SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Further Results on Bilinear Behavior Formulation of Finite State Machines

Jumei YUE[1], Yongyi YAN[2*], Zengqiang CHEN[3] & He DENG[2]

[1]*College of Agricultural Engineering, Henan University of Science and Technology, Luoyang* 471000*, China;*
[2]*College of Information Engineering, Henan University of Science and Technology, Luoyang* 471000*, China;*
[3]*College of Artificial Intelligence, Nankai University, Tianjin* 300071*, China*

Appendix A Propositions used in the body of the letter

A.1: Definition 1 (STP of matrices) [1]: For two matrices of arbitrary dimensions $M \in \mathbf{M}_{m\times n}$ and $N \in \mathbf{M}_{p\times q}$, their STP, denoted by $M \ltimes N$, is defined as follows.

$$M \ltimes N := (M \otimes I_{s/n})(N \otimes I_{s/p}), \tag{A1}$$

where s is the least common multiple of n and p, $\otimes$ is the Kronecker product of matrices.

Remark 1. The STP of matrices is a generalization of the tradition multiplication of matrices. In fact the STP of matrices reduces to the traditional one when $n = p$. Therefore, the matrix multiplication defaults to the STP of matrices in this letter.

A.2: Proposition 1 (Pseudo commutative law of vectors) [1]:For column vectors $X \in \mathrm{R}^m$ and $Y \in \mathrm{R}^n$, the following pseudo commutative law holds.

$$\begin{aligned} W_{[m,n]} \ltimes X \ltimes Y = Y \ltimes X, \\ W_{[n,m]} \ltimes Y \ltimes X = X \ltimes Y, \end{aligned} \tag{A2}$$

where $W_{[m,n]}$ is an $mn \times mn$ matrix defined as follows.

$$w_{((I,J),(i,j))} := \delta_{i,j}^{I,J} = \begin{cases} 1, & I = i \text{ and } J = j, \\ 0, & \text{Otherwise.} \end{cases}$$

A.3: Definition 2 (Finite state machine) [2]: A finite state machine (FSM) is a five-tuple $M = (X, \Sigma, f, x_0, X_F)$, where $X = \{x_1, x_2, \cdots, x_n\}$ and $\Sigma = \{a_1, a_2, \cdots, a_m\}$ are state set and alphabet (or say input set), respectively; f is a state transition function from $X \times \Sigma$ to 2^X (2^X is the power set of X); $x_0 \in X$ is an initial state of M; and $X_F \subset X$ is a set of accepting states. An accepting state $x_i \in X_F$ means that if an input sequence $u(t) = a_{i_1}a_{i_2}\cdots a_{i_t} \in \Sigma^*$ moves M from initial state x_0 to state x_i , then the sequence $u(t)$ is accepted by M, where Σ^* denotes the set of all the finite strings on Σ.

An FSM M is said to be a deterministic one if $|f(x, a)| \leqslant 1$ for each $x \in X$ and each $a \in \Sigma$. Otherwise, M is called a nondeterministic one. The string $x_{i_1} \overset{e_{i_1}}{\to} x_{i_2} \overset{e_{i_2}}{\to} x_{i_3} \overset{e_{i_3}}{\to} \cdots \overset{e_{i_k}}{\to} x_{i_k}$ is called a path from state x_{i_1} to state x_{i_k}.

A.4: Proposition 2 (T-step transform dynamics): Let $M = (X, \Sigma, f, x_0, X_F)$ be an FSM, where $X = \{x_1, x_2, \cdots, x_n\}$, $\Sigma = \{a_1, a_2, \cdots, a_m\}$, f is the state transition function. The t-step evolution that M reads input sequence $u(t) = a_{i_1}a_{i_2}\cdots a_{i_t} \in \Sigma^*$ at state x_i can be formulated as follows.

$$\begin{aligned} & f(x_i, u(t)) \\ & = f(x_i, a_{i_1}a_{i_2}\cdots a_{i_t}) \\ & = f(f(x_i, a_{i_1}), a_{i_2}\cdots a_{i_t}) \\ & = f(f(f(x_i, a_{i_1}), a_{i_2}), a_{i_3}\cdots a_{i_t}) \\ & \vdots \\ & = \underbrace{f(f\cdots f(f(f}_{t}(x_i, a_{i_1}), a_{i_2}), \cdots, a_{i_{t-1}}), a_{i_t}). \end{aligned}$$

A.5: Proposition 3 (Block extraction of a matrix): Let A be a $n \times rp$ matrix, then

$$A \ltimes \delta_p^i = \mathrm{Blk}_i^p(A). \tag{A3}$$

A.6: Notations used in the body of this paper.

- δ_n^i : the i-th column of the identity matrix I_n.
- Δ_n: the set $\{\delta_n^1, \delta_n^2, \cdots, \delta_n^n\}$.
- $\delta_n^{i_1,i_2,\cdots,i_k}$: the column vector whose i_1-th, i_2-th, $\cdots$, i_k-th components are 1s, others are 0s.
- $\mathrm{Col}_i(A)$: the i-th column of matrix A.
- $\mathrm{Row}_i(A)$: the i-th row of matrix A.
- $A(i, j)$: the (i, j)-th element of matrix A.

* Corresponding author (email: yyyan@mail.nankai.edu.cn)

Appendix B Proofs in the body of the letter.

B.1: Proof of Theorem 1

The proof will be completed by two stages. In the first stage, we will prove the single step evolution of M reading a single alphabet $u(t) = a_j$ at state x_i, that is

$$x(2) = F \ltimes x_i \ltimes a_j. \tag{B1}$$

According to Definition 2, without loss of generality, set the reachable state packet of a_j from x_j is $x(2) = \{x_{j_1}, x_{j_2} \cdots, x_{j_l}\}$. The vector form of $x(2)$ is

$$\begin{aligned} x(2) &= \begin{matrix}[0\cdots 0 & 1 & 0\cdots 0 & 1 & 0\cdots 0 & 1 & 0\cdots 0]^T_{1\times n} \\ & \uparrow & & \uparrow & & \uparrow & \\ & j_1 & & j_2 & \cdots & j_l & \end{matrix} \\ &= \delta_n^{j_1, j_2, \cdots, j_l}. \end{aligned} \tag{B2}$$

On one hand, we know from Definition 2 that when M reads an input alphabet $u(1) = a_j$ at state x_i the following equations (B3) holds.

$$\begin{aligned} &f(x_i, a_j) = x_{j_1}, \\ &f(x_i, a_j) = x_{j_2}, \\ &\cdots \\ &f(x_i, a_j) = x_{j_l}. \end{aligned} \tag{B3}$$

The vector form of (B3) is

$$f(x_i, a_j) = x(2). \tag{B4}$$

On the other hand, from Definition 4, we know that the column-wise form of the structure matrix F is just

$$\begin{aligned} \mathrm{Col}_j(F_i) &= \begin{matrix}[0\cdots 0 & 1 & 0\cdots 0 & 1 & 0\cdots 0 & 1 & 0\cdots 0]^T_{1\times n} \\ & \uparrow & & \uparrow & & \uparrow & \\ & j_1 & & j_2 & \cdots & j_l & \end{matrix} \\ &= \delta_n^{j_1, j_2, \cdots, j_l} \\ &= x(2). \end{aligned} \tag{B5}$$

Finally, by Proposition 3, we get

$$\begin{aligned} F \ltimes x_i \ltimes a_j &= F \ltimes \delta_n^i \ltimes \delta_m^j \\ &= [F_1 \cdots F_i \cdots F_n] \ltimes \delta_n^i \ltimes \delta_m^j \\ &= F_i \ltimes \delta_m^j. \end{aligned} \tag{B6}$$

By Proposition 3 again, equation (B6) can be further simplified as

$$F \ltimes x_i \ltimes a_j = F_i \ltimes \delta_m^j = \mathrm{Col}_j(F_i). \tag{B7}$$

Combining equations (B7), (B5) and (B2), we have that

$$x(2) = \mathrm{col}_j(F_i) = F \ltimes x_i \ltimes a_j. \tag{B8}$$

This completes the first stage.

We next use inductive method to prove the second stage, i.e., equation $x(t+1) = F^t \ltimes x_i \ltimes u(t)$. Suppose M reads sequence $u(t) = a_{i_1} a_{i_2} \cdots a_{i_t} \in \Sigma^*$ at state x_i, then by Proposition 2, we have

$$\begin{aligned} &f(x_i, u(t)) \\ &= f(x_i, a_{i_1} a_{i_2} \cdots a_{i_t}) \\ &= f(f(x_i, a_{i_1}), a_{i_2} \cdots a_{i_t}) \\ &= f(f(f(x_i, a_{i_1}), a_{i_2}), a_{i_3} \cdots a_{i_t}) \\ &\vdots \\ &= \underbrace{f(f \cdots f(f(f}_{t}(x_i, a_{i_1}), a_{i_2}), \cdots, a_{i_{t-1}}), a_{i_t}). \end{aligned} \tag{B9}$$

By equation (B8)) and inductive method, the above equation (B9) can be rewritten as

$$\begin{aligned}
&f(x_i, u(t)) \\
&= \underbrace{f(f\cdots f(f}_{t}(x_i, a_{i_1}), a_{i_2}), \cdots, a_{i_{t-1}}), a_{i_t}) \\
&= \underbrace{f(f\cdots f}_{t-1}(F \ltimes x_i \ltimes a_{i_1}, a_{i_2}\cdots a_{i_{t-1}}), \cdots, a_{i_{t-1}}), a_{i_t}) \\
&= \underbrace{f(f\cdots f}_{t-2}(F \ltimes F \ltimes x_i \ltimes a_{i_1} \ltimes a_{i_2}, a_{i_3}), a_{i_{t-1}}), a_{i_t}) \\
&\vdots \\
&= f(\underbrace{F \ltimes F \ltimes \cdots \ltimes F}_{t-1} \ltimes x_i \ltimes a_{i_1} \ltimes a_{i_2} \ltimes a_{i_3} \ltimes \cdots \ltimes a_{i_{t-1}}, a_{i_t}) \\
&= \underbrace{F \ltimes F \ltimes \cdots \ltimes F}_{t} \ltimes x_i \ltimes a_{i_1} \ltimes a_{i_2} \ltimes a_{i_3} \ltimes \cdots \ltimes a_{i_{t-1}} \ltimes a_{i_t} \\
&= F^t \ltimes x_i \ltimes u(t).
\end{aligned} \tag{B10}$$

This completes the proof.

B.2: Proof of Theorem 2

It is easy to see that $\sum_{k=1}^{m^t} \mathrm{Blk}_i^n(F^t)(j,k) \neq 0$ is equivalent to that $\mathrm{Row}_j(\mathrm{Blk}_i^n(F^t)) \neq \vec{0}$, $\vec{0}$ is the zero vector. We first prove the necessity.

(Necessity) If state x_j is t-controllable from state x_i, according to Definition 7, there exists a sequence $u(t) = a_{i_1}a_{i_2}\cdots a_{i_t} \in \Sigma^*$ moving M to state x_j from state x_i. By Definition 2, we have

$$x_j = f(x_i, u(t)). \tag{B11}$$

Using Theorem 1, equation (B11) can be rewritten as

$$x_j = F^t \ltimes x_i \ltimes u(t). \tag{B12}$$

Let the vector forms of states x_i and x_j are δ_n^i and δ_n^j, respectively, and let the vector form of sequence $u(t)$ be $u(t) = a_{i_1}a_{i_2}\cdots a_{i_t} = \delta_m^{i_1} \ltimes \delta_m^{i_2} \ltimes \cdots \ltimes \delta_m^{i_r} = \delta_{m^t}^l$.

Substitute the vector forms of x_i, x_j and $u(t)$ into equation (B12), we get

$$\delta_n^j = F^t \ltimes \delta_n^i \ltimes \delta_{m^t}^l. \tag{B13}$$

According to Proposition 3, equation (B13) can be represented as

$$\delta_n^j = \mathrm{Blk}_i^n(F^t) \ltimes \delta_{m^t}^l, \tag{B14}$$

in which, $\mathrm{Blk}_i^n(F^t)$ is the i-th block of the n-column-equal-division of F^t.

Using Proposition 3 again, equation (B14) can be further expressed as

$$\delta_n^j = \mathrm{Col}_l(\mathrm{Blk}_i^n(F^t)).$$

Thus there is at least one non-zero element in $\mathrm{Row}_j\left(\mathrm{Blk}_i^n(F^t)\right)$, i.e., $\mathrm{Row}_j(\mathrm{Blk}_i^n(F^t)) \neq \vec{0}$. The necessity is then proved.

(Sufficiency). If $\mathrm{Row}_j(\mathrm{Blk}_i^n(F^t)) \neq \vec{0}$, without loss of generality, suppose the k-th element of $\mathrm{Row}_j(\mathrm{Blk}_i^n(F^t))$ is $n_k (n_k \in N^+)$, that is, the (j,k)-th element of F^t is n_k, i.e.,

$$\mathrm{Blk}_i^n(F^t)(j,k) = n_k. \tag{B15}$$

By equation (B15), we know there are n_k different paths from state x_i to state x_j, therefore there is an input string $u(t) = a_{i_1}a_{i_2}\cdots a_{i_t}$ in Σ^* such that M can move to x_j from x_i by $u(t)$. According to Definition 7, state x_j is t-controllable from state x_i. The sufficiency is proved.

B.3: Proof of Proposition 4

Consider the arbitrary value of v_1 in Δ_m, without loss of generality, take $v_1 = \delta_m^k$ for example. According to (A3), we have

$$\begin{aligned}
&D_{[m,n]} \ltimes v_1 \ltimes v_2 \\
&= (\mathbf{1}_m \otimes I_n) \ltimes v_1 \ltimes v_2 \\
&= [I_n, \cdots, I_n] \ltimes \delta_m^k \ltimes v_2 \\
&= I_n \ltimes v_2 \\
&= v_2.
\end{aligned}$$

Thus $D_{[m,n]} \ltimes v_1 \ltimes v_2 = v_2$.

It follows from equation (A2) that

$$\begin{aligned}
&D_{[n,m]} \ltimes W_{[m,n]} \ltimes v_1 \ltimes v_2 \\
&= D_{[n,m]} \ltimes v_2 \ltimes v_1 \\
&= v_1.
\end{aligned}$$

The proof is then completed.

B.4: Proposition 5 (Extracting components from the STP of vectors): For $v_1, v_2, \cdots v_n \in \Delta_m$, assume that the STP of them is

$$\ltimes_{i=1}^{n} v_i = \delta_{m^n}^{s}, \tag{B16}$$

then

$$v_i = D_{[m^{n-1}, m]} \ltimes W_{[m^i, m^{n-i}]} \ltimes \delta_{m^n}^{s}, \quad (i = 1, 2, \cdots, s),$$

where $D_{[m^{n-1},m]}$ is $D_{[m,n]} = \mathbf{1}_m \otimes I_n$, $\mathbf{1}_m = [\underbrace{1, \cdots, 1}_{m}]$.

Proof: Left multiplying both sides of equation (B16) by $W_{[m^i,m^{n-i}]}$ yields

$$W_{[m^i,m^{n-i}]} v_1 \cdots v_{i-1} v_i v_{i+1} \cdots v_n = W_{[m^i,m^{n-i}]} \delta_{m^n}^{s},$$

By Proposition 1, we see that

$$v_{i+1} \cdots v_n v_1 \cdots v_{i-1} v_i = W_{[m^i,m^{n-i}]} \delta_{m^n}^{s}. \tag{B17}$$

Left multiplying both sides of equation (B17) by $D_{[m^{n-1},m]}$ results in

$$D_{[m^{n-1},m]} v_{i+1} \cdots v_n v_1 \cdots v_{i-1} v_i = D_{[m^{n-1},m]} W_{[m^i,m^{n-i}]} \delta_{m^n}^{s}.$$

According to Proposition 4, we get

$$v_i = D_{[m^{n-1},m]} W_{[m^i,m^{n-i}]} \delta_{m^n}^{s}.$$

The proof is completed.

B.5: Proof of Theorem 3

It is clear that the following equation

$$\prod_{s \in \{i_1, i_2, \cdots, i_k\}} \mathrm{Blk}_i^n(F^t)(s, l) \neq 0. \tag{B18}$$

is equivalent to

$$\mathrm{Col}_l(\mathrm{Blk}_i^n(F^t)) = \begin{array}{ccccccccc} [0\cdots0 & n_1 & 0\cdots0 & n_2 & 0\cdots0 & n_k & 0\cdots0]_{1\times n}^{T}. \\ & \uparrow & & \uparrow & & \uparrow & \\ & i_1 & & i_2 & \cdots & i_k & \end{array} \tag{B19}$$

where n_{i_j} are positive numbers; F is the structure matrix of M.

We next use (B19) to prove the theorem (the reason why we use (B18) as the conclusion and use (B19) to prove is that (B18) has explicit algebraic meaning, while (B19) is easy to describe the proof).

(Necessity) If the t-step reachable state packet of input sequence $u(t)$ from state x_i is $R(x_i, u(t)) = \{x_{i_1}, x_{i_2}, \cdots, x_{i_k}\}$, by Definition 3, the following equation holds for any $j \in \{1, 2, \cdots, k\}$.

$$f(x_i, u(t)) = x_{i_j}. \tag{B20}$$

According to Theorem 1, we get

$$\begin{aligned} F^t x_i u(t) &= x(t+1) \\ &= \begin{array}{ccccccccc} [0\cdots0 & n_1 & 0\cdots0 & n_2 & 0\cdots0 & n_k & 0\cdots0]_{1\times n}^{T}. \\ & \uparrow & & \uparrow & & \uparrow & \\ & i_1 & & i_2 & \cdots & i_k & \end{array} \end{aligned} \tag{B21}$$

Let the vector forms of x_i and $u(t)$ be δ_n^i and $\delta_{m^t}^l$, respectively. Then the left side of equation (B21) is

$$F^t \ltimes \delta_n^i \ltimes \delta_{m^t}^l. \tag{B22}$$

By Proposition 3, equation (B22) can be represented as

$$\mathrm{Blk}_i^n(F^t) \ltimes \delta_{m^t}^l = \mathrm{Col}_l(\mathrm{Blk}_i^n(F^t)). \tag{B23}$$

From equations (B23) and (B21), we get

$$\mathrm{Col}_l(\mathrm{Blk}_i^n(F^t)) = \begin{array}{ccccccccc} [0\cdots0 & n_1 & 0\cdots0 & n_2 & 0\cdots0 & n_k & 0\cdots0]_{1\times n}^{T}. \\ & \uparrow & & \uparrow & & \uparrow & \\ & i_1 & & i_2 & \cdots & i_k & \end{array}$$

The necessity is proved.

(Sufficiency) Suppose that

$$\mathrm{Col}_l(\mathrm{Blk}_i^n(F^t)) = \begin{array}{ccccccccc} [0\cdots0 & n_1 & 0\cdots0 & n_2 & 0\cdots0 & n_k & 0\cdots0]_{1\times n}^{T}. \\ & \uparrow & & \uparrow & & \uparrow & \\ & i_1 & & i_2 & \cdots & i_k & \end{array}$$

Using Proposition 3 twice, we have

$$\mathrm{Col}_l(\mathrm{Blk}_i^n(F^t)) = \mathrm{Blk}_i^n(F^t) \ltimes \delta_{m^t}^l = F^t \ltimes \delta_n^i \ltimes \delta_{m^t}^l. \tag{B24}$$

Set $X = \{x_1, x_2, \cdots, x_n\}$ and $\Sigma = \{a_1, a_2, \cdots, a_m\}$. Let the vector forms of x_i and a_j be δ_n^i and δ_m^j, respectively, where $i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, m$, then there must be t alphabets $a_{i_1}, a_{i_2}, \cdots, a_{i_t}$ satisfying equation $a_{i_1} \ltimes a_{i_2} \ltimes \cdots \ltimes a_{i_t} = \delta_{m^t}^l$. Note that $u(t) = a_{i_1} \ltimes a_{i_2} \ltimes \cdots \ltimes a_{i_t} = \delta_{m^t}^l$, then equation (B24) can be rewritten as

$$\begin{array}{c} [0\cdots0 \quad \underset{\substack{\uparrow \\ i_1}}{n_1} \quad 0\cdots0 \quad \underset{\substack{\uparrow \\ i_2}}{n_2} \quad 0\cdots0 \quad \underset{\substack{\uparrow \\ i_k}}{n_k} \quad 0\cdots0]_{1\times n}^T = F^t x_i u(t). \end{array} \tag{B25}$$

By Theorem 1, we have that equation (B25) is the result of the transition of M when M reads $u(t) = a_{i_1} \ltimes a_{i_2} \ltimes \cdots \ltimes a_{i_t}$ at state x_i. In other words, M reaches the state packet $\{x_{i_1}, x_{i_2}, \cdots, x_{i_k}\}$, i.e. $R(x_i, u(t)) = \{x_{i_1}, x_{i_2}, \cdots, x_{i_k}\}$. Then the sufficiency is proved.

B.6: Computational complexity analysis of Algorithms 1 and 2

The time computational complexity of Algorithms 1 and 2 is $O(n)$, where n is the number of the states of the considered FSM. The analysis is as follows.

Let us first look at Algorithm 1. Steps 1 and 2 are executed only once throughout the algorithm. Step 3 runs at most n times in the worst situation where all of the n elements of $\mathrm{Row}_j\left(\mathrm{Blk}_i^n(F^t)\right)$ are non-zero. For the same reason, the set K contains at most n elements. Consequently, Steps 4 and 5 repeat n times in the worst case, respectively. These show that the loop iterations of Algorithm 1 are total $3n$ times in the worst situation. Therefore the time complexity is $O(n)$.

For Algorithm 2, it is easy to get, using the similar way, that the worst time complexity is also $O(n)$.

References

1 Cheng D, Qi H, Zhao Y. An Introduction to Semi-tensor Product of Matrices and Its Applications. Singapore: World Scientific Publishing Co. Pte. Ltd., 2012:23-26

2 Yan Y, Yue J, Fu Z, et al. Algebraic criteria for finite automata understanding of regular language. Frontiers of Computer Science, 2019,13: 1148-1150