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

Seaman, William Joseph
1984
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 '44_

[Signature]

Professor in Charge

[Signature]

Chairman of Department
The author deeply appreciates the guidance and encouragement of Dr. Gerhard Rayna.
Table of Contents

Chapter 1 .................................................. 2
Chapter 2 .................................................. 14
Chapter 3 .................................................. 29
Figures ..................................................... 42
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.
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 $p_i$ is called an input place for the transition $t_j$ if there is an arc from $p_i$ to $t_j$. It is called an output place if there is an arc from $t_j$ to $p_i$.

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

A transition $t_j$ 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 $p_i$ and $t_j$. (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

\[(1.1) \quad m = [m(LP) m(WR) m(WW) m(R) m(W) m(S)]\]

\[\text{and, eg, } m(R) \text{ is the marking at place } R.\]

Thus, the initial marking \(m_0\) is

\[(1.2) \quad m_0 = [n \ 0 \ 0 \ 0 \ 0 \ n]\]

Initially, only the transitions \(t_1\) and \(t_2\) are enabled.
Suppose \(t_1\) fires. Then we obtain the new marking:

\[m_1 = [(n-1) \ 1 \ 0 \ 0 \ 0 \ n] \quad \text{At this point, } t_2 \text{ and } t_3 \text{ are} \]
enabled. Suppose $t_3$ fires. We obtain the new marking $m_2 = [(n-1)\ 0\ 0\ 1\ 0\ (n-1)]$. If $t_2$ now fires: $m_3 = [(n-2)\ 0\ 1\ 1\ 0\ (n-1)]$.

Note that it is now impossible for $t_4$ to fire, since $m(S) = n-1$ implies $t_4$ is not enabled. This is comforting, since otherwise we could fire $t_4$ 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 $m_0$ if and only if there is a sequence of $0$ or more transition firings that change the marking of the Petri net from $m_0$ to $m$.

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

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

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

If $p_i$ is neither an input nor an output place for $t_j$, then $W[i,j] = 0$.

If $p_i$ is both an input and an output place, then $W[i,j] = k_2 - k_1$, where $k_2$ is the marking on the arc from $t_j$ to $p_i$ and $k_1$ is the marking on the arc from $p_i$ to $t_j$.

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

If we denote the $j$th column of $W$ as $W_j$, then if the current marking is $m_1$ and if $t_j$ fires, then the new marking $m_2$ is related to $m_1$ by:

$$m_2 = m_1 + W_j$$

Further, if $t_k$ now fires we obtain $m_3$, where $m_3 = m_2 + W_k = m_1 + W_j + W_k$. In general, if $m$ is reachable from $m_0$, then $m = m_0 + \sum W_j$.

In matrix notation, if $m$ is reachable from $m_0$ then there is a vector $x \geq 0$ such that $m = m_0 + Wx$.

($x \geq 0$ means each entry of $x$ is non-negative.)
In our example,

\[
W = \begin{bmatrix}
-1 & -1 & 0 & 0 & 1 & 1 \\
1 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & -1 & -n & 1 & n
\end{bmatrix}
\]

and if \( m \) is reachable from \( m_0 \), then \( m \geq 0 \) and

\[
m = [n \ 0 \ 0 \ 0 \ 0 \ n] + Wx \text{ for some vector } x \geq 0.
\]

Note: In the general case, \( m \geq 0 \) and \( m = m_0 + Wx \) where \( x \geq 0 \) is necessary, but not sufficient, for \( m \) to be reachable from \( m_0 \). (In Chapter 3, we show that for the above example if \( m \geq 0 \) and \( m = m_0 + Wx \) for some \( x \), then \( m \) is reachable from \( m_0 \).)

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

**Thm. 1:** If \( q \cdot W = 0 \) and if \( m \geq 0 \) is reachable from \( m_0 \),

then \( q \cdot m = q \cdot m_0 \).

**Pf:** Merely multiply the equation

\[
m = m_0 + Wx
\]

by \( q \).

The Invariant method consists of 2 steps:

1. Find the most general vector \( q \) such that \( q \cdot W = 0 \).
2. (This is a standard problem in linear algebra. Note that \( q \)
involves \((I-r)\) arbitrary constants where \(I = \text{number of rows of } W\)
and \(r = \text{rank of } W\).)

(2) Specialize the constants to obtain specific vector(s) \(q\) such that
\[
g \cdot m = q \cdot m_0 \text{ is } "\text{interesting}"\]
(This step is ad hoc, but frequently obvious.)

In our example, the most general \(q\) is:
\[
g = a \cdot [1 \ 1 \ 1 \ 0 \ -(n-1) \ -1] + b \cdot [0 \ 0 \ 0 \ 1 \ n \ 1] \]
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 \(m\). Since \(m \geq 0\), 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 \(m_0\) is given. Suppose that for each transition \(t\) there is a firing sequence (that depends on \(t\)) that produces a marking \(m_t\) such that \(t\) is now enabled.

Claim: \(g \cdot m_0 = g \cdot m\) for each marking reachable from \(m\) if and only if \(g \cdot W = 0\).

Proof: The "if part" is Thm. 1. "Only if": for the transition \(t\), reach a marking \(m_1\) such that \(t\) is enabled. Fire \(t\), obtaining \(m_2\) where
\[ m_2 = m_1 + W_t \] and \( W_t \) is the column of \( W \) that corresponds to the transition \( t \). Then \( q \cdot m_1 = q \cdot m_0 \) and \( q \cdot m_2 = q \cdot m_0 \) implies \( q \cdot W_t = 0. \) Doing this for each transition \( t \), we obtain \( q \cdot W = 0. \)

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

Assuming each transition is, then all statements of the form "if \( m \) is reachable from \( m_0 \), then the components of \( m \) satisfy \( a_1 \cdot m_1 + \ldots + a_r \cdot m_r = c \) (where \( a_i, c \) are specified constants)" can be obtained by specializing the general solution \( q \cdot W = 0. \)

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 \( m_0 \).

In our example, the assertions would be:

1.5a) \( m(W) \leq 1 \)

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

Unfortunately, we cannot carry out step (2).
Suppose \( m(W) = 0 \), \( m(R) = 1 \), and \( t_4 \) is enabled. Then after firing \( t_4 \), (1.5b) is false.

We will see, in fact, that the assumption that \( m(R) > 0 \) and \( t_4 \) 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) \( m(W) \leq 1 \)  

(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) \( m(S) \leq n \)  

Pf.: If (i)-(v) are true before \( t_1 \) (or \( t_2 \)) fires, they are true after firing since firing \( t_1 \) (or \( t_2 \)) does not change \( m(R) \), \( m(W) \) or \( m(S) \).

If (i) - (v) hold and \( t_3 \) is enabled, we conclude that before \( t_3 \) fires:

\[ 1 \leq m(S) \text{ (} t_3 \text{ is enabled)} \]

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

So, after \( t_3 \) fires, it is clear that (i) - (v) still hold.

If (i)-(v) hold and \( t_4 \) is enabled:

\[ m(S) = n \text{ (} (v) \text{ and } t_4 \text{ is enabled)} \]
\[ m(w) = 0 \quad ((iv) \text{ and } m(S)=n) \]
\[ m(R) = 0 \quad ((iii)\text{and } m(S)=n) \]

Thus, after \( t_4 \) 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., \( m(T) = [1 \ 0 \ 1 \ 0 \ 1] \) indicates philosophers 1, 3, 5 are currently thinking. The total marking of the net is then a vector of vectors; i.e.,

\[
(1.7) \quad m = [ \ m(T) \ m(E) \ m(F) ]
\]

( \( m \) is the concatenation of 3 vectors. In actuality, \( m \) 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 \( t_1 \) 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 \( i \)th component of the marking at place \( p \) caused by firing transition \( t \) with color \( j \).

In Fig. 3,

\[
A = B = \begin{bmatrix}
1 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 \\
\end{bmatrix}
\]

In general, if \( m \geq 0 \) and \( m \) is reachable from \( m_0 \) then
\[ (1.9) \quad \mathbf{m} = \mathbf{m}_0 + \mathbf{W} \mathbf{x} \quad \text{for some } \mathbf{x} \geq 0 \]

where \( \mathbf{W} \) is the incidence matrix.

For Fig. 3, (1.9) is

\[
(1.10) \quad \mathbf{m} = \mathbf{m}_0 + \begin{bmatrix} -\mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{bmatrix} \cdot \mathbf{x}
\]

(1.9) is the same as (1.3) except that now \( \mathbf{m} \) (and \( \mathbf{m}_0 \)) is a "vector whose components are themselves vectors" and \( \mathbf{W} \) is a "matrix whose entries are themselves matrices."

In our example, \( \mathbf{m} \) is the "3-vector" each of whose components is a "5-vector." Alternatively, we may simply view \( \mathbf{m} \) as an "ordinary" vector having 15 components. Similarly, for \( \mathbf{W} \). I.e., (1.9) is precisely 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

\[
(1.11) \quad \mathbf{q} = [ q_1 \quad q_2 \quad q_3 ] \quad \text{such that } \mathbf{q} \mathbf{W} = \mathbf{0}
\]
i.e., \[ -q_1 + q_2 - q_3 \mathbf{A} = 0 \]
and \[ q_1 - q_2 + q_3 \mathbf{A} = 0 \]. Thus,

\[
(1.12) \quad \mathbf{q} = [(q_2 - q_3 \mathbf{A}) \quad q_2 \quad q_3]
\]
where \( q_2, q_3 \) are arbitrary is the most general solution.

If \( \mathbf{m} = [ \mathbf{m}(T) \quad \mathbf{m}(E) \quad \mathbf{m}(F) ] \) and if \( \mathbf{m} \) is reachable from \( \mathbf{m}_0 \)

\[
(1.13) \quad (q_2 - q_3 \mathbf{A}) \mathbf{m}(T) + q_2 \mathbf{m}(E) + q_3 \mathbf{m}(F)
\]

\[ = (q_2 - q_3 \mathbf{A}) \mathbf{1} + q_3 \mathbf{1} \]
Note that (1.13) can hold for all \( q_2 \) and \( q_3 \) if and only if

\[
(1.14) \quad m(T) + m(E) = 1 \quad \text{and} \quad
(1.15) \quad -A m(T) + m(F) = -A 1 + 1
\]

The scalar equations corresponding to (1.15) are

\[
(1.16)
\begin{align*}
1 + m_1(F) & = m_1(T) + m_5(T) \\
1 + m_2(F) & = m_1(T) + m_2(T) \\
1 + m_3(F) & = m_2(T) + m_3(T) \\
1 + m_4(F) & = m_3(T) + m_4(T) \\
1 + m_5(F) & = m_4(T) + m_5(T)
\end{align*}
\]

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.
2.1 Inadequacy of the Invariant Method

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

\[
\begin{pmatrix}
1 & -1 \\
-1 & 1 & \ldots & 1 \\
0 & 1
\end{pmatrix}
\]

\[m = [1 \ 0 \ 0] + \begin{pmatrix}
-1 & 1 \\
0 & 1
\end{pmatrix} \cdot x\]

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 \(t_2\) 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 transition \(t\) 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 \(p_i\) is both an input and
an output place for \(N_i\) - many transitions, the above construction
will produce \(M\) new places and \(M\) new transitions, where \(M\) is the
sum of all \(N_i\).

The new net \(P_2\) is equivalent to the original net \(P_1\) in the
following sense:

Give both nets the same initial marking $m_0$. (More carefully: $P_1$ is given the marking $m_0$; $P_2$ is given the marking $m_1$, where $m_1$ is 0 on the new places and coincides with $m_0$ on the old places.)

(i) If $m$ is reachable from $m_0$ on the net $P_1$, then $m$ is also reachable from $m_0$ on the net $P_2$. (Refer to Fig. 5: if, e.g., $m$ is reached via the firing sequence $t_1, t_2, t_3, \ldots, tk$ (on $P_1$), then $m$ can be reached via the firing sequence $t_1, t_1', t_2, t_2', t_3, t_3', \ldots tk, tk'$ (on $P_2$).

(ii) Suppose $m'$ is reachable from $m_0$ on the net $P_2$. We define $m$ to be the marking on $P_2$ reached by firing every new transition in $P_2$ that is enabled when $P_2$ has the marking $m'$. (Note that $m$ is 0 at each new place.) Claim: $m$ is reachable from $m_0$ on the net $P_1$. (Suppose $m$ (on $P_2'$) is reached (e.g.) via the firing sequence:

$$(2.2) \; t_1, 't_1', t_2, 't_2', t_3, 't_3', \ldots tk, 'tk'$$

where the primes denote a sequence of firings of new transitions. Then $m$ is reached (on $P_1$) via the firing sequence:

$$(2.3) \; t_1, t_2, t_3, \ldots tk$$

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., $P_2$ is $P_1$ with a "visibility delay" introduced. Comparing (2.2) and (2.3), we see that if $t_j$ on $P_2$ is enabled, then $t_j$ on $P_1$ 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:

\[(2.4) \ m_1 + m_2 = 1 \text{ for every marking reachable from } [1 \ 0 \ 0]. \text{ (use } (2.1)) \text{ This is the only invariant.}\]

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

\[
\begin{bmatrix}
1 & -1 & 0 \\
-1 & 1 & 0 \\
0 & -1 & 2 \\
0 & 2 & -2
\end{bmatrix} \mathbf{x}
\]

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) \[ W = \begin{bmatrix} \text{row I} & \text{col J} \end{bmatrix} \]

\[ \begin{array}{c c c c c}
\text{col J} & & & & \\
\text{col J+1} & & & & \\
k2-k1 & & & & \\
\end{array} \]

(2.7) \[ W' = \begin{bmatrix} \text{row I} & \text{row I+1} \end{bmatrix} \]

\[ \begin{array}{c c c c c c c c c c c}
\text{col J} & \text{col J+1} & & & & & & & & & \\
0 & & & & & & & & & & 0 \\
0 & & & & & & & & & & 0 \\
0 & & & & & & & & & & 0 \\
k2 & -k1 & k2 & & & & & & & & \\
\end{array} \]

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

The following is easily verified:

**Thm:** If \([q_1 q_2 \ldots q_I] W = 0\),

then \([q_1 q_2 \ldots q_I q_I'] W' = 0\).

Conversely, if \([q_1 q_2 \ldots q_I q_I'] W' = 0\),

then \(q_I' = q_I\) and \([q_1 q_2 \ldots q_I] W = 0\).

This means that all of the invariants for \( P_2 \) can be obtained from the invariants for \( P_1 \) 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

\[
\begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix} + \begin{bmatrix}
-1 \\
1 \\
1
\end{bmatrix} \geq 0
\]

(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 \(\geq 0\). 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:

\[
\begin{bmatrix}
  m_1 \\
  m_2 \\
  m_3
\end{bmatrix}
= \begin{bmatrix}
  1 \\
  0 \\
  0
\end{bmatrix}
+ \begin{bmatrix}
  -1 & 1 & -1 & 1 \\
  1 & -1 & 0 & 0 \\
  0 & 0 & 1 & -1
\end{bmatrix} x
\]

(2.11)

for each reachable marking.

We obtain as the only invariant:

(2.12) \[ m_1 + m_2 + m_3 = 1 \]

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: \( \geq, =, >, <, \leq \) (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).

One way: 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:

\[(2.13) \quad V > c \text{ can be recast as } V \geq c + 1\]
\[V \leq c \quad \text{is equivalent to } V \leq c, \quad \text{and } Vc \geq Z-c\]
\[V < c \quad \text{is equivalent to } V < c, \quad \text{and } Vc \geq Z-c+1\]

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 \( V \geq 0 \). If \( V \geq 0 \) 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:

\[
\begin{bmatrix}
 n \\
 0 \\
 1
\end{bmatrix} + \begin{bmatrix}
 -1 & 1 & -1 & 1 \\
 0 & 0 & 1 & -1 \\
 1 & 1 & -1 & 1 \\
 Z-1 \\
 -1 & 1 & 1 & -1
\end{bmatrix} \cdot \begin{bmatrix}
 m_1 \\
 m_2 \\
 m_3 \\
 m_R \\
 m_{RC}
\end{bmatrix}
\]

where \( m = [m_1 \ m_2 \ m_3 \ m_R \ m_{RC}] \)

By a standard calculation, we obtain the invariants:

(2.15) \( m_1 + m_2 + m_3 = n \)

(2.16) \( m_3 + m_R = m_2 + 1 \)

(2.17) \( m_R + m_{RC} = Z \)

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 \( m_2 = 0 \) or \( m_3 = 0 \); 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 a 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

\[
W' = \begin{bmatrix}
\text{Col J} \\
\text{row (I+1)} & 0 & 0 & \ldots & \ldots & \ldots & 0 \\
\text{row (I+2)} & 0 & 0 & \ldots & \ldots & \ldots & 0
\end{bmatrix}
\]

(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
Define $Q = \{ q : q \cdot W = 0 \}$ and $Q' = \{ q' : q' \cdot W'' = 0 \}$.

Claim: If $q_1, \ldots, q_k$ is a basis for $Q$, then $q_1', \ldots, q_k', r_1, r_2$ is a basis for $Q'$ where we define $q_j' = q_j$ with four 0 entries added at the end, $j = 1, \ldots, k$ and $r_1 = [0 \ldots 0 1 0 1 0]$ and $r_2 = [0 \ldots 0 0 1 0 1]$

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

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

(i) $q_j \cdot m = q_j \cdot m_0$

$j = 1, \ldots, k$

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 deducible using the Invariant method are:

(i) If \( g \cdot W = 0 \) and if \( m \) is reachable from \( m0 \), then \( g \cdot m = g \cdot m0 \)

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

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

\[ m(W) = 0 \text{ or } 1 \]

\[ m(W) = 1 \implies [m(R) = 0 \text{ and } m(S) = 0] \]

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

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 \( m_0 \) be the initial marking, \( W \) the incidence matrix, and suppose \( q \cdot W = 0 \). We must show that \( q \cdot m = q \cdot m_0 \) (for each reachable marking \( m \)) can be proved using Keller's method.

1. \( q \cdot m = q \cdot m_0 \) is trivially true initially.

2. Suppose \( q \cdot m = q \cdot m_0 \) before the enabled transition \( t \) fires. Show \( q \cdot m = q \cdot m_0 \) after \( t \) fires. So, let \( m_B \) be the marking before firing and \( m_A \) the marking after firing. By hypothesis, (2.20) \( q \cdot m_B = q \cdot m_0 \).

Then (2.21) \( m_A = m_B + W_t \), where \( W_t \) is the column of \( W \) that corresponds to transition \( t \). Multiplying (2.20) by \( q \) and recalling that \( q \cdot W = 0 \), we obtain

\[
q \cdot m_A = q \cdot m_B.
\]

(2.20) now implies that \( q \cdot m_A = q \cdot m_0 \)

Remarks: (1) The invariant \( q \cdot m = q \cdot m_0 \) 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.

Example: (see Fig. 9)

Using the invariant method we get:

\[
\begin{bmatrix}
  m_1 \\
  m_2 \\
  m_3
\end{bmatrix} =
\begin{bmatrix}
  1 \\
  0 \\
  0
\end{bmatrix} +
\begin{bmatrix}
  -1 & 1 \\
  1 & -1 \\
  2 & 0
\end{bmatrix} . x
\]

and the only invariant is

\[ m_1 + m_2 = 1 \]

i.e. the Invariant method can tell us nothing about the marking at p3.

Using Keller's method we can prove:

\[ m_1 + m_2 = 1 \]

\[ m_3 \text{ is even} \]

\[ m_3 \geq m_2 \]
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 statements of the form:

(3.1) "Property X holds for every marking \( m \) that is reachable from \( m_0 \)."

There is a third way to prove this statement: explicitly list all markings reachable from \( m_0 \) 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 $m_0$ and incidence matrix $W$. We inductively define a set $R(m_0)$:

1. $m_0$ is in $R(m_0)$
2. if $m$ is in $R(m_0)$ and if $W_i$ is a column of $W$ such that $(m + W_i) \geq 0$, then $(m + W_i)$ is in $R(m_0)$.

It is easy to see (refer to sec 1.2) that $R(m_0)$ is exactly the set of markings reachable from $m_0$. A program for determining $R(m_0)$ is given in Appendix 1.

Example: consider the Petri net of Fig. 1.

(3.2) $m_0 = [n \ 0 \ 0 \ 0 \ 0 \ n]$ and

\[
W = \begin{bmatrix}
-1 & -1 & 0 & 0 & 1 & 1 \\
1 & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & -1 \\
0 & 0 & -1 & -n & 1 & n
\end{bmatrix}
\]

It can be shown (App. 2) that the cardinality of $R(m_0)$ is:

(3.4) $\frac{1}{6}(n + 9n + 14n + 6)$

For $n = 3$, $R(m_0)$ was determined using App. 1 and it was verified (by inspection of each marking in $R(m_0)$) that $m_4 = 0$ or $m_5 = 0$; $m_5 \leq 1$ (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 \( R( m_0 ) \).

### 3.2 Modifying the Invariant Method

Note the following:

(3.5) \( m \) is reachable from \( m_0 \) if and only if there is a sequence

\[
\begin{align*}
\sum_{j=1}^{M} W_{ij} &\geq 0 \quad \text{for} \quad 1 \leq k \leq M \\
\sum_{j=1}^{M} W_{ij} &= m - m_0 \\
\text{and} \quad \sum_{j=1}^{M} W_{ij} &= m
\end{align*}
\]

(see sec. 1.1)

(The firing sequence \( \{ t_{ij} \} \) then transforms the marking from \( m_0 \) to \( m \).) Determining that \( m \) is reachable from \( m_0 \) 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) \( m = m_0 + W x \quad \text{for some vector} \ x \).

The converse, however, is false: the satisfaction of (3.6)
does not imply \( \mathbf{m} \) is reachable from \( \mathbf{m}_0 \).

**Example:** (see Fig. 10)

\[
\mathbf{m} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} \cdot \mathbf{x}
\]

is satisfied by

\[
\mathbf{m} = \begin{bmatrix} N \ (N+1) \ (N+1) \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} N & N & N \end{bmatrix}, \quad \text{for every positive integer} \ N. \ \text{However, the only markings reachable from} \ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \ \text{are} \ \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \ \text{and} \ \begin{bmatrix} 3 & 0 & 0 \end{bmatrix}.
\]

Thus, (3.6) is a necessary, but not sufficient, condition that \( \mathbf{m} \) be reachable from \( \mathbf{m}_0 \). Nevertheless, (3.6) is useful in determining reachability:

**Step 1:** Does there exist \( \mathbf{x} \) such that (3.6) holds? If not, then \( \mathbf{m} \) is not reachable from \( \mathbf{m}_0 \). If so, determine \( \mathbf{X} \), then

**Step 2:** see if (i) and (ii) can be satisfied where

\[
\sum_{j} e_{ij} = \mathbf{x}. \quad (\mathbf{e}_k \ \text{has} \ 1 \ \text{in position} \ k \ \text{and} \ 0 \ \text{elsewhere}.)
\]

**Summary:** Step 1 is used to narrow our search for a firing sequence that will transform \( \mathbf{m}_0 \) into \( \mathbf{m} \). Step 2 then searches through this smaller list of candidates.

**Thm.** There is \( \mathbf{x} \) that satisfies (3.6) if and only if \( \begin{bmatrix} \mathbf{q} \cdot \mathbf{m} \\ \mathbf{q} \cdot \mathbf{m}_0 \end{bmatrix} \) for every \( \mathbf{q} \) such that \( \mathbf{q} \mathbf{W} = \mathbf{0} \)

**Pf:** easy exercise in linear algebra.

**Remark:** The Invariant method stops after step 1. After finding all \( \mathbf{q} \) such that \( \mathbf{q} \mathbf{W} = \mathbf{0} \), the Invariant method then deals
with R', where \( R' = \{ m: q \cdot m = q \cdot m_0 \ \text{for all} \ q \ \text{such that} \ q \cdot W = 0 \} \)

\( R' \) is a superset of \( R(\ m_0 \) \), so there may be properties that hold in \( R(\ m_0 \) \) but not in \( R' \).

The Invariant method cannot prove these properties.

**Example 1:** Suppose \( m_0 \) and \( W \) are given by (3.2) and (3.3).

Then there is \( x \) that satisfies (3.6) if and only if:

\[
\begin{align*}
(3.7) \quad n &= m_1 + m_2 + m_3 + m_4 + m_5 \quad \text{and} \\
&= m_4 + n + m_5 + m_6
\end{align*}
\]

If (3.7) is satisfied, then the most general \( x \) satisfying (3.6) is

\[
(3.8) \quad x = \begin{bmatrix}
m_2 + m_4 + C_5 \\
m_3 + m_5 + C_6 \\
m_4 + C_5 \\
m_5 + C_6 \\
C_5 \\
C_6
\end{bmatrix}
\]

where \( C_5 \) and \( C_6 \) are arbitrary constants. (Again, this is a standard linear algebra calculation.)

To recapitulate: \( m \) is not reachable from \( m_0 \) unless (3.7) is satisfied. If (3.7) is satisfied, then \( m \) is reachable provided we can specialize \( C_5 \) and \( C_6 \) in (3.8) so that there is a "legal firing sequence that sums to \( x \)."

(We may assume \( m \geq 0 \), since otherwise there is no possibility of finding a legal firing sequence that sums to \( x \).)

One choice that works: (\( C_5 = C_6 = 0 \)
\[ x = (m_2 + m_4) e_1 + (m_3 + m_5) e_2 \]

\[ + m_4 e_3 + m_5 e_4 \]

i.e., fire \( t_1 (m_2 + m_4) \) times; then fire \( t_2 (m_3 + m_5) \) times; then fire \( t_3 (m_4) \) times; then fire \( t_4 (m_5) \) times.

We have verified

**Result:** For the Petri net of Fig. 1, \( m \) is reachable from \( m_0 \) if and only if \([ m \geq 0 \text{ and } (3.7) \text{ is satisfied.} \) i.e., we have determined \( R( m_0 ) \) 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 \( t_1' \) and \( t_3' \).

(see Fig. 5). We obtain the following:

**Example 2:**

(3.9) \[ m = [m_1 \ m_2 \ m_3 \ m_R \ m_{R'} \ m_{Rc} \ m_{Rc'}] \]

(3.10) \[ m_0 = [n \ 0 \ 0 \ 1 \ 0 \ (Z-1) \ 0] \]

\[
\begin{array}{ccccccc}
\text{t1} & \text{t1'} & \text{t2} & \text{t3} & \text{t3'} & \text{t4} \\
1 & -1 & 0 & 1 & -1 & 0 & 1 \\
2 & 1 & 0 & -1 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 1 & 0 & -1 \\
\end{array}
\]

(3.11)

\[
\begin{array}{ccccccc}
W = R & -1 & 2 & -1 & -1 & 0 & 1 \\
R' & 2 & -2 & 0 & 0 & 0 & 0 \\
Rc & -1 & 0 & 1 & -(Z-1) & Z & -1 \\
Rc' & 0 & 0 & 0 & Z & -Z & 0 \\
\end{array}
\]
Then there is a vector $x$ satisfying (3.12) $m = m_0 + Wx$ if and only if

(3.13a) $m_1 + m_2 + m_3 = n$
(3.13b) $-m_2 + m_3 + m_R + m_{R'} = 1$
(3.13c) $m_R + m_{R'} + m_{Rc} + m_{Rc'} = Z$

Assuming $m$ satisfies (3.13), then $x$ satisfies (3.12) if and only if

$$x = \begin{bmatrix} m_2 + C_1 \\ m_2 + C_1 - \frac{1}{2}m_{R'} \\ C_1 \\ m_3 + C_2 \\ m_3 + C_2 - \frac{1}{2}Zm_{Rc'} \\ C_2 \end{bmatrix}$$

where $C_1$ and $C_2$ are arbitrary constants.

Is there a choice of $C_1$ and $C_2$ such that the resulting $x$ is the "sum of a legal firing sequence"? ... sometimes.

**Lemma 1:** If $m$ is reachable from $m_0$ then $m_{Rc'}$ is a multiple of $Z$ and $m_{R'}$ is a multiple of 2.

**Pf.:** If $m$ is reachable from $m_0$, there is a choice of $C_1$ and $C_2$ such that the entries of (3.14) are non-negative integers.

**Lemma 2:** If $m$ is reachable from $m_0$, then $m_{Rc'} = 0$ or $Z$.

**Pf.:** Use lemma 1 and (3.13c)

**Lemma 3:** If $m$ is reachable from $m_0$, then

$\#t_1 \geq \#t_1'$

35
\[\#t_1 \geq \#t_2\]

\[\#t_3 = \#t_3' \quad \text{or} \quad \#t_3 = \#t_3' + 1\]

\[\#t_3 \geq \#t_4\]

(where \(\#t_i\) = number of firings of transition \(t_i\) in the firing sequence used to reach \(m\) from \(m_0\))

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

\[x = [\#t_1 \#t_1' \#t_2 \#t_3 \#t_3' \#t_4]\]

**Lemma 4:** If \(m\) is reachable from \(m_0\), then \(m_2 = 0\) or \(m_3 = \).

**Pf.:** Suppose not. Then there is an intermediate marking \(M\) such that

(i) \(M_2 = 0\), \(t_3\) is enabled

(ii) \(M_3 = 0\), \(t_1\) is enabled

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

**Suppose (i):** by (3.11), \(M_2 = \#t_1 - \#t_2 > 0\) and \(M_3 = \#t_3 - \#t_4 = 0\).

Then \(m_{Rc} = (Z - 1) - \#t_1 + \#t_2\)

\[- (Z - 1)\#t_3 + Z\#t_3' - \#t_4\]

Recall that \(m_{Rc}\) is initially \((Z - 1)\))

Thus, \(m_{Rc} = (Z - 1) - (\#t_1 - \#t_2)\)

\[- Z(\#t_3 - \#t_3') + \#t_3 - \#t_4 < (Z - 1)\]

\((\#t_1 - \#t_2 = M_2 > 0; \#t_3 = \#t_3' \quad \text{by} \quad \text{lemma 3;}\)

\(\#t_3 - \#t_4 = M_3 = 0.\)

But \(m_{Rc} < (Z - 1)\) implies \(t_3\) is not enabled. (See (3.11)).
This is a contradiction.

The case (ii) is an exercise.

Lemma 5: If \( m \) is reachable from \( m_0 \) then either \([m_2 = mR' = 0]\) or \([m_3 = mRc' = 0]\).

Pf.: (i) Suppose \( m_3 = 0 \). Then (3.13b) implies \( mR + mR' = m_2 + 1 > 0 \).

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

(ii) Suppose \( m_3 > 0 \). Then lemma 4 implies \( m_2 = 0 \). \((3.13b)\) now implies \( mR + mR' = 0 \). i.e., \( mR' = 0 \).

Lemma 6: If \( m \) is reachable from \( m_0 \) then \( m_3 = 0 \) or 1.

Pf.: \( m_3 > 0 \) implies \( m_2 = 0 \). (lemma 4) \((3.13b)\) now implies \( m_3 = 1 \).

Thm. \( R(m0) \) consists precisely of:

\[
\begin{bmatrix}
  n & 0 & 0 & 1 & 0 & (Z-1) & 0 \\
  (n-1) & 0 & 1 & 0 & 0 & Z & 0 \\
  (n-1) & 0 & 1 & 0 & 0 & 0 & Z \\
  (n-x) & x & 0 & (x+1-2y) & 2y & (Z-1-x) & 0
\end{bmatrix}
\]

where \( 0 \leq x \leq n \) and \( 0 \leq 2y \leq 1 + x \).

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

(a) \([m_2 = mR' = 0]\) or

(b) \([m_3 = mRc' = 0]\)

Suppose (a) holds. Then

\((3.13a)\) yields: \( m1 + m3 = n \)
(3.13b) yields: \( m_3 + m_R = 1 \).

(3.13c) yields: \( m_R + m_{Rc} + m_{Rc'} = Z \)

**Case 1a:** \( m_3 = 0 \) implies \( m_R = 1 \). This now implies (lemma 2)

\[ m_{Rc'} = 0. \]

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

**Case 2a:** \( m_3 = 1 \) implies \( m_R = 0 \). This implies \( m_{Rc} + m_{Rc'} = Z \).

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 \( m_0 \).

Suppose (b) holds. Then

(3.13) yields: \( m_1 + m_2 = n \)

\[ m_R + m_{R'} = m_2 + 1 \]
\[ m_R + m_{R'} + m_{Rc} = Z \]

By lemma 1, \( m_{R'} = 2y \). Denoting \( m_2 \) as \( x \), the only candidates for reachability in case (b) are:

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

where \( 0 \leq x \leq n \) and \( 0 \leq 2y \leq x+1 \)

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

\((t_1, t_1')\) fired \( x \) times; then

\((t_1, t_2)\) fired \( y \) times

**Remark:** Finding \( R( m_0 ) \) 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 \( m_0 \), and a "desirable" set of markings \( MD \), where each marking in \( MD \) is reachable from \( m_0 \) 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 \( t_m \) is enabled. This definition is clearly inadequate.)

One way of showing deadlock is impossible: show that if \( m \) is reachable from \( m_0 \), then \( m_0 \) is reachable from \( m \).

(In [2], \( m_0 \) would be called a "home state.") This implies that if \( m_1 \) and \( m_2 \) are both reachable from \( m_0 \), then \( m_2 \) is reachable from \( m_1 \).

For example 1: If \( m \) is reachable from \( m_0 \), then \( m_0 \) is reachable from \( m \). (See (3.2), (3.3) and (3.7) and result 1).

Proof:

We must show that if

\[
\begin{align*}
(3.15) \quad m & \geq 0 \\
 & = m_1 + m_2 + m_3 + m_4 + m_5 \\
 & = m_4 + n \cdot m_5 + m_6
\end{align*}
\]

then \( m_0 = \begin{bmatrix} n & 0 & 0 & 0 & 0 & n \end{bmatrix} \) is reachable from \( m \).

(3.15) implies there is vector \( x \) such that

\[
(3.16) \quad m_0 = m + W x
\]

where \( W \) is given by (3.3).

In fact, the most general \( x \) satisfying (3.16) is
where \( C_1 \) and \( C_2 \) are arbitrary constants.

Can we specialize \( C_5 \) and \( C_6 \) so that the resulting \( x \) is the "sum of a legal firing sequence" beginning with the initial state \( m_0 \)?

The "smallest" choice is \( C_1 = m_2 + m_4; \) \( C_2 = m_3 + m_5 \)

Then

\[
\begin{bmatrix}
0 \\
0 \\
m_2 \\
m_3 \\
m_2 + m_4 \\
m_3 + m_5
\end{bmatrix}
\]

The following firing sequence is legal and sums to \( x \)

- Fire \((t_6)\) \( m_5 \) times;
- fire \((t_5)\) \( m_4 \) times;
- fire \((t_3)\) \( m_2 \) times;
- fire \((t_5)\) \( m_2 \) times;

(At this time we have reached

\[
[(n-m_3) \ 0 \ m_3 \ 0 \ 0 \ n]\begin{bmatrix} m_1 & m_2 & m_3 & m_4 & m_5 & m_6 \end{bmatrix})
\]

- fire \((t_4, t_6)\) \( m_3 \) times.
For example 2: If \( m \) is reachable from \( m_0 \), then \( m_0 \) is reachable from \( m \).

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

Proof: that \( m_0 = [n \ 0 \ 0 \ 1 \ 0 \ (Z-1) \ 0] \) can be reached from the first three markings listed in the theorem is an exercise.

To show \( m_0 \) 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: $m_0(LP) = n = m_0(S)$
$m_0\text{ (other)} = 0$

Fig. 1
For convenience, several places have been drawn more than once.

Initial marking: $m(T_i) = 1; m(F_i) = 1; m(E_i) = 0; \ 1 \leq i \leq 5$

Fig. 2
Fig. 3
Initial marking: $m_1 = 1; m_2 = 0 = m_3$

Fig. 4a
Initial marking: \( m_1 = 1; \) \( m_2 = 0 = m_3 \)

Fig. 4b
Initial marking: $m_1 = 1; m_2 = m_3 = m_3' = 0$

Fig. 4c
Initially: \( m_1 = n; m_2 = 0 = m_3; R = 1 \)

Fig. 6.
\[ t_1 \quad V := V - c \]
\[ t_2 \quad V := V + c \]
\[ t_3 \quad \text{When } V \geq c \quad V := V + d \]

Fig. 7a
\[t_4 \text{ When } V < c \quad V := V + d\]

\[t_5 \text{ When } V = c \quad V := V - d\]

Fig. 7b
For convenience, place R has been drawn twice.

Initial marking: \( m_1 = n \); \( m_2 = 0 = m_3 \); \( m_R = 1 \); \( m_{Rc} = Z - 1 \geq n \)

Fig. 8

52
Initial marking: $m_1 = 1; m_2 = 0 = m_3$

Fig. 9
Initial marking: $m_1 = 0; m_2 = 1 = m_3$

Fig. 10
Bibliography


Appendix 1

The following program will generate all markings in Reach ($m_0$). The program terminates with list = Reach ($m_0$) if and only if Reach ($m_0$) is a finite set.

BEGIN

Initialize a queue to empty;

Initialize a list to empty;

Enqueue ($m_0$);

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 ($m + W_i >= 0$) AND ($m + W_i$) is neither in the list nor in the queue

THEN Enqueue ($m + W_i$)

END

END.
Appendix 2

Def. 1: We define $F(n) =$ number of distinct vectors $\mathbf{m} = [m_1 \ m_2 \ m_3 \ m_4 \ m_5 \ m_6]$ that satisfy:

(i) $m_i$ is a non-negative integer for each $i$

(ii) $n = m_1 + m_2 + m_3 + m_4 + m_5$

(iii) $n = m_4 + n \cdot m_5 + m_6$

($n > 0$)

Def. 2: We define $G(n) =$ number of distinct vectors $\mathbf{m} = [m_1 \ m_2 \ m_3]$ that satisfy:

(i) $m_i$ is a non-negative integer for each $i$

(ii) $n = m_1 + m_2 + m_3$

($n \geq 0$)

Lemma 1: For each $n \geq 0$, $G(n) = (n + 1) \cdot (n + 2)/2$

Proof: Assign to $m_3$ the integer value $j$, where $0 \leq j \leq n$. Then $(n - j)$ units remain for assignment to $m_1$ and $m_2$. This latter assignment can be done in $(n - j + 1)$ ways.

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 $m_5$ may have only the values 0 or 1. If we assign $m_5$ the value 1, then we must assign $m_4$ and $m_6$ the value 0. We must then assign $m_1$, $m_2$, $m_3$ values such that $n = m_1 + m_2 + m_3 + 1$.

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

If we assign $m_5$ the value 0, then $m_6 = n - m_4$ and $n = m_1 + m_2 + m_3 + m_4$.

Assign $m_4$ the value $k$, where $p < k < n$. Now we must assign $m_1$, $m_2$, $m_3$ values such that $n = m_1 + m_2 + m_3 + k$. This latter assignment can be done in $G(n - k)$ many ways.

**Thm.**: For each $n \geq 0$, $F(n) = \frac{n^3 + 9n^2 + 14n + 6}{6}$

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

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.