathcal{A}) = h(R)$. The question is:

"Does there always exist a subautomaton of Sat(R) of minimal rank?"

Or, put another way:

"Is $h(R) = \min\{\mathrm{rank}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathrm{Sat}(R),\ T(\mathcal{A}) = R\}$?"

This has been studied in [3] where they used the saturated automaton $\mathcal{A}_1(R)$ instead of the minimal Sat(R).

Definition 6.7:

Given $\quad \mathrm{Det}^{-}(R) = (\hat{P}, \Sigma, \delta_D, p_e, F_D)$,

define $\quad \mathcal{A}_1(R) = (P^1, \Sigma, M_1, P_0^1, F_0^1)$,

where $\quad P^1 = \{P' \mid \emptyset \neq P' \subseteq \hat{P}\}$, $\quad P_0^1 = \{P' \mid p_e \in P'\}$, $\quad F_0^1 = \{P' \mid P' \subseteq F_D\}$

$(P', a, P'') \in M_1 \iff \delta_D(P', a) \subseteq P''$, $\quad a \in \Sigma$.

Given $\quad \mathcal{A} = (Q, \Sigma, M, S, F)$ (not necessarily accepting R),

define $\quad f_R^{\mathcal{A}} : Q \to P^1$ by

$f_R(q') = \delta_D(p_e, \mathrm{Pr}^{\mathcal{A}}(q'))$.

Let $\mathcal{A}_k(R)$ denote the automaton with k duplicates of each state in $\mathcal{A}_1(R)$.

Note: By modifying $\mathcal{A}_1(R)$ and $\mathcal{A}_k(R)$ to allow e-transitions, and to allow empty subsets of $\hat{P}$, $\mathcal{A}_k(R)$ are saturated automata in our sense. (Because f_R is an homomorphism when $T(\mathcal{A}) \subseteq R$.) The main difference between $\mathcal{A}_1(R)$ and Sat(R), is that $\mathcal{A}_1(R)$ uses all subsets of $\hat{P}$, while Sat(R) uses only those subsets P of $\hat{P}$ which are closed ($P=\bar{P}$). The closure operation is also the only difference between $f_R^{\mathcal{A}}$ and $f_1^{\mathcal{A}}$.

In [3] the version of the question was: "Does there exist a k (uniform in R or recursive in R) such that

$$h(R) = \min\{\text{rank}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathcal{A}_k(R),\ T(\mathcal{A}) = R\}\text{?"}$$

Their conclusion was (p. 280): "In fact, for any integer $t > 0$, an example of an event R_t can be constructed, such that no partial automaton of $\mathcal{A}_i(R)$ where $1 < i \leqslant t-1$ recognizes R_t and has rank $h(R_t)$; ... "

This should imply that no uniform k is possible, and a fortiori that $h(R) = \min\{\text{rank}(\mathcal{A}) \mid \mathcal{A} \subseteq \text{Sat}(R), T(\mathcal{A})=R\}$ is in general wrong.

However, they do not give any expression for R_t, but they do give an example of R_t for $t = 2$, (ex. 6.5 in [3]). But it turns out that with R_2 as in ex. 6.5, there does in fact exist an automaton $\mathcal{A}_{min}$ of minimal rank and

$$\mathcal{A}_{min} \subseteq \text{Sat}(R) \subseteq \mathcal{A}_1(R).$$

This shows that though Sat(R) is in general smaller than $\mathcal{A}_1(R)$, this need not be a drawback.

We now turn to the example.

<u>Ex. 6.1</u> (ex. 6.5 from [3])

Let $R_2 = T(\mathcal{A})$ where $\mathcal{A}$ is given in fig. 6.1 a). ($\mathcal{A} = \text{Det}(R_2)$.)

One natural expression for R is:

$E = (11^*(0\vee 2) \vee (0\vee 2)(1\vee 2) \vee (0\vee 2)01^*(0\vee 2))^*$, $h_\alpha(E) = 2$, since $\text{rank}(\mathcal{A}) = 2$.

But $h(R) = 1$, because we have the automaton $\mathcal{A}'$ of minimal rank as shown in fig. 6.1 b). $\mathcal{A}'$ corresponds to the expression

$$E' = ((0\vee 1\vee 2)^* 1\ 1(0\vee 2) \vee e)\cdot E''$$

where $$E'' = [(0\vee 2)((1\vee 2) \vee 0((0\vee 2)\vee 1(0\vee 2) \vee 1((0\vee 2) \vee 1(0\vee 2))]^*$$

$\mathcal{A}'$ is a subautomaton of $\mathcal{A}_2(R)$, but not a subautomaton of $\mathcal{A}_1(R)$.

$\mathcal{A}'$ is constructed by splitting the loop $(p_1,1,p_1)$ in a transition $(p_1,1,p_1')$ where p_1 and p_1' are duplicates (relative to outgoing transitions).

We expect that all homomorphisms $f: \mathcal{A}' \to \mathcal{A}_1(R)$ (at least f_{R_2}, f_1, f_2) will give

$$f(p_1) = f(p_1')$$

and thus give us the undesired loop back.

But there are other ways to get rid of the loop $(p_1,1,p_1)$, e.g. by modifying $\mathcal{A}$, not to $\mathcal{A}'$, but to the nondeterministic $\mathcal{A}''$ shown in fig. 6.1 c).

$\mathcal{A}''$ is of minimal rank, and

$\mathcal{A}''$ is a subautomaton of $\text{Sat}(R)$.

We will not exhibit all of $\text{Sat}(R_2)$ here, but we will give the maximal pairs (represented by their first component), and the e-transitions. Together with $f_1(\mathcal{A}')$, this should indicate of how $\mathcal{A}''$ was found.

Fig. 6.1 a)

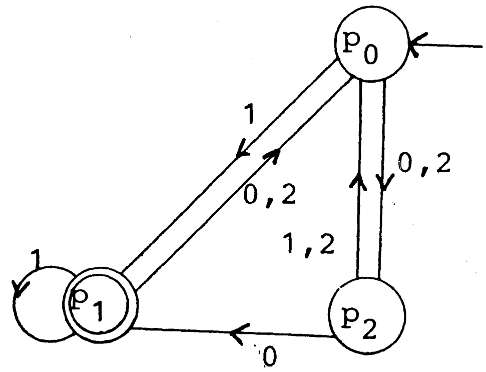


Fig. 6.1 b)

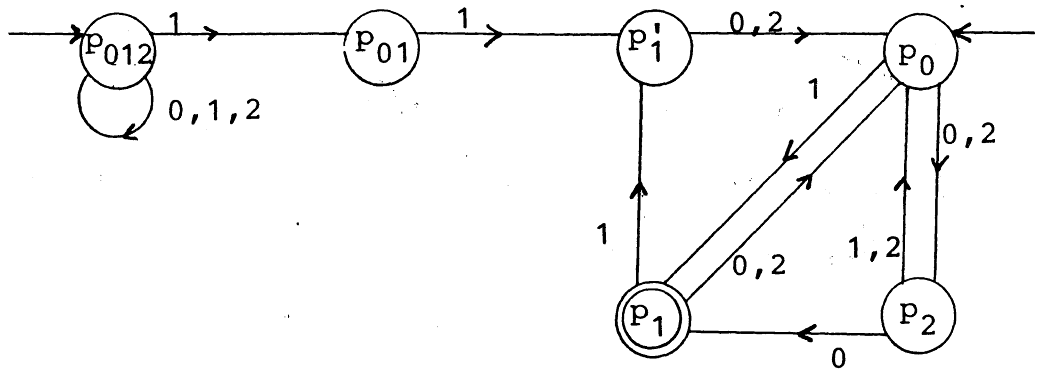


Fig. 6.1 c)

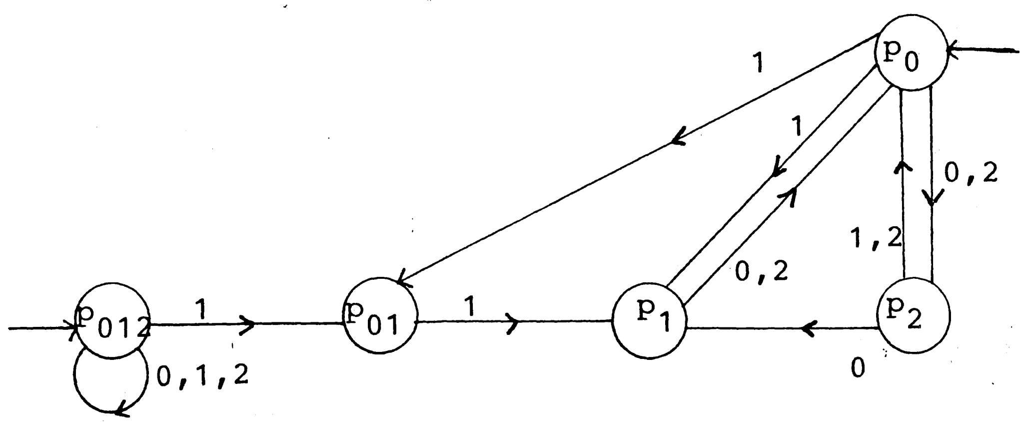
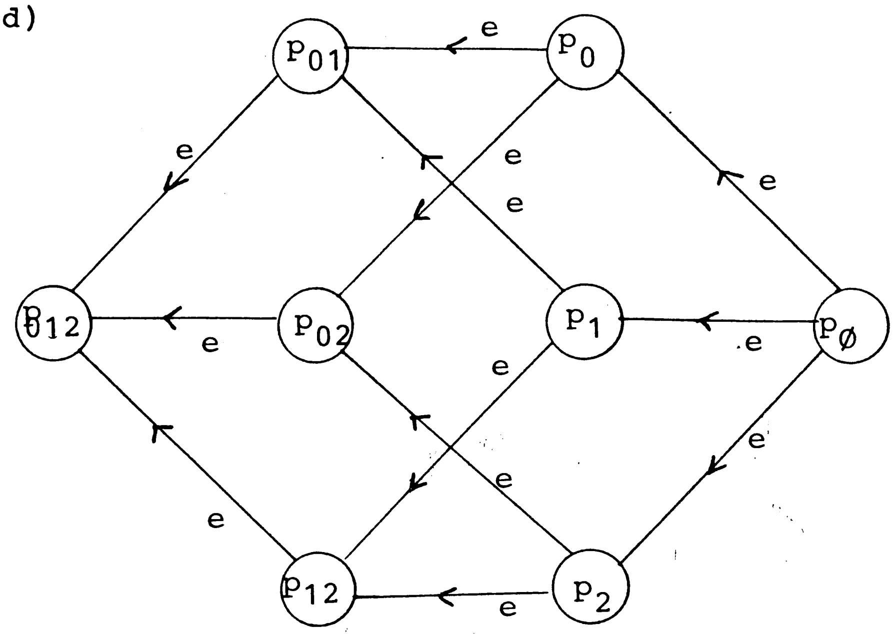


Fig. 6.1 d)



We have searched for other examples which possibily could give the final answer "no" to our question, but in vain.

We have also searched for a proof that the answer in "yes", but this seems far from easy.

We will end the paper with some examples where subautomata of Sat(R) of minimal rank does exist.

<u>Ex. 6.2</u>: Events where $\mathrm{Det}^-(R)$ or $\mathrm{BDet}^-(R)$ is of minimal rank. Then $f_1(\mathrm{Det}^-(R))$ or $f_2(\mathrm{BDet}(R))$ is also of minimal rank. E.g. ex. 4.3 in [3].

<u>Ex. 6.3</u>: The event $\bar{R}$ defined in the proof of corollary 5.6 in [3], is such that

$$\mathrm{rank}(\mathrm{Det}^-(\bar{R})) - h(\bar{R}) > k.$$

$\bar{R} = \bar{R}_1 \cup \ldots \cup \bar{R}_n$ where $n > k+2$ and

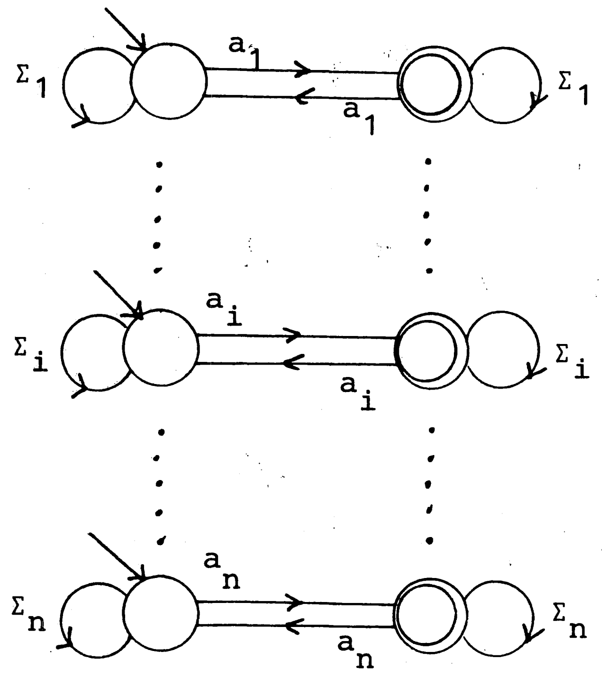
$\bar{R}_i$ = all words over $\{a_1,\ldots,a_n\}^*$ where the occurences of a_i is odd.

Then $\bar{R} = (\bar{R})^T$ and also $\mathrm{rank}(\mathrm{BDet}^-(\bar{R})) - h(\bar{R}) > k$.

However the automaton $\mathscr{A}$ shown in fig. 6.3 is of rank 2 and accepts $\bar{R}$.

We have (at least one) $f: \mathscr{A} \to \mathrm{Sat}(\bar{R})$, and it is easy to show that all f must be injective (otherwise $f(\mathscr{A})$ accepts too much). If we assume $h(\bar{R}) = 2$, then $f(\mathscr{A}) \subseteq \mathrm{Sat}(R)$ is of miminal rank.

Fig. 6.3.



$\Sigma_i=\{a_1,\ldots,a_n\}-\{a_i\}$

Ex. 6.4 (Due to Stål Aanderaa, originally constructed with the hope that it would give us the answer "no".)

Let $\quad R = (0(12)^*(21)^*0 \vee 01 \vee 10 \vee 12 \vee 21$
$\vee\ 0(12)^*212 \vee 212(21)^*0 \vee 212212)^*.$

We compute the root of R ([2])

$$\sqrt{\bar{R}} = (00 \vee 01 \vee 0120 \vee 012210 \vee 012212$$
$$\vee\ 0210 \vee 0212 \vee 10 \vee 12 \vee 21 \vee 2120$$
$$\vee\ 212210 \vee 212212)$$
$$= (w_1 \vee \ldots \vee w_n)$$

Since $R = \sqrt{R}^*$, $h(R) = 1$. We have a simple automoton $\mathscr{A}$ of minimal rank shown in fig. 6.4 a).

Fig. 6.4 a)

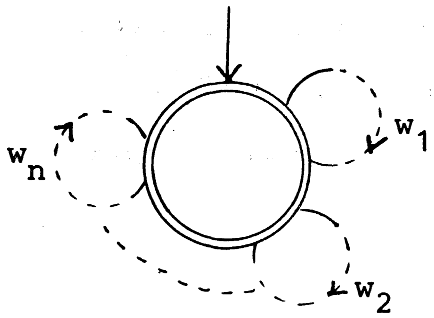
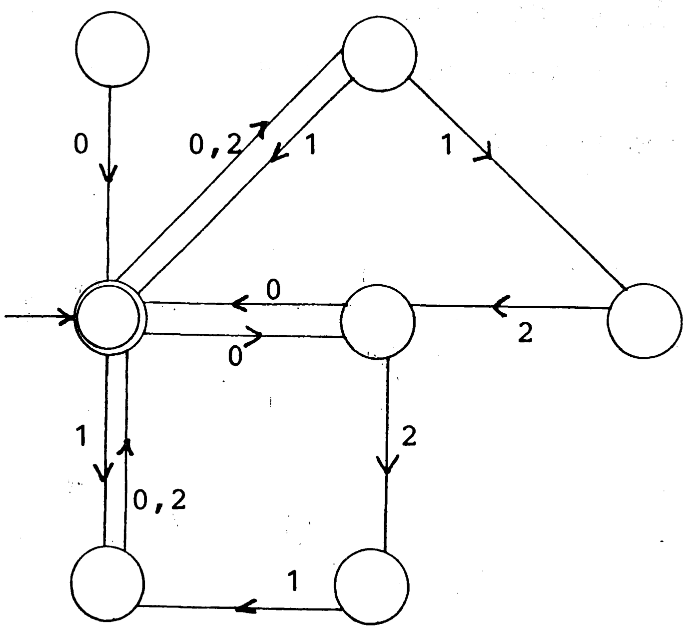


Fig. 6.4 b)



Neither Det(R) nor BDet(R) are of minimal rank.
If we choose $w = 012210 \in \sqrt{R}$, both $P_D(w)$ and $P_B(w)$ contain a loop, however, Core(R) is loopfree.

We have $f: \mathcal{A} \to \mathrm{Sat}(R)$ which is not injective, but preserves the rank. So $f(\mathcal{A})$ shown in fig. 6.4 b) is of minimal rank, and $f(\mathcal{A}) \subseteq \mathrm{Sat}(R)$.

<u>Ex. 6.5</u> (from [2])
It would have been nice if $h(R) = h(\sqrt{R}) + 1$, because $R = (\sqrt{R})^*$.
But even if $h(R) < h(\sqrt{R})+1$, this does not give us a counterexample, e.g. $R = e \vee 1(0 \vee 11)^* 1$ has $h(R) = 1$ and $\sqrt{R} = 10^* 1$ so $h(R) < h(\sqrt{R})+1$.

But $\mathrm{Det}^-(R)$ is of rank 1, and thus $f_1(\mathrm{Det}^-(R)) \subseteq \mathrm{Sat}(R)$ is of minimal rank.

<u>Ex. 6.6</u> (This is fig. 6 a-e from [14])

Let $R = (00^*1 \vee 101)^*$.

Let $\mathcal{A}$ be the rank minimal automaton found in fig. 6 e) [14]. Here shown as fig. 6.6 a) (slightly modified).

$f: \mathcal{A} \to \mathrm{Sat}(R)$ is not injective, but preserves the rank.

$f(\mathcal{A})$, shown in fig. 6.6 b), is thus of minimal rank and $f(\mathcal{A}) \subseteq \mathrm{Sat}(R)$.

Fig. 6.6 a)

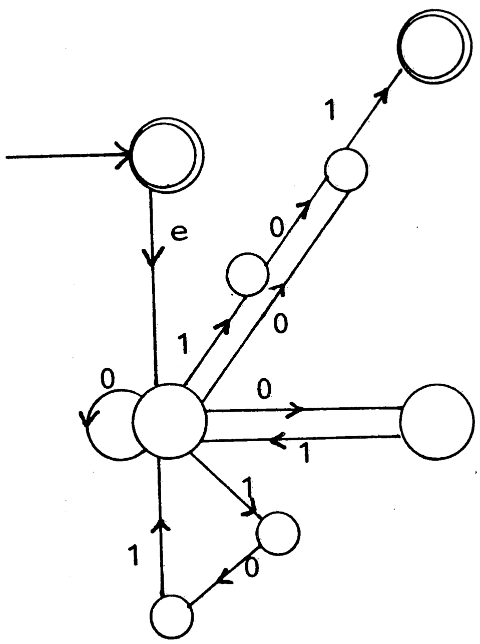
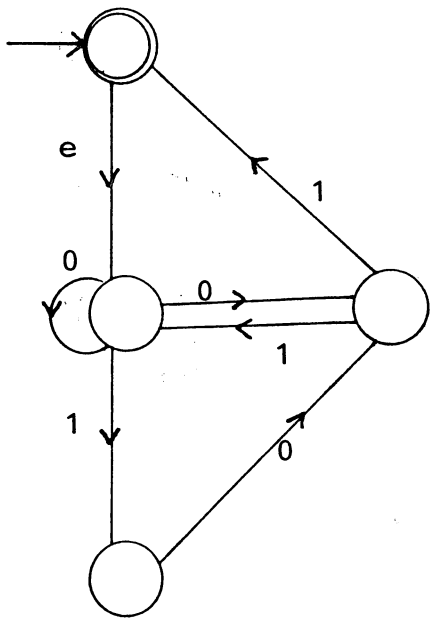


Fig. 6.6 b)



Examples 6.3, 6.4, and 6.5 show that even when the events R are "complicated", R does not answer the question: "Is $h(R) = \min\{\mathrm{rank}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathrm{Sat}(R),\ T(\mathcal{A}) = R\}$?" in the negative. Thus the question remains open.

Acknowledgements:

This paper and [12] are based on my cand.scient. thesis in mathematics at The University of Oslo.

I want to thank professor Stål Aanderaa who was my adivsor during my cand.scient studies.

The concept of minimal saturated automata and its construction is fully due to him.

His good ideas and clearifying examples were of great help to me.

Thanks also to professor Jens Erik Fenstad for useful comments on the manuscript.

References

[1] Brzozowski, J.: "Derivatives of regular expressions", J.ACM, vol. 5, pp 481-494, October 1964.

[2] Brzozowski, J.: "Roots of star events", J. ACM, vol. 14, pp 466-477, 1967.

[3] Cohen, R.S., Brzozowski, J.: "General properties ot star height of regular events", J. Computer System Sciences 4 (1970), pp 260-280.

[4] Cohen, R.S.: "Star height of certain families of regular events", J. Computer System Sciences 4 (1970), pp 281-297.

[5] Cohen, R.S.: "Techniques for establishing star height of regular sets", Math.System Theory (USA), Vol. 5, No. 2 (1971) pp 97-114.

[6] Cohen, R.S.: "Rank-non-increasing transformations on transition graphs", Information and Control 20 (1972), pp 93-113.

[7] Eggan, L.C.: "Transition graphs and the star height of regular events", Michigan Math. J. 10 (1963), pp 385-397.

[8] Hashiguchi, K.: "Regular languages of star height one", Information and Control 53 (1982), pp 199-210.

[9] Hashiguchi, K., Honda, N.: "The star height of reset-free events and stirictly locally testable events", Information and Control 40 (1979), pp 267-284.

[10] Hopcroft, J.E. Ullman, J.D.: "Introduction to automata theory, languages, and computation".

[11] Kameda, T., Weiner, P.: "On the state minimization of nondeterministic finite automata", IEEE Transact, on Computers C-19 No 7 1970, pp 617-627.

[12] Kristiansen, L.: Regular languages and minimal saturated automata, Preprint Series 4 (1986), University of Oslo.

[13] McNaughton, R.: "The loop complexity of pure-group events", Information and Control 11 (1967), pp 167-176.

[14] McNaughton, R.: "An introduction to regular expressions", Applied Automata Theory 1968, pp 35-54.

[15] McNaughton, R.: "The loop complexity of regular events", Information Science 1 (1969), pp 305-328.

[16] Meyer, A.R.: "A note on star free events", J. ACM, Vol 16, No 2 (1969), pp 220-225.

[17] Thomas, W.: "Remark on the star height problem", Theoretical Computer Sci. 13 (1981), pp 231-237.