1-1-1984

An Analysis of Some Proof Methods for Parallel Programs on Petri Nets.

William Joseph Seaman

An Analysis of Some Proof Methods for Parallel Programs on Petri Nets

by

William Joseph Seaman

A Thesis

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Computing Science

Lehigh University

1984

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

Sept. 20 '84

date

Professor in Charge

Chairman of Department

The author deeply appreciates the guidance and encouragement of Dr. Gerhard Rayna.

Table of Contents

Abstract

Two techniques for demonstrating the correctness of parallel programs are analyzed and compared: Keller's method and the Invariant method.

It is shown that Keller's method is the more powerful: any fact proveable using the Invariant method is also proveable using Keller's method, but not conversely.

It is known that the Invariant method is generally too coarse to handle a Petri net having a transition whose input and output places intersect. It is shown that it is possible to construct an equivalent net such that no transition has this property, but that the Invariant method will still fail on this new net.

It is shown how, under certain conditions, a transition system can be transformed to a Petri net. It is seen that if the Invariant method failed for the transition system it will also fail on the Petri net.

An attempt is made to construct a procedure that is more general than the Invariant method and more mechanical than Keller's method. The analysis of several examples indicates that the procedure is not yet sufficiently mechanical.

Chapter 1

1.1 Introduction to Petri Nets

A thorough discussion of Petri nets is given in [3]. We give here a very brief and informal introduction.

A Petri net consists of places (drawn as circles), transitions (drawn as bars), and arcs (labelled with positive integers. An unlabelled arc is implicitly labelled 1). In addition, each place is marked with a non-negative integer. (An unmarked place has the implicit mark 0).

A place pi is called an input place for the transition tj if there is an arc from pi to tj. It is called an output place if there is an arc from tj to pi.

A transition tj is enabled if, for each input place pi of the transition tj, the marking at pi is greater than or equal to the label on the arc from pi to tj.

A transition tj may fire only if it is enabled. A firing causes the markings at the input places to be decreased and the markings at the output places to be increased. The amount by which the marking is changed is equal to the label on the arc between pi and tj. (Since a transition may fire only if it is enabled, the markings at every place are always non-negative).

Example (see Fig. 1, which appears on p. 320 of [1]). In the readers/writers problem there are n processes, each of which may want to read or write to a data item. Any number of processes may read simultaneously, but a process that desires to write must have exclusive access--no other processes may write or read while the writing process is accessing the data item. One way to solve this problem is:

n permission slips are constructed.

In order to read, a process must obtain one permission slip. To write, a process must obtain n permission slips. This (alleged) solution is modelled in Fig. 1.

Note the abbreviations:

LP ... Local Processing

WR ... Waiting to Read

R ... Reading

WW ... Waiting to Write

S ... Slips

Initially, S is marked n, indicating there are n available permission slips. LP is marked n, indicating all processes are doing local processing. All other places are marked 0. The marking of all places can be represented as a vector in which

(1.1) $\underline{m}$ = [m(LP) m(WR) m(WW) m(R) m(W) m(S)]

and, eg, m(R) is the marking at place R .

Thus, the initial marking $\underline{m0}$ is

(1.2) $\underline{m0}$ = [n 0 0 0 0 n]

Initially, only the transitions t1 and t2 are enabled. Suppose t1 fires. Then we obtain the new marking:

$\underline{m1}$ = [(n-1) 1 0 0 0 n] At this point, t2 and t3 are enabled. Suppose t3 fires. We obtain the new marking m2 = [(n-1) 0 0 1 0 (n-1)]. If t2 now fires: m3 = [(n-2) 0 1 1 0 (n-1)].

Note that it is now impossible for t4 to fire, since m(S) = n-1 implies t4 is not enabled. This is comforting, since otherwise we could fire t4 and have one process in W and one in R ... violating our requirement for mutual exclusion.

We would like to prove that the following always holds: m(W) = 0 or 1; if m(W) = 1, then m(R) = 0.

There are at least three ways to prove this:

(i) Invariant method [1]

(ii) Keller's method [2]

(iii) list all possible markings and show that the above holds for each marking (see Chap. 3).

1.2 The Invariant method

Returning to the general case, we say the marking m is reachable from m0 if and only if there is a sequence of 0 or more transition firings that change the marking of the Petri net from m0 to m.

If a Petri net has I places and J transitions, we define the IxJ incidence matrix W as follows:

If pi is an input place (and only an input place) for the transition tj, then W[i,j]= -k, where k is the marking on the arc from pi to tj.

If pi is an output place (and only an output place) for the transition tj, then W[i,j] = k, where k is the marking on the arc from tj to pi.

If pi is neither an input nor an output place for tj, then W[i,j] = 0.

If pi is both an input and an output place, then W[i,j]= k2 - k1, where k2 is the marking on the arc from tj to pi and k1 is the marking on the arc from pi to tj.

Summary: W [i,j] gives the net change in the marking at place pi caused by the firing of transition tj.

If we denote the jth column of W as $\underline{Wj}$, then if the current marking is $\underline{m1}$ and if tj fires, then the new marking $\underline{m2}$ is related to $\underline{m1}$ by:

$$\underline{m2} = \underline{m1} + \underline{Wj}$$

Further, if tk now fires we obtain $\underline{m3}$, where $\underline{m3} = \underline{m2} + \underline{Wk} = \underline{m1} + \underline{Wj} + \underline{Wk}$. In general, if $\underline{m}$ is reachable from $\underline{m0}$, then $\underline{m} = \underline{m0} + \sum_i \underline{Wj}_i$.

In matrix notation, if $\underline{m}$ is reachable from $\underline{m0}$ then there is a vector $\underline{x} >= \underline{0}$ such that $\underline{m} = \underline{m0} + W \underline{x}$.

($\underline{x}$ >= $\underline{0}$ means each entry of $\underline{x}$ is non-negative.)

In our example,

$$(1.4)\quad W = \begin{bmatrix} -1 & -1 & 0 & 0 & 1 & 1 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & -n & 1 & n \end{bmatrix}$$

and if $\underline{m}$ is reachable from $\underline{m0}$, then $\underline{m} >= \underline{0}$ and $\underline{m} = [n\ 0\ 0\ 0\ 0\ n] + W\,\underline{x}$ for some vector $\underline{x} >= \underline{0}$.

Note: In the general case, $\underline{m} >= \underline{0}$ and $\underline{m} = \underline{m0} + W\,\underline{x}$ where $\underline{x} >= \underline{0}$ is necessary, but not sufficient, for $\underline{m}$ to be reachable from $\underline{m0}$. (In Chapter 3, we show that for the above example if $\underline{m} >= \underline{0}$ and $\underline{m} = \underline{m0} + W\,\underline{x}$ for some $\underline{x}$, then $\underline{m}$ is reachable from $\underline{m0}$.)

The above vector equation leads to the Invariant method, which is based on the following theorem:

Thm. 1: If $\underline{q} \cdot W = \underline{0}$ and if $\underline{m} >= \underline{0}$ is reachable from $\underline{m0}$, then $\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$.

Pf: Merely multiply the equation

$$\underline{m} = \underline{m0} + W\,\underline{x} \quad \text{by } \underline{q}\ .$$

The Invariant method consists of 2 steps:

(1) Find the most general vector $\underline{q}$ such that $\underline{q} \cdot W = \underline{0}$. (This is a standard problem in linear algebra. Note that $\underline{q}$

involves (I-r) arbitrary constants where I = number of rows of W and r = rank of W.)

(2) Specialize the constants to obtain specific vector(s) $\underline{q}$ such that

$$\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0} \text{ is "interesting"}$$

(This step is ad hoc, but frequently obvious.)

In our example, the most general $\underline{q}$ is:

$$\underline{q} = a\cdot[1 \quad 1 \quad 1 \quad 0 \quad -(n-1) \quad -1] +$$
$$b\cdot[0 \quad 0 \quad 0 \quad 1 \quad n \quad 1]$$

where a, b are arbitrary constants. The "interesting" choice is a = 0 and b = 1 which leads to:

$$m(R) + n\cdot m(W) + m(S) = n$$

for all reachable markings $\underline{m}$. Since $\underline{m} >= \underline{0}$, the above equation implies m(W) = 0 or 1; if m(W) = 1, then m(R) = 0. I.e., we have just proved mutual exclusion for our alleged solution to the readers/writers problem.

Under certain conditions, there is a converse to Thm.1:

Thm. 2: The initial marking $\underline{m0}$ is given. Suppose that for each transition t there is a firing sequence (that depends on t) that produces a marking $\underline{mt}$ such that t is now enabled.

Claim: $\underline{q} \cdot \underline{m0} = \underline{q} \cdot \underline{m}$ for each marking reachable from $\underline{m}$ if and only if $\underline{q} \cdot W = \underline{0}$.

Proof: The "if part" is Thm. 1. "Only if": for the transition t, reach a marking $\underline{m1}$ such that t is enabled. Fire t, obtaining $\underline{m2}$ where

$\underline{m2} = \underline{m1} + \underline{Wt}$ and $\underline{Wt}$ is the column of W that corresponds to the transition t. Then $\underline{q} \cdot \underline{m1} = \underline{q} \cdot \underline{m0}$ and $\underline{q} \cdot \underline{m2} = \underline{q} \cdot \underline{m0}$ implies $\underline{q} \cdot \underline{Wt} = 0$. Doing this for each transition t, we obtain $\underline{q} \cdot W = \underline{0}$.

Checking that each transition is "enable-able" is generally easy.

Assuming each transition is, then all statements of the form "if $\underline{m}$ is reachable from $\underline{m0}$, then the components of $\underline{m}$ satisfy a1• m1 + ... + aI•mI = c (where ai, c are specified constants)" can be obtained by specializing the general solution $\underline{q}$ of $\underline{q} \cdot W = \underline{0}$.

1.3 Keller's Method

This method makes no use of linear algebra. Instead, a set of statements is asserted. To prove these assertions are always true:

(1) verify by inspection that the assertions are true when the net has the initial marking.

(2) Assuming that the assertions are currently true and assuming the transition t is enabled, show that the assertions are true after t is fired. Do this for every transition t.

Assuming (1) and (2) have been verified, it is clear that the assertions hold for each marking reachable from $\underline{m0}$.

In our example, the assertions would be:

(1.5a) m(W) <= 1

(1.5b) m(W) = 0 or m(R) = 0

Unfortunately, we cannot carry out step (2).

Suppose m(W) = 0, m(R) = 1, and t4 is enabled. Then after firing t4, (1.5b) is false.

We will see, in fact, that the assumption that m(R)>0 and t4 is enabled is impossible. This fact, though, cannot be deduced from the assertions (1.5).

The major difficulty with Keller's method is that it is non-mechanical: it is necessary to discover a superset of assertions for which it is possible to carry out (2). Then the invariance of the original assertions follows at once.

For our example, an appropriate set of assertions is:

(1.6) (i) m(W) < = 1

(ii) m(W) = 0 or m(R) = 0

(iii) m(W) = 0 implies m(S) + m(R) = n

(iv) m(W) = 1 implies m(S) = 0

(v) m(S) < = n

Pf.: If (i)-(v) are true before t1 (or t2) fires, they are true after firing since firing t1 (or t2) does not change m(R), m(W) or m(S).

If (i) - (v) hold and t3 is enabled, we conclude that before t3 fires:

1 < = m(S) (t3 is enabled)

and m(W) = 0 (from (i), (iv) and 1<=m(S)).

So, after t3 fires, it is clear that (i) - (v) still hold.

If (i)-(v) hold and t4 is enabled:

m(S) = n ((v) and t4 is enabled)

m(W) = 0 ((iv) and m(S)=n)

m(R) = 0 ((iii)and m(S)=n)

Thus, after t4 is fired, (i)-(v) are still true.

t5 and t6: exercise.

Comparison: Keller's method is stronger than the Invariant method: any result proveable using the Invariant method is proveable using Keller's method, but not conversely. (sec. 2.3) However, the Invariant method is mechanical, while Keller's method requires the ingenuity to determine a superset of assertions for which the inductive step can be carried out.

1.4 Colored Petri Nets

As an example, consider the "dining philosophers" problem discussed in [1]:

Five philosophers are seated at a circular table and between each pair of philosophers is a single fork. Each philosopher alternately thinks and eats (spaghetti, presumably). In order to eat, a philosopher must use two forks: the two forks on either side of himself. This can be modelled as a Petri net in the usual way (see Fig. 2).

Note that each philosopher has a THINK place and an EAT place; each fork has a FORK place it occupies when not in use.

Constructing the Petri net is easy but tedious. An alternative is to construct a much smaller Petri net using "colors" (Fig. 3).

Fig. 3 differs from a "plain" Petri net in the following ways:

(i) For each place, the marking is a vector instead of a scalar. eg., $\underline{m}$ (T) = [1 0 1 0 1] indicates philosophers 1, 3, 5 are currently thinking. The total marking of the net is then a vector of vectors; i.e.,

(1.7) $\underline{m} = [\ \underline{m}(T)\ \ \underline{m}(E)\ \ \underline{m}(F)]$

($\underline{m}$ is the concatenation of 3 vectors. In actuality, $\underline{m}$ has 15 components.)

(ii) Each transition can fire with any of the "colors" 1, 2, 3, 4 or 5. In Fig. 3 there are actually 10 possible firings. Note that it may happen, e.g., that t1 is enabled for color 3 but not for color 4.

(iii) Each arc is now labelled by a matrix instead of a scalar. (Unlabelled arcs have the implicit label I, where I is the identity matrix.)

Specifically, if there is an arc from place p to transition t, then the arc is labelled A(p,t) where A(p,t)[i,j] is, by definition, the change in the ith component of the marking at place p caused by firing transition t with color j.

In Fig. 3,

$$(1.8) \qquad A = B = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

In general, if $\underline{m} >= \underline{0}$ and $\underline{m}$ is reachable from $\underline{m0}$ then

(1.9) $\underline{m} = \underline{m0} + W\,\underline{x}$ for some $\underline{x} >= \underline{0}$

where W is the incidence matrix.

For Fig. 3, (1.9) is

$$(1.10)\quad \underline{m} = \underline{m0} + \begin{bmatrix} -I & I \\ I & -I \\ -A & A \end{bmatrix} \cdot \underline{x}$$

(1.9) is the same as (1.3) except that now $\underline{m}$ (and $\underline{m0}$) is a "vector whose components are themselves vectors" and W is a "matrix whose entries are themselves matrices."

In our example, $\underline{m}$ is the "3-vector" each of whose components is a "5-vector." Alternatively, we may simply view $\underline{m}$ as an "ordinary" vector having 15 components. Similarly, for W. I.e., (1.9) is <u>precisely</u> the same type of equation as (1.3). We can thus apply the Invariant method to (1.9).

For the example of Fig. 3 we want to find

(1.11) $\underline{q} = [\ \underline{q1} \quad \underline{q2} \quad \underline{q3}\ ]$ such that $\underline{q}\,W = \underline{0}$,

i.e., $-\underline{q1} + \underline{q2} - \underline{q3}\,A = \underline{0}$

and $\underline{q1} - \underline{q2} + \underline{q3}\,A = \underline{0}$. Thus,

(1.12) $\underline{q} = [(\underline{q2} - \underline{q3}\,A) \quad \underline{q2} \quad \underline{q3}\ ]$

where $\underline{q2}$, $\underline{q3}$ are arbitrary is the most general solution.

If $\underline{m} = [\ \underline{m}(T) \quad \underline{m}(E) \quad \underline{m}(F)]$ and if $\underline{m}$ is reachable from $\underline{m0}$ $=[\ \underline{1} \quad \underline{0} \quad \underline{1}\ ]$ (where $\underline{1} = [1 \quad 1 \quad 1 \quad 1 \quad 1]$), then

(1.13) $(\underline{q2} - \underline{q3}\,A)\,\underline{m}(T) + \underline{q2}\ \underline{m}(E) + \underline{q3}\ \underline{m}(F)$

$= (\underline{q2} - \underline{q3}\,A)\,\underline{1} + \underline{q3}\ \underline{1}$

Note that (1.13) can hold for all $\underline{q2}$ and $\underline{q3}$ if and only if

(1.14) $\underline{m}(T) + \underline{m}(E) = \underline{1}$ and

(1.15) $-A\,\underline{m}(T) + \underline{m}(F) = -A\,\underline{1} + \underline{1}$

The scalar equations corresponding to (1.15) are

(1.16)

$$1 + m1(F) = m1(T) + m5(T)$$

$$1 + m2(F) = m1(T) + m2(T)$$

$$1 + m3(F) = m2(T) + m3(T)$$

$$1 + m4(F) = m3(T) + m4(T)$$

$$1 + m5(F) = m4(T) + m5(T)$$

This indicates that for each pair of adjacent philosophers, at least one is thinking; i.e., no pair of adjacent philosophers can be eating simultaneously.

Summary: For a system with a high degree of regularity a colored Petri net is a more compact model than a plain Petri net. The Invariant method may be applied to a colored Petri net. The calculations appear to be easiest if we continue to view the incidence matrix as a "matrix whose entries are matrices" as in [1]. Since colored Petri nets are not substantially different from plain Petri nets, we will restrict future discussion to plain nets.

Chapter 2

2.1 Inadequacy of the Invariant Method

Consider the Petri nets in Figš 4a and 4b. Both nets have the initial marking [1 0 0]. If we use the Invariant method, we obtain for both nets:

$$(2.1) \qquad \underline{m} = [1 \quad 0 \quad 0] + \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \underline{x}$$

i.e., the Invariant method makes no distinction between the two nets. However, the nets are quite different: for the net in Fig. 4a the transition t2 can never fire, while in Fig. 4b it can fire infinitely often. This indicates that we should not apply the Invariant method to a net having a place that is both an input and an output place for the same transition. Instead, we will construct an "equivalent" net for which the input places and output places are disjoint for each transition.

One possibility:

For each trantition t and for each place p that is both an input and an output place for t, perform the transformation indicated in Fig. 5. Note that if place pi is both an input and an output place for Ni - many transitions, the above construction will produce M new places and M new transitions, where M is the sum of all Ni.

The new net P2 is equivalent to the original net P1 in the following sense:

Give both nets the same initial marking m0. (More carefully: P1 is given the marking m0; P2 is given the marking m1, where m1 is 0 on the new places and coincides with m0 on the old places.)

(i) If m is reachable from m0 on the net P1, then m is also reachable from m0 on the net P2. (Refer to Fig. 5: if, e.g., m is reached via the firing sequence t1, t2, t3, ..., tk (on P1), then m can be reached via the firing sequence t1, t1', t2, t2', t3, t3',...tk, tk' (on P2).

(ii) Suppose m' is reachable from m0 on the net P2. We define m to be the marking on P2 reached by firing every new transition in P2 that is enabled when P2 has the marking m' . (Note that m is 0 at each new place.) Claim: m is reachable from m0 on the net P1. (Suppose m (on P2) is reached (e.g.) via the firing sequence:

(2.2) t1, ''', t2, ''', t3, ''', ... tk, ''' where the primes denote a sequence of firings of new transitions. Then m is reached (on P1) via the firing sequence:

(2.3) t1, t2, t3, ... tk

A formal proof can be given using induction on k. An informal argument (see Fig. 5):Imagine that the marks in p' are actually in p, but invisible. Firing t' makes these marks visible; i.e., P2 is P1 with a "visibility delay" introduced. Comparing (2.2) and (2.3), we see that if tj on P2 is enabled, then tj on P1 is also enabled.)

Unfortunately, this construction produces a net for which the Invariant method is still inadequate:

For the net of Fig. 4a, we obtain:

(2.4) m1 + m2 = 1 for every marking reachable from [1 0 0]. (use (2.1)) This is the only invariant.

If we apply the above construction to Fig. 4a we obtain the net in Fig. 4c and the resulting equation:

$$(2.5) \qquad \underline{m} = [1 \quad 0 \quad 0 \quad 0] + \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & -2 \end{bmatrix} \cdot \underline{x}$$

The Invariant method applied to (2.5) still yields (2.4) as the only invariant.

The above result is typical. To be specific, suppose the net has I places and J transitions. Suppose only pI is both an input and an output place for the same transition. Further, suppose pI is "bad" only for the transition tJ.

Let W be the incidence matrix for the original net, P1, and let W' be the incidence matrix for the new net, P2, constructed as in Fig. 5. Then:

$$
(2.6)\quad W = \begin{array}{r} \\ \\ \\ \\ \text{row } I \end{array}
\begin{bmatrix} & \text{Col } J \\ & \\ & \\ & \\ & k2-k1 \end{bmatrix}
$$

$$
(2.7)\quad W' = \begin{array}{r} \\ \\ \\ \\ \\ \text{row } I \\ \text{row } I+1 \end{array}
\begin{bmatrix} & \text{Col } J & \text{Col } J+1 \\ & & 0 \\ & & 0 \\ & & 0 \\ & & 0 \\ & -k1 & k2 \\ 0\ \ 0 \ldots 0 \ldots 0 & k2 & -k2 \end{bmatrix}
$$

(corresponding blank entries of W and W' are equal)

The following is easily verified:

<u>Thm</u>: If [q1 q2 ... qI]W = $\underline{0}$,

then [q1 q2 ... qI qI] W' = $\underline{0}$.

Conversely, if [q1 q2 ... qI qI'] W' = $\underline{0}$,

then qI' = qI and [q1 q2 ...qI] W = $\underline{0}$.

This means that all of the invariants for P2 can be obtained from the invariants for P1 merely by replacing mI by (mI + mI'). With respect to the Invariant method, the new net P2 is not helpful. For other algebraic calculations, the new net is helpful (see chap. 3). In particular, inspection of Fig. 4a (or 4c) indicates that no transition is enabled, so only the initial marking is reachable. However, if we disregard the figure and consider only eqn. (2.1) it (erroneously) appears that the marking [0 1 1] is reachable from [1 0 0] since

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} >= \underline{0}$$

(i.e., t2 is enabled for the marking [1 0 0].)

If we consider (2.5), we see that no other marking is reachable from [1 0 0 0]. (No transition can fire since there is no column of W such that [1 0 0 0] + that column >= $\underline{0}$. I.e., no transition is enabled when the net has the marking [1 0 0 0].) Summary: The Invariant method has difficulties when the net has a place that is both an input and an output place for the same transition. Constructing the obvious equivalent net does not remove the difficulty.

2.2 Transition Systems

Transition systems are described carefully in [2]. Briefly, a transition system is a Petri net with conditions and assignments labelling some of the transitions.

In a transition system, a transition is enabled if it is enabled in the "marking sense" (i.e., each input place has a marking >=the label on the arc from the input place to the transition) and if the conditions that appear at the transition are all true. *Firing* a transition causes the usual "marking change" and the execution of the assignments that appear at the transition.

Note that the initial conditions consist of an initial marking and initial assignments.

A transition system model for the readers/writers problem is given in Fig. 6.

Keller's method (see sec. 1.3) is applicable to transition systems as well as to "pure" Petri nets. Again we have the difficulty of choosing an appropriate set of assertions. e.g., for the system of Fig. 6 we would like to prove:

(2.8) (m2 = 0 or m3 = 0) and m3 <= 1 (mutual exclusion)

Unfortunately, *this* set of assertions cannot be proved using Keller's method.

The following set of assertions *can* be proved using Keller's method:

(2.10) m2 = 0 or m3 = 0

m3 <=1

R = 0 if and only if m3 = 1

(m2 > 0 or R >= 1) implies (m2 = R-1)

Of course, since (2.10) is true it then follows immediately that (2.9) is true.

Can we avoid the problem of choosing an appropriate set of assertions?

As a first attempt, we may view the transition system as a Petri net and then use the Invariant method (i.e., we completely ignore the conditions and assignments that appear at the transitions).

For Fig. 6 we obtain:

$$(2.11)\qquad \begin{bmatrix} m1 \\ m2 \\ m3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \cdot \underline{x}$$

for each reachable marking.

We obtain as the only invariant:

$$(2.12)\qquad m1 + m2 + m3 = 1$$

In particular, we cannot prove (2.8) using the Invariant method. Ignoring the specifications at the transitions is too drastic.

As a second attempt, we will construct a Petri net "equivalent" to the original transition system.

We will restrict ourselves to transition systems that satisfy:

(i) all conditions are of the form: V r c, where V is a variable, c is a positive constant and r is one of the relational operators: >=, =, >, <, <= (but not <>).

(ii) all assignments are of the form:

V := V + c, where V is a variable and c is a constant, possibly negative.

Note that because of (ii) we will not treat the system in Fig. 6. Rather, we will treat the related system where "R:= 0" is replaced by "R: = R-1" and "R: = 1" is replaced by "R:= R + 1."

If the transition t has the condition "WHEN V >= c," do the following:

if there is no place labelled V, draw one. Draw an arc from V to t labelled c and an arc from t to V labelled c and erase the condition"WHEN V >= c."

If the transition t has the condition "WHEN V = c," we have some difficulty. In effect, a Petri net can deal only with the question "does the variable V have a value >= the constant c?" (i.e., is the marking at place V>= the label c on the arc from V to t?). Thus, we must reformulate an equality condition as an inequality condition(s).

<u>One way:</u> Introduce a new variable Vc with initial value equal to (Z minus the initial value of V.) If we can ensure that Vc + V = Z always, then the condition "V = c" is equivalent to "Vc >= Z-c and V >= c."

How should we choose Z? If we choose Z small, then the non-negativity requirements on V and Vc (we will construct places labelled Vc and V) plus the constraint Vc + V = Z will restrict V to small values. Since such a restriction on V is not necessarily imposed by the original transition system, we must not choose Z small.

Instead, we will choose Z to be an unspecified, large (but finite) integer.

Some of the transformations are given in Figs. 7a and 7b.

Note that:

(2.13) V > c can be recast as V >= c + 1

V <=c " " " " Vc >= Z-c

V < c " " " " Vc >= Z-c+1

Note that the resulting Petri net will have a place that is both an input and an output place for the same transition.

Note that the original transition system may have no explicit requirement that V be non-negative, but the "equivalent" Petri net does. This may be of no importance if the original system implicitly guarantees that V >= 0. If V >= 0 is not guaranteed, then the "equivalent" Petri net is apparently more restrictive than the original transition system.

Finally, we have not given a precise definition of "equivalent." We merely observe that on an intuitive level the above construction produces an "equivalent" system.

The Petri net corresponding to Fig. 6 is shown in Fig. 8. [For convenience, R has been drawn twice. There is actually only one place labelled R.]

If we now apply the Invariant method to Fig. 8 we obtain:

(2.14)
$$\underline{m} = \begin{bmatrix} n \\ 0 \\ 0 \\ 1 \\ Z-1 \end{bmatrix} + \begin{bmatrix} -1 & 1 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix} \cdot \underline{x}$$

where $\underline{m}$ = [m1 m2 m3 mR mRc]

By a standard calculation, we obtain the invariants:

(2.15) $m1 + m2 + m3 = n$

(2.16) $m3 + mR = m2 + 1$

(2.17) $mR + mRc = Z$

We attempted to choose transformations that would preserve (2.17), and we have succeeded.

(2.15) is really a statement that processes are neither created nor destroyed.

Clearly, there is nothing in the above invariants that will guarantee that m2 = 0 or m3 = 0; i.e., in this example the Invariant method does not succeed on the "equivalent" Petri net.

We can show that in general the transformations in Fig. 7 do not produce à Petri net for which the Invariant method is useful.

Ex.: Consider a transition system such that

(i) no place is both an input and an output place for the same transition

(ii) there are I places and J transitions

(iii) only transition J has a condition.

That condition is: "WHEN V = c."

There is no assignment at transition J.

If we ignore the condition we obtain an incidence matrix W.

Suppose we now use Fig. 7. This introduces two new places V and Vc. These are neither input nor output places for any transition except J. For transition J, V (and Vc) is both an input and an output place. The incidence matrix W for this new Petri net is

$$
\begin{array}{lc}
 & \text{Col } J \\
W' = & \left[\begin{array}{ccccccccccc} & & & & W & & & & & & \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \end{array}\right]
\begin{array}{l} \\ \text{row } (I+1) \\ \text{row } (I+2) \end{array}
\end{array}
$$

(V is the (I+1)st place; Vc is the (I+2)nd.)

If we now use sec. 2.1 to construct a Petri net such that no place is both an input and an output place for the same transition we obtain a net whose incidence matrix W" is

$$
(2.18)\quad W'' = \begin{array}{c} \\ \\ \\ \\ \\ \\ (I+1) \\ (I+2) \\ (I+3) \\ (I+4) \end{array}
\left[\begin{array}{ccccc}
 & & J & (J+1) & (J+2) \\
 & & & 0 & 0 \\
 & & & '' & '' \\
 & W & & '' & '' \\
 & & & '' & '' \\
 & & & 0 & 0 \\
0 & \ldots\ 0 & -c & c & 0 \\
0 & \ldots\ 0 & -(Z-c) & 0 & (Z-c) \\
0 & \ldots\ 0 & c & -c & 0 \\
0 & \ldots\ 0 & (Z-c) & 0 & -(Z-c)
\end{array}\right]
$$

(V' is the (I+3)rd place; Vc' is the (I+4)th)

Define $Q = \{ \underline{q} : \underline{q} \cdot W = \underline{0} \}$

and $Q' = \{ \underline{q'} : \underline{q'} \cdot W'' = \underline{0} \}$

Claim: If $\underline{q1}, \ldots, \underline{qk}$ is a basis for Q, then

$\underline{q1'}, \ldots, \underline{qk'}, \underline{r1}, \underline{r2}$ is a basis for Q' where we define

$\underline{qj'} = \underline{qj}$ with four 0 entries added at the end, j = 1, ... k

and $r1 = [0 \ldots 0 \quad 1 \quad 0 \quad 1 \quad 0]$

$r2 = [0 \ldots 0 \quad 0 \quad 1 \quad 0 \quad 1]$

(The above follows easily from the fact (easily verified) that $\underline{q'} \cdot W'' = \underline{0}$ implies $\underline{q'} = [\ \underline{q} \quad a\ b\ a\ b]$, where $\underline{q} \cdot W = \underline{0}$; a, b are constants).

This says that for our newest net the "basic invariants" are:

(i) $\underline{qj} \cdot \underline{m} = \underline{qj} \cdot \underline{m0}$

j = 1, ..., k

and

(ii) mV + mV' = K1

mVc + mVc' = K2

where K1 and K2 are constants.

(i) is a "basic invariant" set for the original transition system. If the original set (i) of basic invariants was inadequate for proof purposes, then the set consisting of (i) and (ii) is also inadequate.

2.3 Keller's Method vs the Invariant Method

Suppose we have a Petri net with initial marking m0 and incidence matrix W. The only facts deduceable using the Invariant method are:

(i) If q W = 0 and if m is reachable from m0, then q • m = q • m0

(ii) facts deduceable from facts in (i) e.g., in sec. 1.2 we derived a "type (i) fact" (2.19) m(R) + n•m (W) + m(S) = n.

From this (and the non-negativity of markings) we can deduce the "type (ii) facts"

m(W) = 0 or 1

m(W) = 1 implies [m(R) = 0 and m(S) = 0]

m(W) = 0 implies [m(R) + m(S) = n]

The above observation indicates that if we have a method X which can always be used to prove a set of facts that includes (i), then any fact that can be proved using the Invariant method can also be proved using method X; i.e. method X is "stronger" than the Invariant method.

Thm.: For any Petri net, Keller's method is stronger than the Invariant method.

Pf: Let $\underline{m0}$ be the initial marking, W the incidence matrix, and suppose $\underline{q}\ W = \underline{0}$. We must show that $\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$ (for each reachable marking $\underline{m}$) can be proved using Keller's method.

(1) $\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$ is trivially true initially.

(2) Suppose $\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$ before the enabled transition t fires. Show $\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$ after t fires. So, let $\underline{mB}$ be the marking before firing and $\underline{mA}$ the marking after firing. By hypothesis, (2.20) $\underline{q} \cdot \underline{mB} = \underline{q} \cdot \underline{m0}$.

Then (2.21) $\underline{mA} = \underline{mB} + \underline{Wt}$, where $\underline{Wt}$ is the column of W that corresponds to transition t. Multiplying (2.20) by $\underline{q}$ and recalling that $\underline{q}\ W = \underline{0}$, we obtain

$\underline{q} \cdot \underline{mA} = \underline{q} \cdot \underline{mB}$. (2.20) now implies that $\underline{q} \cdot \underline{mA} = \underline{q} \cdot \underline{m0}$

Remarks: (1) The invariant $\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$ is generally discovered more easily with the Invariant method than with Keller's method since the Invariant method is purely mechanical. The above theorem merely indicates that the invariant, once discovered, can always be proved using Keller's method.

(2) The following may be the best way to analyze a Petri net:

(i) use the Invariant method to deduce certain facts (the invariants) mechanically.

(ii) if these facts are insufficient for proof purposes, use Keller's method.

(3) The converse of the theorem is false. There are facts proveable by Keller's method that cannot be proved by the Invariant method.

<u>Example:</u> (see Fig. 9)

Using the invariant method we get:

$$\begin{bmatrix} m1 \\ m2 \\ m3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \\ 1 & -1 \\ 2 & 0 \end{bmatrix} \cdot \underline{x}$$

and the only invariant is

$$m1 + m2 = 1$$

i.e. the Invariant method can tell us <u>nothing</u> about the marking at p3.

Using Keller's method we can prove:

$$m1 + m2 = 1$$

m3 is even

$$m3 >= m2$$

Chapter 3

3.1 An Alternative Method

We have seen that the Invariant method is sometimes too weak to analyze a Petri net, while Keller's method requires us to "guess" a proper set of assertions. I.e., the Invariant method is mechanical, but not general; Keller's method is general, but not mechanical. Note that both methods typically attempt to prove statments of the form:

(3.1) "Property X holds for every marking m that is reachable from m0".

There is a third way to prove this statement: explicitly list all markings reachable from m0 and then verify by exhaustive inspection that property X holds for each marking.

Remarks: (1) There is no method stronger than the above. I.e., if (3.1) is indeed true, then (3.1) can be proven using the above technique (assuming we can list all the reachable markings).

(2) There is no need to "guess" any set of assertions. Thus, our new technique does not suffer the deficiencies of either Keller's method or the Invariant method.

(3) The method, however, does have limitations. Although there is a straightforward method for determining all reachable markings, the method does not terminate when the number of reachable markings is infinite. Even in the finite case, the time to list all reachable markings may be prohibitive.

Def.: Let P be a Petri net such that no place is both an input and an output place for the same transition. The net P has initial marking $\underline{m0}$ and incidence matrix W. We inductively define a set R($\underline{m0}$):

(i) $\underline{m0}$ is in R($\underline{m0}$)

(ii) if $\underline{m}$ is in R($\underline{m0}$) and if $\underline{Wi}$ is a column of W such that ($\underline{m}$ + $\underline{Wi}$) >= $\underline{0}$, then ($\underline{m}$ + $\underline{Wi}$) is in R($\underline{m0}$).

It is easy to see (refer to sec 1.2) that R($\underline{m0}$) is exactly the set of markings reachable fom $\underline{m0}$. A program for determining R($\underline{m0}$) is given in Appendix 1.

Example: consider the Petri net of Fig. 1.

(3.2) $\underline{m0}$ = [n 0 0 0 0 n] and

$$(3.3)\quad W = \begin{bmatrix} -1 & -1 & 0 & 0 & 1 & 1 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & -n & 1 & n \end{bmatrix}$$

It can be shown (App. 2) that the cardinality of R($\underline{m0}$) is:

(3.4) $1/6\cdot(n^3 + 9n^2 + 14n + 6)$

For n = 3, R($\underline{m0}$) was determined using App. 1 and it was verified (by inspection of each marking in R($\underline{m0}$)) that

m4 = 0 or m5 = 0; m5<=1 (mutual exclusion)

For larger n, (3.4) indicates the impracticality of App. 1. Further, App. 1 can not handle the case that n is finite, but unspecified. Generally, we need a better way to generate R($\underline{m0}$).

3.2 Modifying the Invariant Method

Note the following:

(3.5) $\underline{m}$ is reachable from $\underline{m0}$ if and only if there is a sequence

$$\underline{Wi}_j \Big|_{j=1}^{j=M}$$ of columns of W such that

$$\text{(i)} \quad \underline{m0} + \sum_{j=1}^{j=k} \underline{Wi}_j >= 0 \text{ for } 1<=k<=M$$

and $$\text{(ii)} \quad \underline{m} = \underline{m0} + \sum_{j=1}^{j=M} \underline{Wi}_j$$

(see sec. 1.1)

(The firing sequence $\{ti_j\}$ then transforms the marking from $\underline{m0}$ to $\underline{m}$.) Determining that $\underline{m}$ is reachable from $\underline{m0}$ by verifying that (i) and (ii) hold is too difficult. Instead, we will attempt to determine reachability by a two-step process.

Note that if (i) and (ii) hold, then

(3.6) $\underline{m} = \underline{m0} + W \underline{x}$ for some vector $\underline{x}$.

The converse, however, is false: the satisfaction of (3.6) does not imply $\underline{m}$ is reachable from $\underline{m0}$.

Example: (see Fig. 10)

$$\underline{m} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} \cdot \underline{x}$$

is satisfied by

$\underline{m}$ = [N (N+1) (N+1)] and $\underline{x}$ = [N N N], for every positive integer N. However, the only markings reachable from [0 1 1] are [0 1 1] and [3 0 0].

Thus, (3.6) is a necessary, but not sufficient, condition that $\underline{m}$ be reachable from $\underline{m0}$. Nevertheless, (3.6) is useful in determining reachability:

Step 1: Does there exist $\underline{x}$ such that (3.6) holds? If not, then $\underline{m}$ is not reachable from $\underline{m0}$. If so, determine $\underline{x}$, then

Step 2: see if (i) and (ii) can be satisfied where $\sum_j \underline{e}i_j = \underline{x}$. ($\underline{ek}$ has 1 in position k and 0 elsewhere.)

Summary: Step 1 is used to narrow our search for a firing sequence that will transform $\underline{m0}$ into $\underline{m}$. Step 2 then searches through this smaller list of candidates.

Thm.: There is $\underline{x}$ that satisfies (3.6) if and only if [$\underline{q} \cdot \underline{m} = \underline{q} \cdot \underline{m0}$ for every $\underline{q}$ such that $\underline{q}\ W = \underline{0}$]

Pf: easy exercise in linear algebra.

Remark: The Invariant method stops after step 1. After finding all $\underline{q}$ such that $\underline{q}\ W = \underline{0}$, the Invariant method then deals with R', where R' = { $\underline{m}$: $\underline{q}$ $\underline{m}$ = $\underline{q}$ $\underline{m0}$ for all $\underline{q}$ such that $\underline{q}$ W = $\underline{0}$ }

R' is a superset of R($\underline{m0}$), so there may be properties that hold in R($\underline{m0}$) but not in R'.

The Invariant method cannot prove these properties.

Example 1: Suppose $\underline{m0}$ and W are given by (3.2) and (3.3). Then there is $\underline{x}$ that satisfies (3.6) if and only if:

(3.7) n = m1 + m2 + m3 + m4 + m5 and

n = m4 + n • m5 + m6

If (3.7) is satisfied, then the most general $\underline{x}$ satisfying (3.6) is

(3.8)

$$\underline{x} = \begin{bmatrix} m2 + m4 + C5 \\ m3 + m5 + C6 \\ m4 + C5 \\ m5 + C6 \\ C5 \\ C6 \end{bmatrix}$$

where C5 and C6 are arbitrary constants. (Again, this is a standard linear algebra calculation.)

To recapitulate: $\underline{m}$ is not reachable from $\underline{m0}$ unless (3.7) is satisfied. If (3.7) is satisfied, then $\underline{m}$ is reachable provided we can specialize C5 and C6 in (3.8) so that there is a "legal firing sequence that sums to $\underline{x}$".

(We may assume $\underline{m}$ >= $\underline{0}$, since otherwise there is no possibility of finding a legal firing sequence that sums to $\underline{x}$.)

One choice that works : (C5 = C6 = 0)

$\underline{x} = (m2 + m4)\ \underline{e1} + (m3 + m5)\ \underline{e2}$
$+ m4\ \underline{e3} + m5\ \underline{e4}$

i.e., fire t1 (m2 + m4) times; then fire t2 (m3 + m5) times; then fire t3 (m4) times; then fire t4 (m5) times.

We have verified

Result: For the Petri net of Fig. 1, m is reachable from m0 if and only if [m >= 0 and (3.7) is satisfied.] i.e., we have determined R(m0) without using Appendix 1.

Remark: In sec. 1.2 we saw that the net of Fig. 1 could be analyzed successfully by the Invariant method. Let us consider Fig. 8, a net for which the Invariant method fails. In order that no place be both an input and an output place for the same transition, introduce new places labelled R' and Rc' and new transitions labelled t1' and t3'.
(see Fig. 5). We obtain the following:

Example 2:

(3.9) $\underline{m}$ = [m1 m2 m3 mR mR' mRc mRc']

(3.10) $\underline{m0}$ = [n 0 0 1 0 (Z-1) 0]

(3.11) W =

	t1	t1'	t2	t3	t3'	t4
1	-1	0	1	-1	0	1
2	1	0	-1	0	0	0
3	0	0	0	1	0	-1
R	-1	2	-1	-1	0	1
R'	2	-2	0	0	0	0
Rc	-1	0	1	-(Z-1)	Z	-1
Rc'	0	0	0	Z	-Z	0

Then there is a vector $\underline{x}$ satisfying (3.12) $\underline{m} = \underline{m0} + W\,\underline{x}$.

if and only if

(3.13a) $$m1 + m2 + m3 = n$$

(3.13b) $$-m2 + m3 + mR + mR' = 1$$

(3.13c) $$mR + mR' + mRc + mRc' = Z$$

Assuming $\underline{m}$ satisfies (3.13), then $\underline{x}$ satisfies (3.12) if and only if

(3.14) $$\underline{x} = \begin{bmatrix} m2 + C1 \\ m2 + C1 - 1/2 \cdot mR' \\ C1 \\ m3 + C2 \\ m3 + C2 - 1/Z \cdot mRc' \\ C2 \end{bmatrix}$$

where C1 and C2 are arbitrary constants.

Is there a choice of C1 and C2 such that the resulting $\underline{x}$ is the "sum of a legal firing sequence"? ... sometimes.

Lemma 1: If $\underline{m}$ is reachable from $\underline{m0}$ then mRc' is a multiple of Z and mR' is a multiple of 2.

Pf.: If $\underline{m}$ is reachable from $\underline{m0}$, there is a choice of C1 and C2 such that the entries of (3.14) are non-negative integers.

Lemma 2: If $\underline{m}$ is reachable fom $\underline{m0}$, then mRc' = 0 or Z.

Pf.: Use lemma 1 and (3.13c)

Lemma 3: If $\underline{m}$ is reachable from $\underline{m0}$, then

#t1 >= #t1'

#t1 >= #t2

#t3 = #t3' or #t3 = #t3' + 1

#t3 >=#t4

(where #ti = number of firings of transition ti in the firing sequence used to reach m from m0)

Pf.: Use (3.14), lemma 2, and note that

x = [#t1 #t1' #t2 #t3 #t3' #t4]

Lemma 4: If m is reachable from m0, then m2 = 0 or m3 = [illegible]

Pf.: Suppose not. Then there is an intermediate marking M such that

(i) M[illegible] M[illegible] = 0, t3 is enabled

(ii) M[illegible], M2 = 0, t1 is enabled

(see (3.11)) Note that M itself is reachable, so lemma 3 applies to M.

Suppose (i): by (3.11), M2 = #t1 - #t2 > 0 and M3 = #t3 - #t4 = 0.

Then mRc = (Z - 1) - #t1 + #t2

- (Z - 1)•#t3 + Z#t3' - #t4

(recall that mRc is initially (Z - 1))

Thus, mRc = (Z - 1) - (#t1 - #t2)

- Z•(#t3 - #t3') + #t3 - #t4 < (Z - 1)

(#t1 - #t2 = M2 > 0; #t3 >= #t3' by lemma 3;

#t3 - #t4 = M3 = 0.)

But mRc < (Z - 1) implies t3 is not enabled. (See (3.11)).

This is a contradiction.

The case (ii) is an exercise.

Lemma 5: If m is reachable from m0 then either [m2 = mR' = 0] or [m3 = mRc' = 0].

Pf.: (i) Suppose m3 = 0. Then (3.13b) implies mR + mR' = m2 + 1>0.

(3.13c) now implies mRc + mRc' < Z. Lemma 2 now implies mRc' = 0.

(ii) Suppose m3 > 0. Then lemma 4 implies m2 = 0. (3.13b) now implies mR + mR' = 0. i.e., mR' = 0.

Lemma 6: If m is reachable from m0 then m3 = 0 or 1.

Pf.: m3 > 0 implies m2 = 0.(lemma 4) (3.13b) now implies m3 = 1.

Thm. R(m0) consists precisely of:

[n	0	0	1	0	(Z-1)	0],
[(n-1)	0	1	0	0	Z	0],
[(n-1)	0	1	0	0	0	Z],
[(n-x)	x	0	(x+1-2y)	2y	(Z-1-x)	0]

where 0 <=x <=n and 0 <=2y <=1 + x.

Pf.: We know m must satisfy (3.13) and

(a) [m2 = mR' = 0] or

(b) [m3 = mRc' = 0]

Suppose (a) holds. Then

(3.13a) yields: m1 + m3 = n

(3.13b) yields: m3 + mR = 1 .

(3.13c) yields: mR + mRc + mRc' = Z

Case 1a: m3 = 0 implies mR = 1. This now implies (lemma 2) mRc' = 0.

i.e., [n 0 0 1 0 (Z-1) 0]

Case 2a: m3 = 1 implies mR = 0. This implies mRc + mRc' = Z. Using lemma 2,

[(n-1) 0 1 0 0 Z 0]

[(n-1) 0 1 0 0 0 Z]

and it is easily verified that these markings are actually reachable from m0.

Suppose (b) holds. Then

(3.13) yields: m1 + m2 = n

mR + mR' = m2 + 1

mR + mR' + mRc = Z

By lemma 1, mR' = 2y. Denoting m2 as x, the only candidates for reachability in case (b) are:

[(n-x) x 0 (x+1-2y) (2y) (Z-1-x) 0]

where 0 <= x <= n and 0 <= 2y <= x+1

That the above is actually reachable can be demonstrated via the firing sequence:

(t1, t1') fired x times; then

(t1, t2) fired y times

Remark: Finding R(m0) was not mechanical after (3.14) was derived.

3.3 Deadlock

Roughly speaking, we say deadlock is possible for a Petri net if there is a marking m', reachable from m0, and a "desirable" set of markings MD, where each marking in MD is reachable from m0 but no marking in MD is reachable from m'.

(In [1], a net is said to be "deadlock-free" if for each reachable marking m at least one transition tm is enabled. This definition is clearly inadequate.)

One way of showing deadlock is impossible: show that if m is reachable from m0, then m0 is reachable from m.

(In [2], m0 would be called a "home state.") This implies that if m1 and m2 are both reachable fom m0, then m2 is reachable from m1.

For example 1: If m is reachable from m0, then m0 is reachable from m. (See (3.2), (3.3) and (3.7) and result 1).

Proof:

We must show that if

(3.15) $\underline{m} >= \underline{0}$

$n = m1 + m2 + m3 + m4 + m5$

$n = m4 + n \cdot m5 + m6$

then $\underline{m0} = [n \quad 0 \quad 0 \quad 0 \quad 0 \quad n]$ is reachable from m.

(3.15) implies there is vector x such that

(3.16) $\underline{m0} = \underline{m} + W\,\underline{x}$

where W is given by (3.3).

In fact, the most general x satisfying (3.16) is

$$(3.17) \qquad \underline{x} = \begin{bmatrix} C1 - m2 - m4 \\ C2 - m3 - m5 \\ C1 - m4 \\ C2 - m5 \\ C1 \\ C2 \end{bmatrix}$$

where C1 and C2 are arbitrary constants.

Can we specialize C5 and C6 so that the resulting $\underline{x}$ is the "sum of a legal firing sequence" beginning with the initial state $\underline{m}$?

The "smallest" choice is C1 = m2 + m4; C2 = m3 + m5

Then $$\underline{x} = \begin{bmatrix} 0 \\ 0 \\ m2 \\ m3 \\ m2 + m4 \\ m3 + m5 \end{bmatrix}$$

The following firing sequence is legal and sums to $\underline{x}$

Fire (t6) m5 times;

fire (t5) m4 times;

fire (t3) m2 times;

fire (t5) m2 times;

(At this time we have reached

[(n-m3) 0 m3 0 0 n] beginning from [m1 m2 m3 m4 m5 m6])

fire (t4, t6) m3 times.

For example 2: If m is reachable from m0, then m0 is reachable from m.

(See (3.10), (3.11) and the theorem.)

Proof: that m0 = [n 0 0 1 0 (Z-1) 0] can be reached from the first three markings listed in the theorem is an exercise.

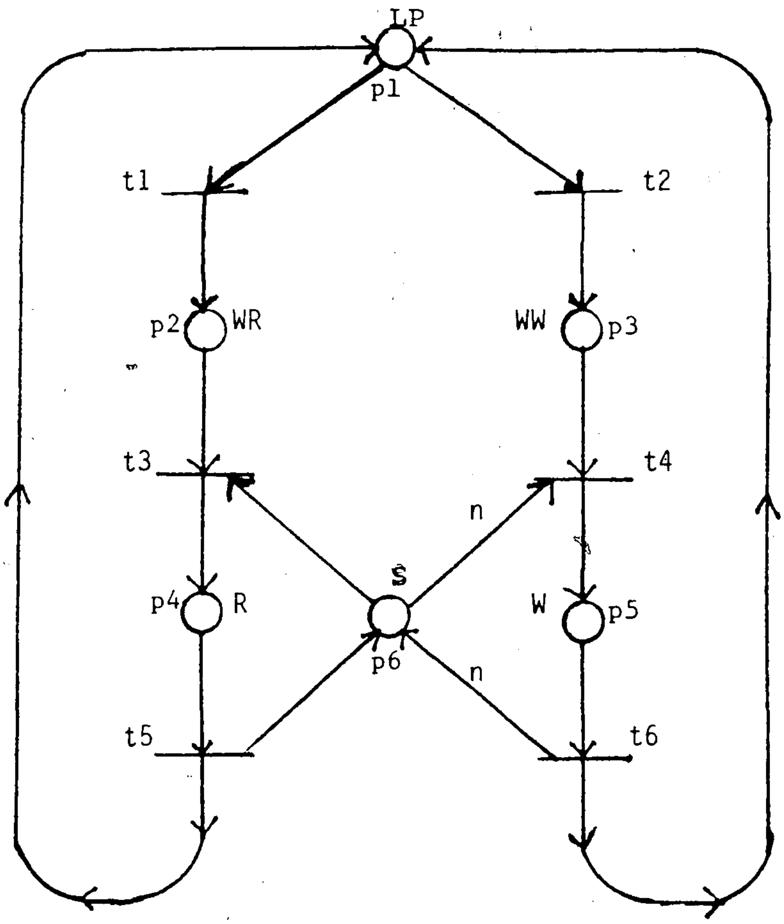
To show m0 can be reached from [(n-x) x 0 (x+1-2y) 2y (Z-1-x) 0]:

Fire (t1') y times reaching:

[(n-x) x 0 (x+1) 0 (Z-1-x) 0]

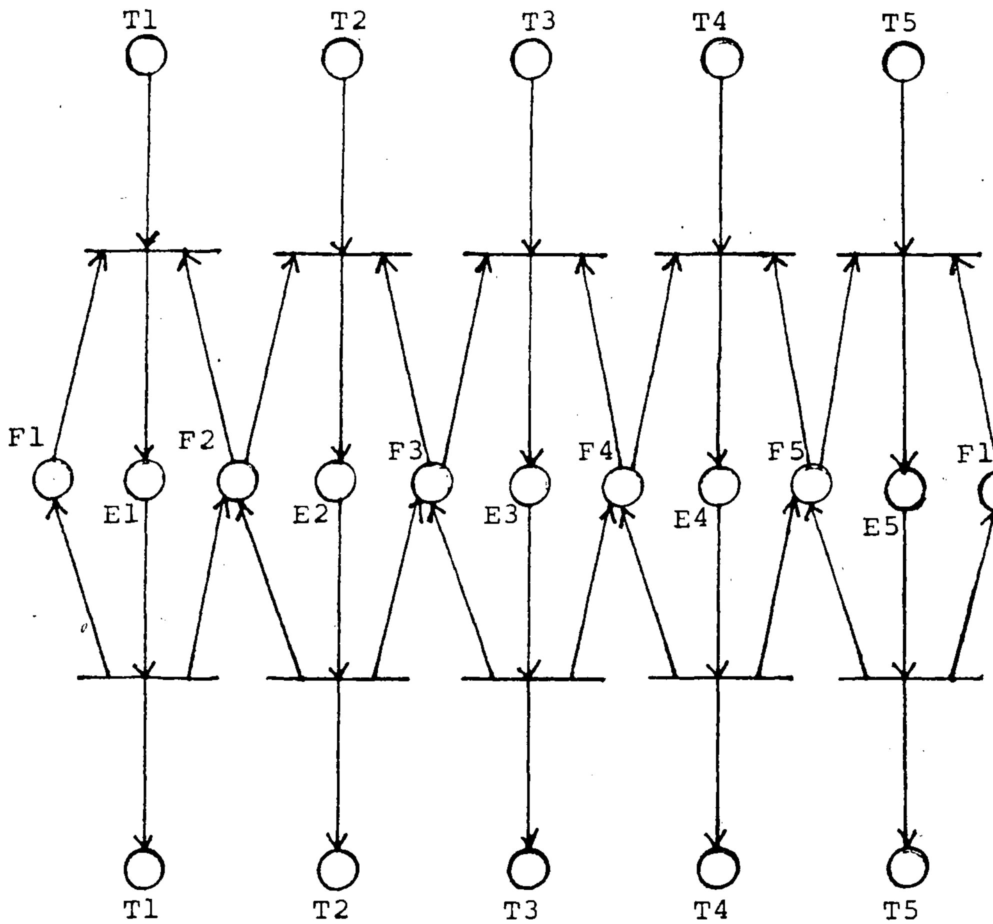
Now fire (t2) x times reaching:

[n 0 0 1 0 (Z-1) 0].



Initial marking: m0(LP) = n = m0(S)
m0 (other) = 0

Fig. 1



For convenience, several places have been drawn more than once.

Initial marking: $m(Ti) = 1$; $m(Fi) = 1$; $m(Ei) = 0$; $1 <= i <= 5$

Fig. 2

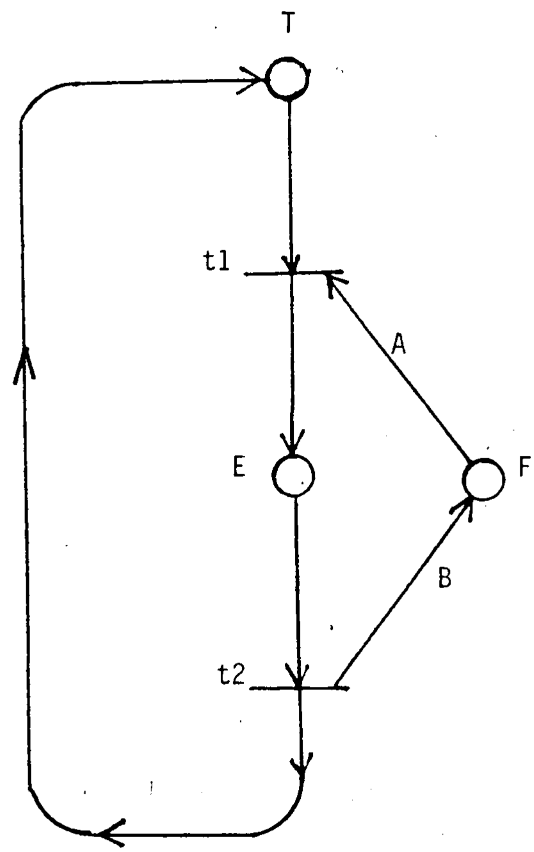
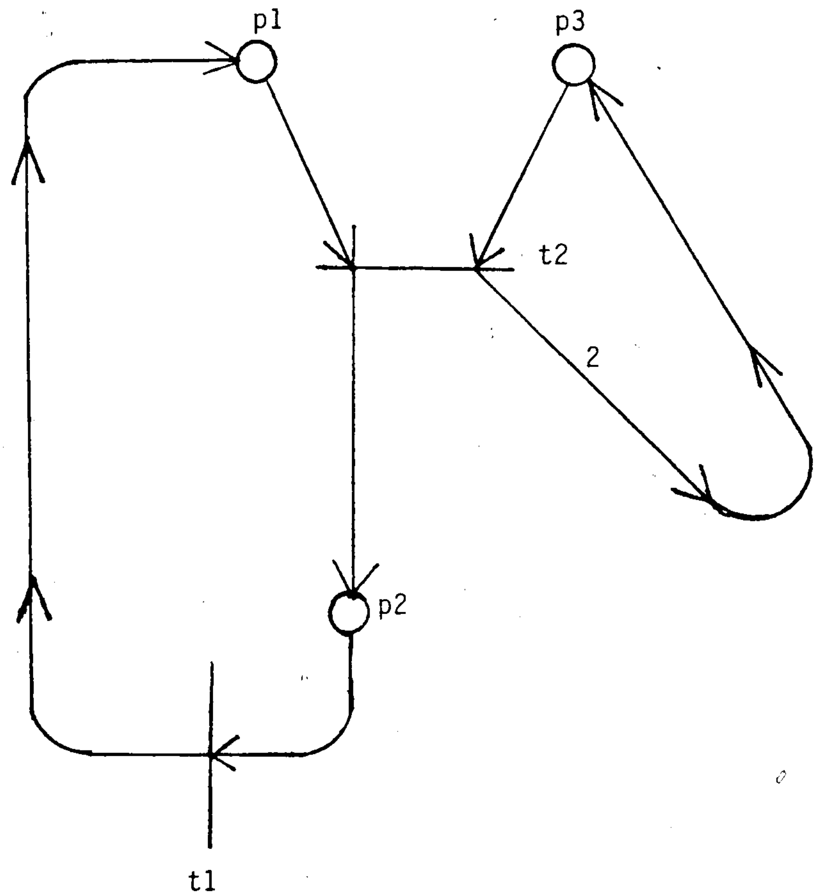
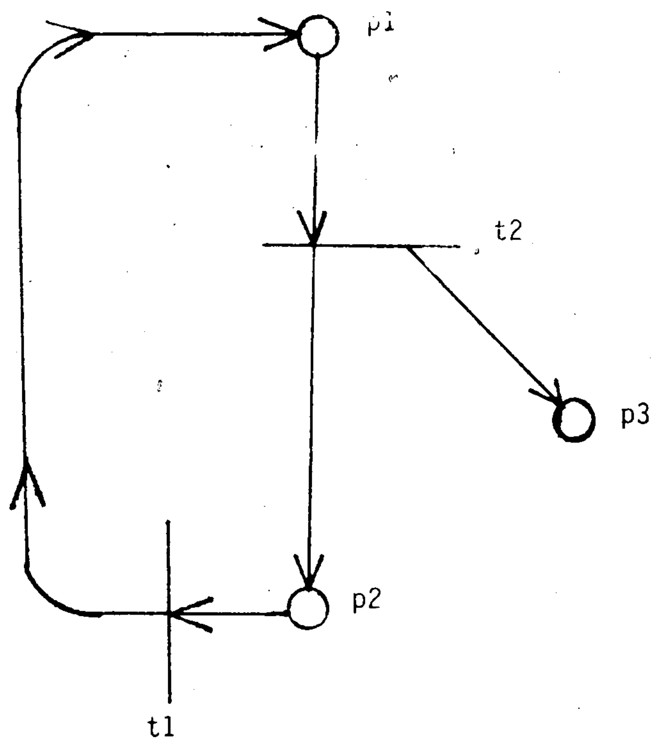


Fig. 3



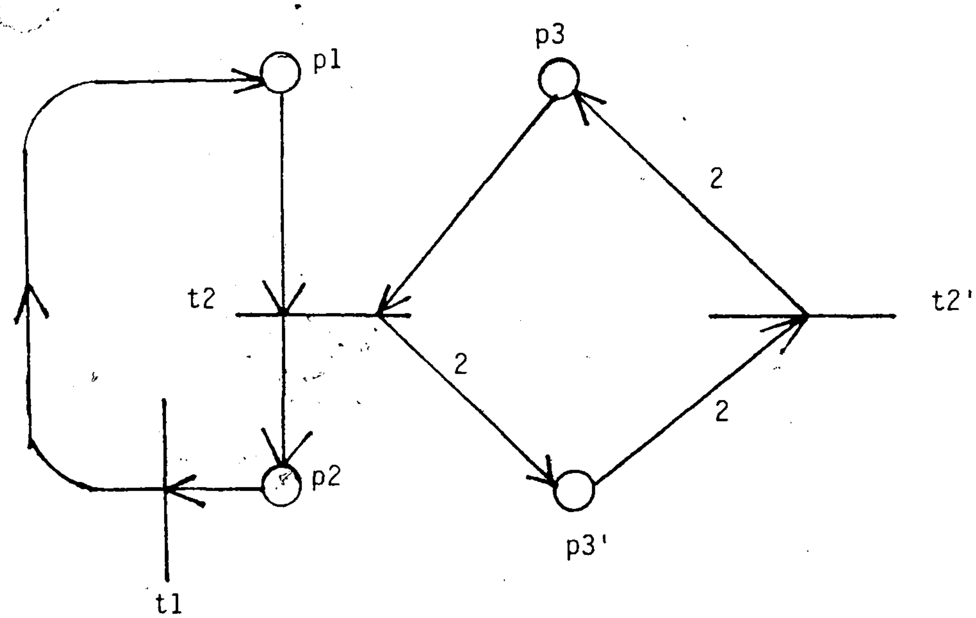
Initial marking: m1 = 1; m2 = 0 = m3

Fig. 4a



Initial marking: m1 = 1; m2 = 0 = m3

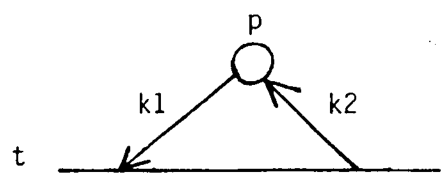
Fig. 4b



Initial marking: m1 = 1; m2 = m3 = m3' = 0

Fig. 4c

Before:



After:

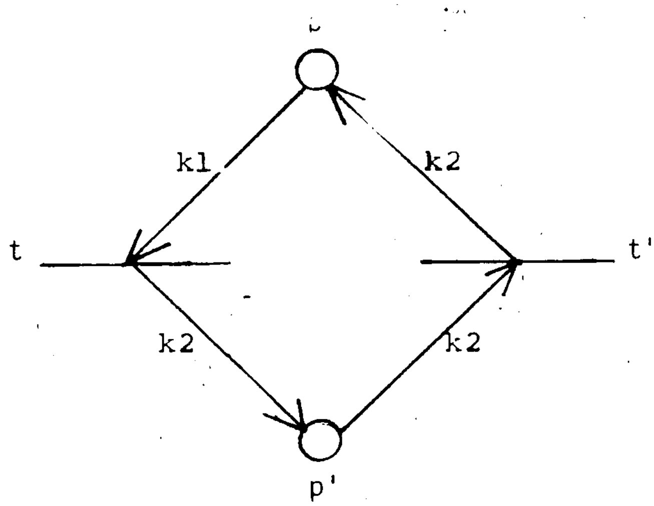
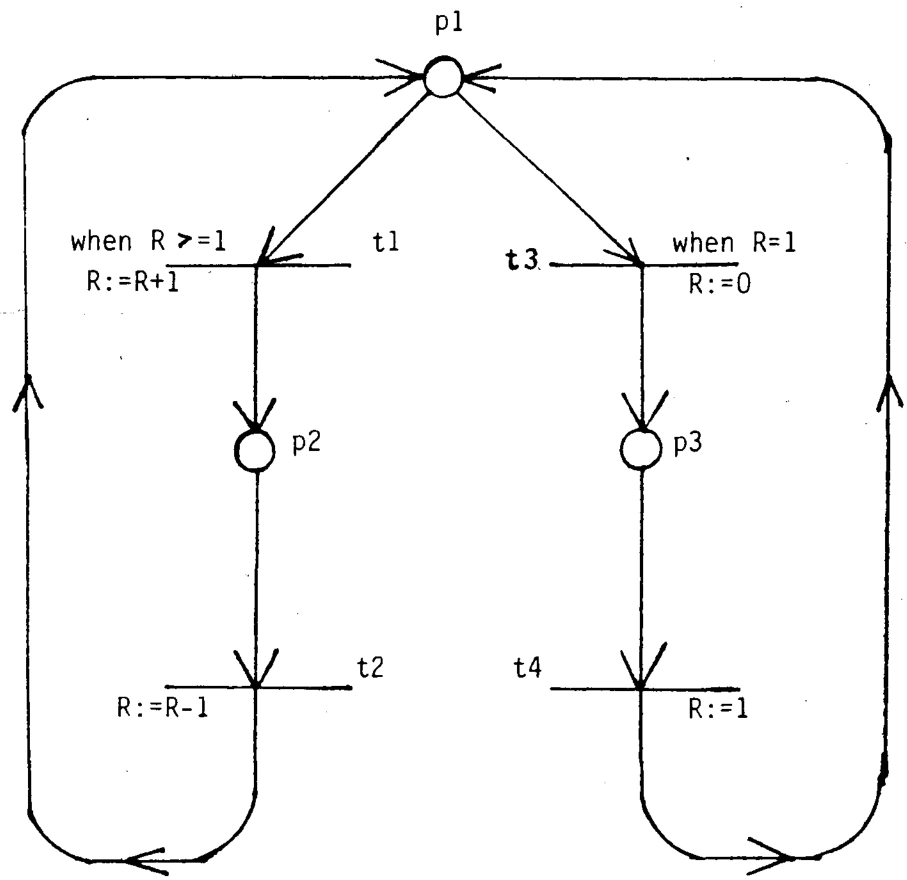


Fig. 5



Initially: m1 = n; m2 = 0 = m3; R = 1

Fig. 6.

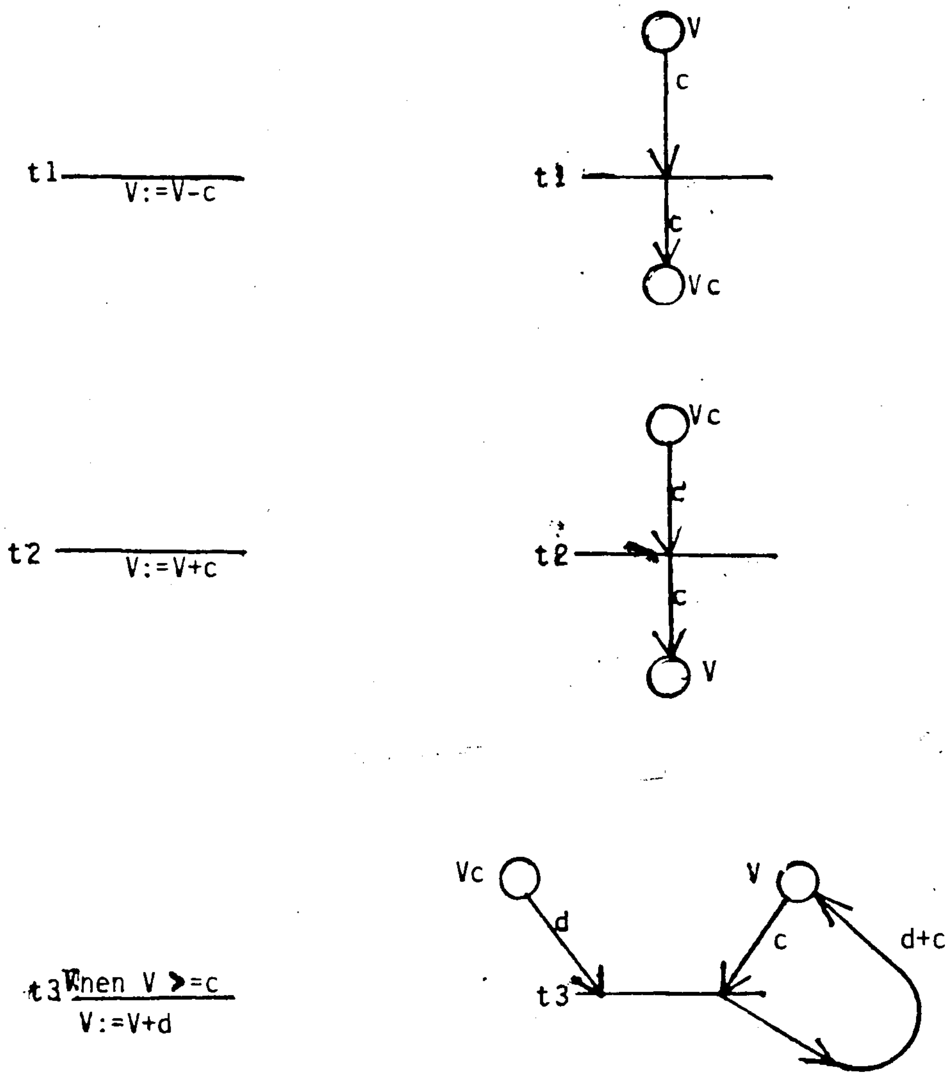
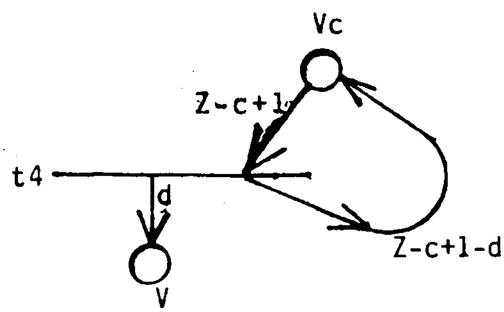


Fig. 7a

t4 When V < c
V:=V+d



t5 When V=c
V:=V-d

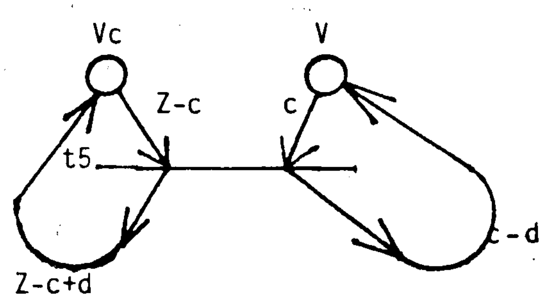
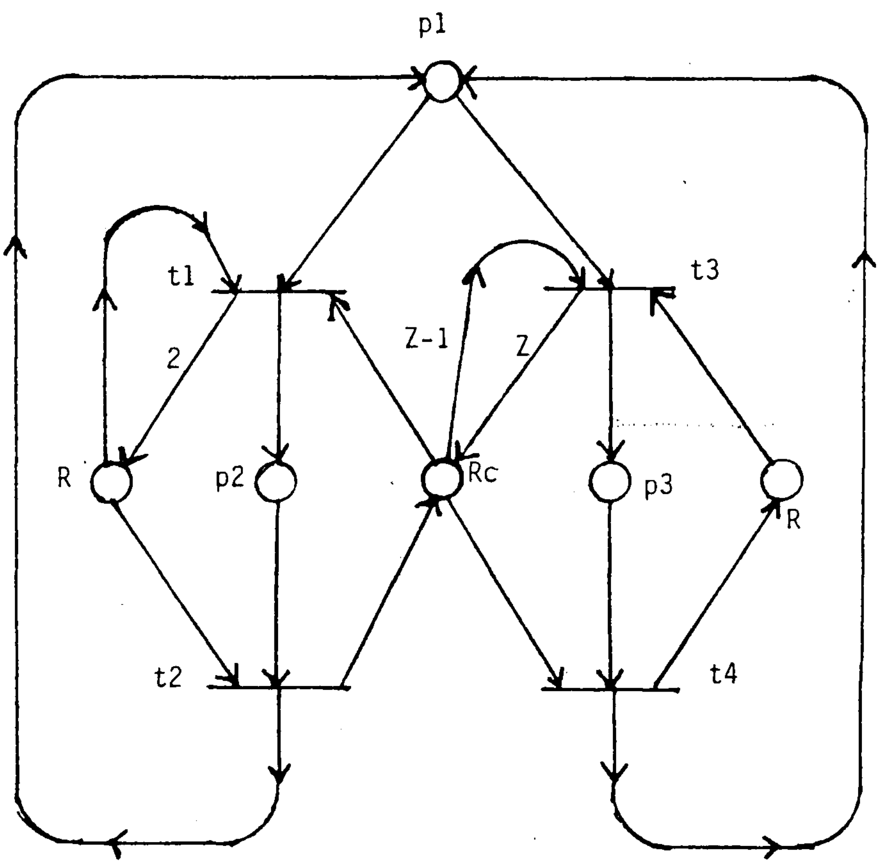


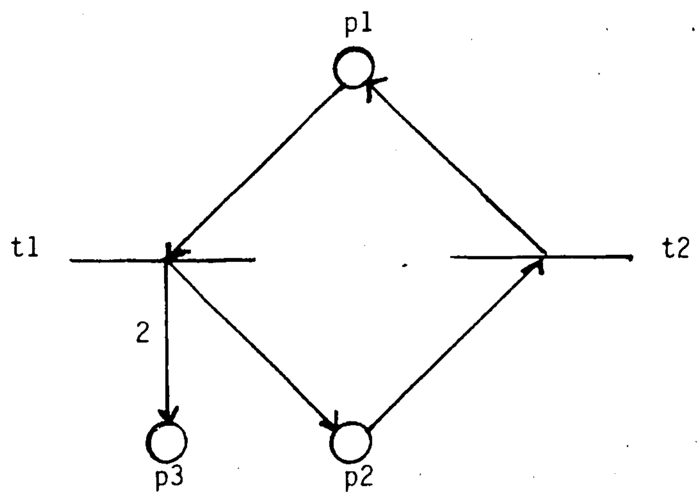
Fig. 7b



For convenience, place R has been drawn twice.

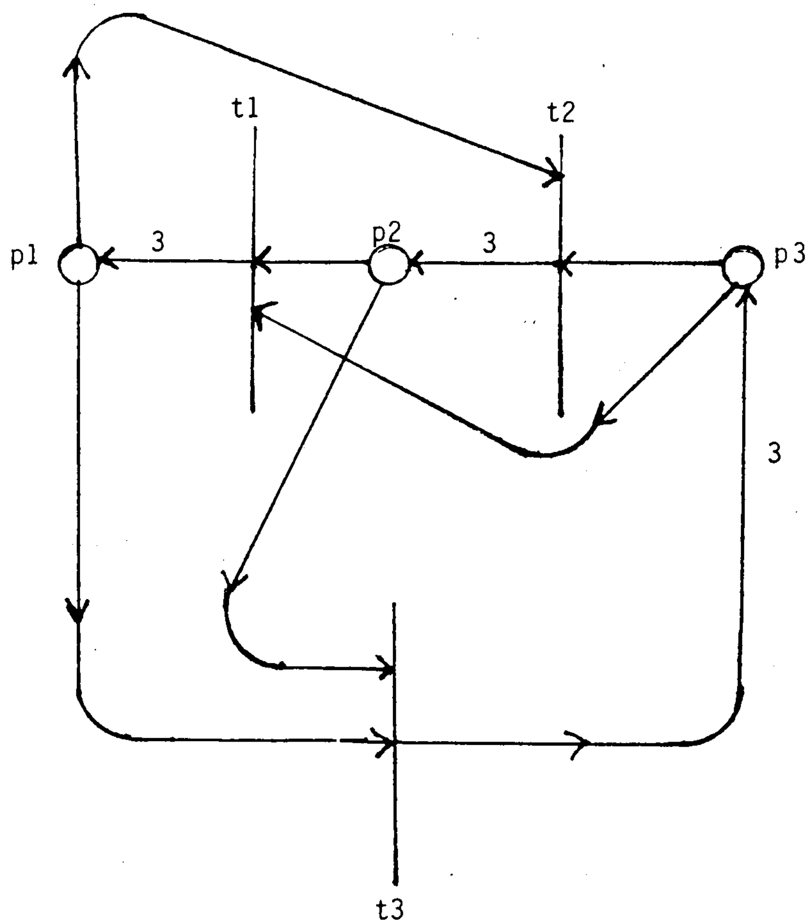
Initial marking: $m1 = n$; $m2 = 0 = m3$; $mR = 1$; $mRc = Z - 1 \gg n$

Fig. 8



Initial marking: m1 = 1; m2 = 0 = m3

Fig. 9



Initial marking: m1 = 0; m2 = 1 = m3

Fig. 10

Bibliography

1. Jensen, K. Coloured Petri nets and the Invariant-method, Theor. Comp. Sci. 14 (1981) 317-336.

2. Keller, R. M. Formal verification of parallel programs. Comm. ACM 19, 7 (July 1976), 371-384.

3. Peterson, J. L. Petri nets, Comput. Surveys 9(3) (1977), 223-252.

Appendix 1

The following program will generate all markings in Reach (m0). The program terminates with list = Reach (m0)if and only if Reach (m0) is a finite set.

```
BEGIN
     Initialize a queue to empty;
     Initialize a list to empty;
     Enqueue ( m0 );
     WHILE the queue is not empty DO
        BEGIN
          Dequeue ( m );
          Append ( m ) to the list;
          FOR i:= 1 to number of columns of W do
                 IF ( m + Wi >= 0) AND ( m + Wi ) is neither
                    in the list nor in the queue
                       THEN   Enqueue ( m + Wi )
        END
END.
```

Apendix 2

Def. 1: We define F(n) = number of distinct vectors m = [m1 m2 m3 m4 m5 m6] that satisfy:

(i) mi is a non-negative integer for each i

(ii) n = m1 + m2 + m3 + m4 + m5

(iii) n = m4 + n•m5 + m6

(n > 0)

Def.2: We define G(n) = number of distinct vectors m = [m1 m2 m3]

that satisfy:

(i) mi is a non-negative integer for each i

(ii) n = m1 + m2 + m3

(n >= 0)

Lemma 1: For each n > = 0, G(n) = (n + 1)•(n + 2)/2

Proof: Assign to m3 the integer value j, where 0 <= j <= n. Then (n - j) units remain for assignment to m1 and m2. This latter assignment can be done in (n - j + 1) ways.

$$\text{Thus, } G(n) = \sum_{j=0}^{n} (n - j + 1)$$

$$= (n + 1) \cdot (n + 2)/2$$

Lemma 2: For each n > 0, $F(n) = G(n - 1) + \sum_{k=0}^{n} G(n - k)$

Proof: (i) and (iii) of Def. 1 imply that m5 may have only the values 0 or 1. If we assign m5 the value 1, then we must assign m4 and m6 the value 0. We must then assign m1, m2, m3 values such that n = m1 + m2 + m3 + 1.

This latter assignment can be done in G(n - 1) ways.

If we assign m5 the value 0, then m6 = n - m4 and n = m1 + m2 + m3 + m4.

Assign m4 the value k, where p < = k < = n. Now we must assign m1, m2, m3 values such that n = m1 + m2 + m3 + k. This latter assignment can be done in G(n - k) many ways.

Thm.: For each n > = 0, $F(n) = (n^3 + 9n^2 + 14n + 6)/6$

Proof: The above is clearly true for n = 0. For n> 0, lemmas 1 and 2 imply $F(n) = n(n + 1)/2 + \sum_{k=0}^{n} (n - k + 1)(n - k + 2)/2$

Now recall: $\sum_{k=0}^{n} k^2 = n(n + 1)(2n + 1)/6$

Vita

William J. Seaman was born to William C. and Margaret A. Seaman on Feb. 19, 1946 in Bethlehem, Pa. He received a B.S. in Engineering Mechanics from Lehigh U. in 1968 and a Ph.D. in Applied Mathematics from MIT in 1973. He is currently a member of the faculty of Muhlenberg College.