# SYMMETRIES OF THE FIRING SQUAD SYNCHRONIZATION 

## problem revealed IN a nested array

J. Philip Benkard<br>Staff Programmer, Quality Assurance International Business Machines Corporation<br>Dept 17W / Bldg 630 Route 52<br>Hopewell Junction, NY 12533-0999

## ABSTRACT

The firing squad problem of automata theory is described and a simple, nonoptimal solution is given. Examination of output traces shows that the solution has at least one obvious symmetry. If the transition matrix is reshaped and partitioned, the resulting nested array not only reveals and explicates more symmetries but suggests ways in which nested arrays can be used to save space.

## INTRODUCTION

A finite state automaton is a device with a finite number each of states, inputs, and outputs. One of the inputs is the current state; one of the outputs is the next state. The other inputs and outputs are external. As time advances in discrete steps, the state of the device at one time is determined entirely by the inputs at the last previous time. The firing squad problem requires the definition of such an automaton, called a soldier, with two external inputs which are the states of his left and right neighbors when arrayed in a firing squad line. The only output is the next state.

From time to time a squad is formed whose size is determined by the prominence of the guest of honor. In view of the inevitable inflation of political values, the possible size of a squad must be considered unlimited. The automata are set to a common initial state, and an officer at one end of common initial state, and an officer at one end of The soldiers, following the rules of finite state automata and each receiving input only from his two neighbors, exchange enough information to fire simultaneously as soon as possible. Firing is represented by a particular state of the automaton. The optimal solution, which takes $2+2 \times \mathrm{N}$ steps, where $N$ is the length of the squad, is quite complex. A simple solution with only twelve internal states will be shown. For squads of length internal states will be shown. For squads_of len
two or greater it requires no more than
$2+3 \times N$ two or
states.

The problem is nontrivial because of the combination of synchronization and finiteness of the individual automata. It is not hard to design soldiers who will fire in sequence down the line. It is not too much harder to synchronize if the automata can "count off"; just like real soldiers When the soldier at the far end finds out that he When the soldier at the far end finds out that he
is number 576273054 , the soldiers can easily count down in sequence and synchronize. But a finite state automaton has a maximum limit to his

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specfic permission.
countability; a squad beyond that limit cannot so synchronize itself.

A much simplified problem will be analyzed first to introduce the method of message transmission. to introduce the method of message transmis symmetries. A solution to the original problem is then presented and analyzed. When the shallow transition matrix, on which the solution is based, is transformed into a depth three array, the symmetries of the internal states of the automata and of the connections between automata are revealed as different forms of reflections.

The shallow matrix, furthermore, is rather sparse. The distribution of significant elements in items and subitems can be used to effect siqnificant storage savings.

## A TRIVIAL CASE

Let the requirement be that only the soldier at the opposite end of the line from the officer must fire his weapon. The command to synch and shoot can be thought of as a shoulder tap of the first soldier. The internal states must be defined so that the shoulder tap is transmitted to the next soldier during each time interval. For this case, three internal state suffice. Before the command is given all soldiers are in a common idling state, which is selfperpetuating; if a soldier and his neighbors are in the idling state, the soldier remains idle for the next time period.

The next state is determined by three values, each of which has a range of three. A 3 by 3 by 3 transition matrix is necessary. The values in the matrix are the same three values with one addition, the firing state. The iterative solution halts as soon as one or more soldiers enter that state so that this state is never an input to another soldier.

The states are represented by characters, which can be used to create a suggestive display. This the possibility of a few 20 by 20 by 20 four byte integer matrices was fraught with hints of W5FULL. The representation turned out to be serendipitous. The characters are in a variable called SFSCH, standing for Simple Firing Squad CHaracters.

SFSCH $\leftrightarrow$ '. = ${ }^{\circ}{ }^{\prime}$
The logarithn symbol represents the firing state; the period, the idling state; the equals, the shoulder tap state. The quad remains in the same state regardless of what happens to his neighbors; for that reason, he is sometimes called a blockhead. Before the iterative solution begins, a blockhead is affixed to each end of the line to provide neighbors for the end soldiers. Blockheads will be found to play a role at other than end points in the general solution.

The transition matrix is called SFSM, for Simple Firing Squad Matrix, It is conveniently displayed by lowering its last two axes

$$
\begin{array}{ll} 
& \text { C[1 } 2 \text { 2] SFSM } \\
\cdots & -- \\
\hdashline-\infty & -\square \\
\cdots & -0
\end{array}
$$

The three axes represent，respectively，a soldier， his left neighbor，and his right neighbor．Thus， the left hand 3 by 3 matrix represents a soldier in state zero（．）．Rows stand for values of the left neighbor；columns，the right．For example，the middle rows of the three displayed matrices represent a soldier whose left neighbor is in state one（ $=$ ）；the left columns，a soldier whose right neighbor is in state zero（．）．

The character in each position gives the state of the soldier at the next point in time．One（ $=$ ） will be the state of the soldier at the next time if the soldier and his right neighbor are in state zero（．），and his left neighbor is in state one（ $=$ ）． Thus an equals sign appears in plane zero，row one， and column zero．The minus sign is used at locations which will never be used；that is，for configurations of states which will never occur．

The character representations can be converted to small integers by finding their indices in SFSCH．The numbers for a particular time can then be used as indices into the transition matrix to find the next internal state．For example，at time zero a squad of length three is displayed

```
\[
\square=\text {. . }
\]
```

Soldier zero is in state one（＝）；his left neighbor is in state two（ $D$ ）；his right，in state zero（．）． For soldier one，the integers are zero（．），one $(=;$ ， and zero（．），respectively．Using these values to select from the transition matrix

$$
\left.\left(e^{* *}(1) 20\right)(010)\right) 0^{* *} \in S F S M
$$

Soldiers zero and one are in idle and shoulder tap states，respectively，one time interval later．The firing squad at this time would be represented

ㅁ．＝．$\square$
The shoulder tap has been passed on from soldier zero to soldier one．

A simple expression creates a few examples
SFSGO＂1＋26
믐 ㅁ．

$01{ }^{\nabla}$
［1］Z世SFSGO Y
$\nabla$ 1987－08－26 13．39．07（GMT－4）
SFSGO，standing for Simple FSGO uses the variables defined above as left argument to the solution to the general problem，FSGO．The code for FSGO appears in the appendix．The right argument is the length of the squad．The reader may wish to confirm that the log symbol at SFSM［1；2；2］is only used for the one man squad．The starting position for a for the one man squad． $\begin{aligned} & \text { The starting position for a } \\ & \text { one man squad is the only instance of an equals sign }\end{aligned}$ between two quads．

The iterated line，FSGO［6］，is the core of the solution．Before iteration starts，the transition matrix is converted to integer form using INDEX $0 F$ so that the look up is not required at each step， and the variable $L$ is set to one less than the and the variable $\begin{aligned} & \text { is set to one less than the } \\ & \text { number of characters，that is，to the index of the }\end{aligned}$ number of characters，that is，to the indek of the
firing state． is ised in ine［6］to detect an firing state
ofcurrence of the firing state（ $\Theta$ ）so that iteration will stop．

The indices for the state of each individual automaton are determined in FSGO［6］by using the squad together with its left and right shifts．The two block heads appear together in both shifted forms．Thus，the blockhead entries in the tramsition matrix，which correspond to a soldier and transition matrix，which correspond to a soldier
one of his neighbors in state two，preserve that one of his neighbors in state
state over each time interval．

THE GENERAL CASE

## If FSGO is called monadically it uses the global variables FSM and FSCH．It also uses the

 sign of its right argument to determine at which end of the line the officer stands．The conceit is simple．Two messages are sent，one traveling at one simple．Two messages are sent，one traveling at one third the speed of the other．The delay in the before each soldier taps the shoulder of an adjacent soldier．When the slow message meets the reflection of the fast，the midpoint of the squad has been found．One or two soldiers，depending on parity， become blockheads，defining two squads of the same length，and fast and slow messages are dispatched from the midpoint in opposite directions．Ablockhead in the general solution can leave his inert state only when one of a set of prefiring conditions is met；at the next moment he fires．
Since a squad of any length can be called as a subdivision of at least one larger squad and will in that case be called from both left and right ends， the solution must be symmetric．

The symmetry of the overall solution and the use of blockheads to define subdivisions are illustrated in the following small example．


The function REP＿IN replaces the dots with blanks， which removes clutter from the figure．There is obvious symmetry in these two figures；one is just the reflection of the other

$$
0 \quad X \equiv \Phi Y
$$

Oops！It is also necessary to reverse the individual characters using

```
        \nablaREVCH[D]D
    [0]
        \nabla
        Z&REVCH Y A Where defined, REVCH"*}\leftrightarrow<REVC
        (,Z)*(' -',FSCH)[(' -',F5-CH)-,Z Z Y ]
        0 1987-08-24 15.08.57 (GMT-4) and
        PFSCH FS-
```



```
        <<<=.\square|->>つき*
```

The high minus sign in the name $F S^{-} C H$ suggests a reversal．The directed symbols are interchanged in the two variables；the symmetric symbols are unchanged．

## $X \equiv$ REVCH $\boldsymbol{X} \mathbf{Y}$

That＇s better！The rich APL symbol set led to characters，originally chosen to save space in an APLSV workspace，which themselves express a symmetry．Figure 1，which follows the text，gives the complete run of a seventeen man squad and the end of a seventy－five man squad．Parts of the smaller example occur in the larger．The seventy－five man squad divides five times．

## DISCOVERING INTERNAL SYMMETRY

Once the notion of the reflected messages and the key sixth line of FSGO were in hand, there remained only the entry of items in the transition matrix. This was done experimentally by changing locations, lines, and planes in the matrix. The first solution looked something like Figure 2. Not much symmetry there.

The reader might have noticed that the array in Figure 2 has length fifteen in each axis rather than length twelve. The first step in improving the solution was to use a trace matrix to keep track of locations in the transition matrix which were actually used during a set of runs of FSGO. The trace matrix was a bit array, originally all zeroes, with the same shape as the transition matrix. A line of code was added to the then current version of FSGO to enter ones in the trace matrixat each time step. It was quickly clear that few of the locations were in fact used, for example, no opposing left and sight arrows are ever adjacent. With this information states were combined, reducing the size to 12 by 12 by 12 . From fifteen to twelve may not seem like much but

### 0.512

$$
1215 \text { +.* } 3
$$

so that for a rank three array, nearly half the space is saved.

The twelve remaining states were of three distinct types, representing respectively, right moving, stationary, and left moving states. The set of characters was reordered to form the variable FSCH given above. Figure 3 shows the transition matrix after each of its axes had been correspondingly reordered.

Little symmetry shows here, but symmetry requires the right point of view. In the form shown in Figure 3, the three axes represent left neighbor, soldier, right neighbor. This means that the rotation control vector in FSGO[6] was 101. That is a reasonable way to arrange a set of consecutive integers, but it is from the point of view of the guy in the middle that symmetry is most readily apparent.

Figure 4 shows the matrix with the first two axes interchanged. This is the form of the matrix used by FSGO. There are some more patterns, but symmetry is still elusive. To appreciate the effect of the axis transposition, the reader is challenged to find in Figure 3 the sort of cut-off rectangle of mostly $\square$ in item [1;2] of Figure 4. A good place to start is the bottom row of Figure 3 .

Now notice that in corresponding items in the top and bottom rows the same characters appear reversed. That is reminiscent of the need to reverse characters to obtain an overall symmetry of the result. Figure 5 is the result of partitioning each item of Figure 4 corresponding to the groups of characters representing left moving, stationary, and right moving states. This matrix, FSMD3 is of depth three and satisfies the following identity

-BT[ם]
$\nabla$


- 1987-08-06 14.03.03 (GMT-4)

In the identity and in line zero of the function the underscore is used between two characters to
designate a single overstruck symbol which is not in the character set. This convention is adopted from an APL2 input convention. Using the symbols $\phi$ and o_for REVCH and BT we can write

This identity expresses the symmetry of the solution transition matrix which underlies the symmetry of results. We shall return to discuss this symmetry later on.

## SOME OBSERVATIONS ON STORAGE

The transition matrix has 1728 leaves of which only 102 are not -'. Sparse array techniques could be used to store this matrix, although for one so small not much would be gained. The arrangement of nontrivial leaves in the matrix is, however, not random. The partitioned matrix FSMD3 has twelve items each of which has nine subitems. Only thirty-two of the one-hundred-eight subitems contain any useful information; those with sixteen minus signs could be replaced by just one. Furthermore, only eight of the thirty-two useful subitens contain instances of more than one character. The other twenty-four could be replaced by the respective character scalars.

The resulting array, FSMRD, is shown in Figure 6. The name stands for Firing Squad Matrix of Ragged Depthe The figure shows the matrix both naked and with each item DISPLAYed using the function provided with APL2. The bare display is catenated to the form with boxes so that rows of items in the two forms are collinear. Since REP_IN has replaced minus signs with blanks, the minus signs which appear in the figure are only those produced by the appear 1 I each contain eight significant characters. Seven of these are boxed; in the case of the eighth an entire item has been replaced with a scalar, so it is DISPLAYed with only a single accompanying minus sign. The middle row contains the other sixteen subitems, with one, eight, four, and three in its respective items.

The system function $\square A T$ can be used to document the space saving of the nested array. With left argument 4 and a variable name as right argument the function returns a two iten vector. The first item is the total number of bytes used to store the array; the second item is the number of bytes used to store actual data. The difference of the two numbers is the nesting overhead.

Here is a short terminal session.
$Y \leftarrow 1 E 64 \times X * 100 \times F S M R D I * F S C H 2 D(10) F S M R D$
NAMES
FSM
FSMRDI $X$


$$
0.1230 .041
$$

The 2 by 3 array in the middle shows the space requirements for all six arrays. The second line in each item, generated by the minus reductions, gives the structure overhead explicitly.

The variable FSMRDI has the same structure as FSMRD, but for each character has been substituted its index in FSCH. INDEX OF is applied with left argument FSCH to each of the scalars in FSMRD argument experimental depth operator, $\underline{D}$, which is shown in the appendix. Since all these indices are small positive integers, the array uses the same amount of space as FSMRD. $X$ and $Y$ are defined to show the effect of changing data clement size. Uniform nesting of the data results in use of about twice as much space as the shallow form uses. The ragged depth form of the array uses about half as ragged depth form of the array uses about half
much as the original flat array. When larger much as the original fategers are used data storage space increases, but overhead is unchanged.

Again referring to the terminal session above, note that of the 212 leaves of the array only 70 are in fact useful, since the remaining i42 are all minus signs, representing locations that are never referred to. 5till, only one-eighth of the original 1728 leaves remain. The inutile items are displayed as blanks, which are underscored by the minus signs generated by the DISPLAY function. They are distributed $32,78,32$, respectively, in They are distributed $32,78,32$, respective he three rows of the ragged array perhaps
structural techniques are competitive with sparse array techniques for some larger problems.

The symmetry identity

still holds for the ragged array because the ENCLOSE of a shallow scalar is the scalar itself.

## ANALYZING SYMMETRY

The twelve items in the partitioned array of Figure 5 each correspond to a state of the center soldier according to the pattern

$$
\begin{aligned}
& \rightarrow \gg \geq \\
& =\cdots \quad 0 \\
& \leftarrow \ll \quad<
\end{aligned}
$$

This fact is suggested by the upper diagran to the right in Figure 5. The leaves of the array correspond to the next state of the automaton. Thus the two left most operations, $\theta$ and REVGH?... interchange respectively the input and output states of the center automaton, For example, if an automaton is in '>', between two ',' states his next state will be ' $s^{\prime}$ '. This is represented by the $' s$ in the 0 ( 0 (1 1)(1) position of the matrix. The three item nested vector could be used as left argument to the PICK function. Each item in the vector selects an item in a level of the transition matrix, beginning at the top. Thus 01 selects item one in row zero, 1 selects subitem one in row one of that item, and the second 11 selects the ' $r^{\prime}$ '. The other, ' ${ }^{\circ}$ ' in item 0,1 ,
corresponding to $a$ '>' between $a$ ' and a ' $\rightarrow$ ' is selected by ( 0 1)(1 0)(2 0). It occurs in the illustrations of FSGO 8 shown earlier. The two operations thus assure that the symmetric cases are handled correctly as far as the center automaton of a triple is concerned.

The remaining two operations take care of the center automaton's view of his neighbors. Let R, 5 , and $L$ stand for right moving, stationary, and left moving states. Then each subitem corresponds to a combination of two of these types of states

| \wl | R | 5 | 1 |
| :---: | :---: | :---: | :---: |
| - |  |  |  |
| 51 | SR | 55 | , |
|  | SR | SS | SL |
| L 1 | LR | LS | LL |

where the rows correspond to the left neighbor; the columns, to the right. In the lower diagram to the right in Figure 5 the breakdown of neighbor states into the three categories is indicated by spaces in the row and column of symbols. If an automaton is in, say, the LS configuration, that is, he has a left moving state on his left and a stationary state on his right, then in the symmetric state he must have a stationary state on his left and a right moving state on his right. To the automaton the important thing is whether the neighbor is bearing an incoming message or an outgoing message. As the diagram just above shows, the operation which makes that switch is BT", the backwards transpose on items. Finally, the $Q^{\cdots}$, transpose on subitems interchanges the left and right automatons. The symmetries of the transition matrix support the symmetry of results mentioned earlier.

## USING THE NESTED TRANSITION MATRIX

The function FSGON, for Firing Squad Go Nested, uses the nested transition matrix. It is similar to FSGO but requires two functions to handle the indexing paths which PICK uses to access array items below the first level. All three functions are shown in the appendix. FSMITR translates the indices into FSM, which are derived directly from the firing squad line, into paths of length three which would work to retrieve items from FSMD3. Since all the paths are of length three, an index error would occur with FSMRD for leaves which are not three levels into the array, that is exactly those scalars whose substitution for larger subarrays led to space savings. The function SCPIC is a variant of PICK which ignores all items in the left argument after a shallow scalar is reached.

CONCLUSIONS
In the above simple analysis of a classic problem, use of a nested array to describe the solution provided two quite different kinds of insight:

- The symmetries of the transition array, which underlie the symmetry of the solution, are easily expressed as reflections at differeat levels of the array. The group of reflections includes one across oach of the axes and diagonals. The symmetry expression holds for the ragged depth version of the matrix without change.
- The transition matrix is a small array, yet use of a nested structure results in saving about one half the storage space. This is achieved with no basic change in the progran which uses it.

APPENDIX

DISP_FS
Miscellaneous functions and data
Z-IM FSGO FSI; GHR; $L$ S Simulate the firing squad A INPUIS: LEFT - Transition matrix / Character list A RIGHT - Squad length; start at right if < 0 $\Phi(0=\square N C$ 'IM')/'IM\&FSM FSCH' A Globals for standard problem.



$Z \leftrightarrow R E V C H$ Y A Where defined, REVCH** $\leftrightarrow$ REVCH

$Z \leftrightarrow B T Y \quad A \leftrightarrow O_{0}, \quad$ Transpose on minor axis
$\mathbf{Z} \leftrightarrow \Phi \oplus \boldsymbol{Y}$


```
FSMD3 =_ ө REVCH\cdots\cdots...... BT* FSMD3
    FSM =_ = , r*,[0]r*, FFSMD3
```



```
    FSMRD =__ e REVCH********* FSMRD
    FSMD3T FSMRDI =_ (CFSCH) थ@(1 0)" FSMD3 FSMRD
```



```
    F5M=__1 0 2&FSMOLM;M;ME-1\downarrowFSCHO_FSCHJ
    OFSCH FSCHO a As originally conceived
->>>2=.ロ|<<e\leq@
->>>2*<<S=10
```

0 SHLST FSFNLST
$\nabla$ 0．$Z \in X$ SFSGO $Y$ A Solution for small squad．
$\qquad$ Z↔SFSM SFSCH FSGO Y

```
\nabla 0. ZEL(F D G)R;MON;T A **_* (DEPTH)
```



```
4. Z\leftarrow,\Phi(4\timesMON&0=\squareNC''L') ' '(= (= L)(=__R)' A
    *)
    ->LD [J RD RJ IFE<C2 The,f-f invocation is as suggested in [Iverson '83, p 44]
```




```
    LD:F&1\downarrow4^0,F A In Case empty F is indicated by ''.
A 10. A P is the power operator, which cannot take as operand.
```



```
f 12. A Alternative using the promoting transform of [Benkard APL84]
```



```
13. A.$
\div15. RJ:ロES 5 4 A
```



```
->17. דSYMM IF 0=ppG A A A ASYmm is the LEVEL operator of [Thomson APL84]
    18. DES 5 3 IF MON^2=OG A ....O Length error.
```





```
A 22. A Alternative using the MESH invocation of B (\).
```



```
->%24. SYMM:->NNEG IF 0\leqG Neg does (IG+1)"S, but stops at
```



$\nabla$ 0. ZEIM FSGON FSI;CHR;L $A$ Simulate the firing squad with nesting.
A 1. A INPUIS: LEFT - Hested transition matrix / Character list
A 2. A RIGHT - Squad length; start at right if < 0



$\nabla$ 0. Z\&FSMITR Y Shallow indices (FSM) $\Rightarrow=>$ depth three paths (FSMD3)
$\begin{array}{r}\circ \\ \hline\end{array}$
$Z \in(c 3$ 4TY[0]), $C[1] 3$ 4T1廿Y
$\nabla$ 0. Z $\leftarrow X$ SCPIC $Y$


$\begin{array}{ll}\rightarrow 3 . & \rightarrow 20==-Z<X[0]=Y \\ \rightarrow 4 . & \rightarrow 2==-Z \leftarrow X[1]\end{array}$

| $\rightarrow$ 4. | $\rightarrow 20==$ |
| :--- | :--- |
|  |  |



The last eighteen lines of the larger squad comprise oight replications
of each half of the smaller squad.
fiGIRE.

| 5 UHL FSIM -0 vee everepeete -. ... -ee •ciesu <br>  ค ¿. 0 . ir : : : :r: nopede900.:- <br>  -000000004.nern <br>  Onntinncominion: |
| :---: |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |




 cerpreperpertiv aicep+eнreperiv Honcoccerceng


" -6+4*0000 inorimbriminhi! 0xcoco000000 0

 $+4+4+40+40$



 menvorventren
 -perepppreceet -rererrorerer-b -recrerevererert ceeperer+erter -cerereppepert -ererererepert -pcequrpertert -



FIGURE 2.


An intermediate version of the transition

Scalars in each item.










ion matrix.
FIGURE 5.

F5MD3 FSMD3
-

: $\overline{1}$


