High-Order Temporal Effects in Finite-State Machine Analysis

Diana Marculescu, Radu Marculescu, Massoud Pedram

Department of Electrical Engineering - Systems
University of Southern California, Los Angeles, CA 90089

Abstract - The objective of this paper is to illustrate, analytically and quantitatively, the effect of high-order temporal correlations on steady-state and transition probability calculations in Finite State Machines (FSMs) analysis. As the main theoretical contribution, it is shown that, if the sequence feeding the target circuit has temporal correlations of order k, then a lag-k Markov chain model of the sequence will suffice to model correctly the joint transition probabilities of the primary inputs and internal states in the target circuit. From the experimental point of view, it is shown that assuming temporal independence or using first-order temporal models is not sufficient, that is, the inaccuracies induced in steady-state and transition probability calculations are significant for most of the analyzed benchmarks.

I. INTRODUCTION

In the last decade, probabilistic approaches have received a lot of attention as a viable alternative to deterministic techniques for analyzing complex digital systems. Logic synthesis [1], verification [2], testing [3] and more recently, low-power design [4] have benefited from using probabilistic techniques. In particular, the behavior of FSMs has been investigated using concepts from the Markov chain theory.

Studying the behavior of the Markov chain provides us with different variables of interest of the original FSM. In this direction [5][6] are excellent references where steady-state and transition probabilities (as variables of interest) are successfully estimated in large FSMs. Both techniques are analytical in nature and, in order to manage complexity, have made some simplifying assumptions, temporal independence on the primary inputs being the most notable one. These assumptions, however, limit the applicability and usefulness of the results. As we will prove in this paper, in the case of FSMs, this is not enough for accurate estimation of transition probabilities. Temporal correlations longer than one time step can affect the overall behavior of the FSM and therefore, result in very different values for the actual transition probabilities compared to those predicted analytically. More interestingly, it will be shown that, if one ignore the effects of finite-order statistics at the primary inputs of the FSM, it is possible to wrongly predict non-zero steady-state probabilities for some transient states which normally occur in the beginning of operation but disappear shortly afterwards, once the machine have reached its steady-state regime. This kind of erroneous prediction for the state occupation probabilities is especially harmful in timing verification, when one tries to find the probability of occurrence of an event related to a state or transition in the circuit.

Addressing these issues, the present paper aims to document the effect of finite-order statistics of the input statistics on the FSM behavior and hopefully, stimulate further research toward incorporating temporal effects in steady-state and transition probability calculations. The analysis itself relies on modeling the binary input streams as high-order Markov sources of information [7] that emit symbols at each time step, according to some probabilistic rule.

From the original results provided here, two are noteworthy for probabilistic FSM analysis:

- If the sequence feeding the target circuit has temporal correlations of order k, then a lag-k Markov chain model of the sequence will suffice to model correctly the joint transition probabilities of the primary inputs and internal states in the target circuit. Consequently, the probability distribution on states can be correctly extracted.
- If the input sequence has order two or higher then modeling it as a lag-one Markov Chain cannot exactly preserve the first-order joint transition probabilities (primary inputs and internal states) in the target circuit.

The paper is organized as follows: in Section II we present the main results about the effect of finite-order statistics on FSM behavior. Section III discusses the perturbation of the input sequence and its effects on the transition probabilities calculation. Section IV gives some experimental results for common sequential benchmarks. Finally, we conclude by summarizing our main contribution.

II. HIGH-ORDER MARKOV SOURCES OF INFORMATION

A. Finite-order Markov chains

First, we focus on the input sequence that feeds a target circuit and we model it as an information source. Without loss of generality, we restrict ourselves to finite binary strings, that is, finite sequences consisting only of 0's and 1's. The set of events of interest is the set S of all finite binary sequences. A particular sequence S_1 in S consists of vectors v_1, v_2,..., v_n (which may be distinct or not), each having a positive occurrence probability. Indices 1, 2,..., n represent the discrete time steps when a particular vector occurs in sequence which is applied to a target circuit.

Proposition 1 [8]. The occurrence probability of a vector string $v = v_1 v_2 ... v_n$ is given by:

$$p(v) = p(v_1) \cdot p(v_2 | v_1) \cdot \ldots \cdot p(v_n | v_{n-1} v_{n-2} \ldots v_1) \quad (1)$$

where the conditional probabilities are uniquely defined by $p(x|v) = p(vx) / p(v)$.

Example 1: Let S_1 and S_2 be two 2-bit sequences, of length 48, as shown in Fig.1a. These two sequences, have exactly the same set of first-order temporal statistics that is, they cannot be distinguished as far as wordwise one-step transition probabilities are concerned. In fact, in Fig.1b we provide the wordwise transition graph for these two sequences. Each node in this graph is associated to a distinct pattern that occurs in S_1 and S_2 (the upmost bit is the most significant one, e.g. in S_1, v_1 = '1', v_2 = '2', v_3 = '3',..., v_{48} = '1'). Each edge represents a valid transition between any two valid patterns and has a nonzero probability associated with it. For instance, the pattern '3' in S_1 and S_2 is always followed by '1' (thus the edge between nodes '3' and '1' has the probability 1) whereas it is equally likely to have either '0', '2' or '1' after pattern '1'. Starting with different initial states and using a random number generator we may, of course, generate other sequences equivalent with S_1 and S_2 as far as the one-step transition probabilities are concerned.

S_1

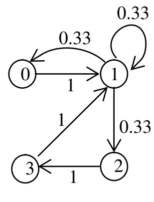
v_{48} ... $v_3v_2v_1$

0110110110000110000001101100000000000011000000000110
1101101101111101101011011010111010111010110101101

S_2

001100001100001100001100001100001100001100001100
011011011011011011011011011011011011011011011011

(a)



(b)

Fig.1

We can then see the graph in Fig.1b as a compact, canonical, characterization of sequences S_1 and S_2. Suppose that we want to compute the occurrence probability of string v = '01 10' that is, the probability that transition $1 \rightarrow 2$ is taking place in S_1. To this effect, we just apply formula (1) and then $p(v) = p(v_1 v_2) = p(v_1) \cdot p(v_2 | v_1)$ which gives us the value of 1/6. If we are interested in finding the two-step transition probability $0 \rightarrow 1 \rightarrow 1$ in S_2, then applying once again (1) we get the value of 1/6.

An important class of information sources is the class of Markov sources [9] which can be conveniently modeled as Markov chains of finite-order.

Definition 1. (lag-k Markov chain). A discrete stochastic process $\{v_n\}_{n \geq 1}$ is said to be a lag-k Markov chain if at any time step $n \geq k+1$:

$$p(v_n | v_{n-1} v_{n-2} \ldots v_1) = p(v_n | v_{n-1} v_{n-2} \ldots v_{n-k}) \tag{2}$$

In particular, any lag-one Markov source, is characterized by the set of internal states (nodes in the corresponding graph representation) and the set of one-step conditional probabilities q_{ij} that give the transition probability from state v_i to the next state v_j. It should be noted that any lag-k Markov chain can be reduced to a lag-one Markov chain based on the following result.

Proposition 2 [8]. If $\{u_n\}_{n \geq 1}$ is a lag-k Markov chain then $\{v_n\}_{n \geq 1}$, where $v_n = (u_n, u_{n+1}, \ldots, u_{n+k-1})$, is a multivariate first-order Markov chain.

B. The effect of finite-order statistics of the input sequences on FSM behavior

Now we turn our attention from the input sequence to the circuit and investigate the effect of input statistics on the transition probabilities (for primary inputs and present state lines) in the target circuit. As shown in Fig.2, we model the 'tuple' (*input_sequence, target_circuit*) by the 'tuple' (*Markov_chain, target_circuit*), where *Markov_chain* models the *input_sequence* and *target_circuit* is the sequential machine where the transition probabilities have to be determined. In what follows, x_n, s_n will denote the inputs and states of the target sequential machine; $p(x_n s_n)$ is the probability that the input is x_n and the state is s_n at time step n. As usual, *'next'* denotes the next state function of the FSM and *'out'* represents the output function.

We are interested in defining the joint probabilities $p(x_n s_n)$ and $p(x_n s_n x_{n-1} s_{n-1})$ because, as we can see in Fig.2, they capture the characteristics of the input (primary inputs and present state lines) that feeds the next state and the output logic of the target circuit. Under the general assumption of stationarity and ergodicity we can prove the following result:

Theorem 1[1]**.** If the sequence feeding a target sequential circuit has order k, then a lag-k Markov chain which correctly models the

[1] All proofs are available upon request.

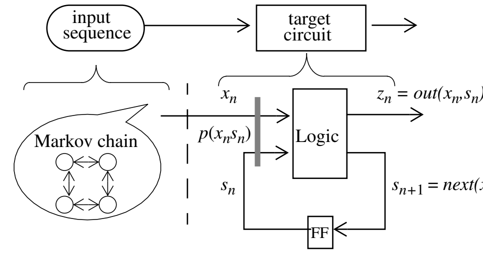


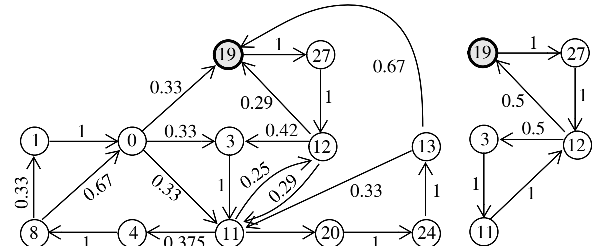
Fig.2

input sequence, also correctly models the k-step conditional probabilities of the primary inputs and internal states in the target circuit, that is:

$$\begin{aligned} p(x_n s_n | x_{n-1} s_{n-1} x_{n-2} s_{n-2} \ldots x_{n-k} s_{n-k}) &= \\ = p(x_n | x_{n-1} x_{n-2} \ldots x_{n-k}) \end{aligned} \tag{3}$$

We note therefore that preserving order-k statistics implies also that order-k statistics will be captured for inputs and states. In general, modeling an order-k source by a lower order source may introduce large inaccuracies as it is shown in the next example.

Example 2: We consider once again the sequences S_1 and S_2 that feed the benchmark *dk17* and illustrate that indeed, if the input sequence has order two, then modeling it as a lag-one Markov chain will *not* preserve the first-order joint transition probabilities (primary inputs and internal states) in the target circuit. We simulated the benchmark *dk17 (*starting with the same initial state *'19')* for both sequences and we present in Fig.3 (Fig.3a is for S_1 and Fig.3b is for S_2) the wordwise transition graphs obtained for the signal lines ($x_n s_n$).



(a) (b)

Fig.3

The benchmark *dk17* has 2 primary inputs and 3 FFs therefore in Fig.3, any node is decimally encoded using 5 bits. For instance, the initial state *'19'* corresponds to the binary code '10011' that is, '10' for primary inputs and '011' for state lines. After that, applying '11' on the primary inputs, the present state becomes '011', therefore we enter the node '11011' = 27 in Fig.3. As we can see, because S_1 can be modeled as a first order Markov source, while S_2 must be modeled as a second order Markov source, the corresponding transition graphs are quite different. From a practical point of view, this means that if one underestimates a high-order source (for instance, assuming that second- or higher-order temporal correlations are not important), then one may end up not preserving even the first-order transition probabilities in the target circuit.

III. PERTURBATION ANALYSIS FOR THE INPUT SEQUENCE

Without loss of generality, we will refer subsequently only to first-order Markov chains but, by virtue of Proposition 2, all results translate easily to lag-k Markov chains. We also assume

.

Table 1: Steady-state probability distributions for the actual input sequences

Circuit	# of primary inputs	# of FFs	# of reached States	State probability distribution
bbara	4	4	3	[0.5 0.25 0.25]
bbsse	7	4	6	[0.23 0.01 0.01 0.05 0.18 0.51]
bbtas	2	3	6	[0.06 0.33 0.33 0.06 0.06 0.17]
cse	7	4	6	[0.88 0.01 0.02 0.04 0.03 0.03]
dk17	2	3	2	[0.67 0.33]
donfile	2	5	5	[0.17 0.17 0.17 0.17 0.33]
ex1	9	5	13	[0.16 0.05 0.20 0.12 0.11 0.01 0.26 0.06 0.001 0.004 0.01 0.01 0.01]
planet	7	6	34	[0.04 0.04 0.01 0.05 0.04 0.01 0.04 0.05 0.01 0.03 0.04 0.01 0.01 0.05 0.05 0.06 0.01 0.01 0.03 0.01 0.09 0.03 0.03 0.01 0.01 0.01 0.03 0.05 0.04 0.01 0.04 0.03 0.03 0.01]
sand	11	5	32	[0.07 0.01 0.01 0.01 0.07 0.03 0.003 0.004 0.02 0.03 0.003 0.01 0.004 0.04 0.02 0.01 0.06 0.04 0.03 0.01 0.01 0.01 0.05 0.06 0.14 0.03 0.03 0.01 0.07 0.03 0.05 0.04]

cardinalities l and m for the input and reachable state spaces, respectively, and introduce a new formulation based on matrices, mostly because this is a convenient notation for our purpose.

What we are interested in is the probability distribution on the state lines, that is the probability that the FSM occupies a particular state. However, the only information that we have at this point is the characterization of the inputs. As we can see in Fig.2, the first-order Markov chain at the primary inputs x_n can be characterized by the matrix of conditional probabilities $Q_X = (q_{ij})_{1 \le i,j \le l}$, where $q_{ij} = p(x_n = v_j | x_{n-1} = v_i)$ and $\{v_1, v_2, \ldots, v_l\}$ represents the whole set of possible vectors at the primary inputs. On the other hand, the Markov chain defined jointly for the primary inputs and state lines (x_n, s_n) can be characterized by the matrix $Q_{XS} = (r_{ip,jq})_{\substack{1 \le i,j \le l \\ 1 \le p,q \le m}}$, where $r_{ip,jq} = p(x_n s_n = v_j u_q | x_{n-1} s_{n-1} = v_i u_p)$ and $\{u_1, u_2, \ldots, u_m\}$ represents the set of reachable states. From this joint characterization we can easily derive the state probabilities as $p(s_n) = \sum_{all\ x_n} p(x_n s_n)$.

Proposition 3. The matrix Q_{XS} can be written in the following form:

$$Q_{XS} = \begin{bmatrix} q_{11}B_1 & q_{12}B_1 & \ldots & q_{1l}B_1 \\ q_{21}B_2 & q_{22}B_2 & \ldots & q_{2l}B_2 \\ \ldots & \ldots & & \ldots \\ q_{l1}B_l & q_{l2}B_l & \ldots & q_{ll}B_l \end{bmatrix} \quad (4)$$

where $\{B_i\}_{1 \le i \le l}$ is a set of $m \times m$ degenerate[1] stochastic matrices defining the next state function for input v_i; specifically, if $B_i = (b^i_{pq})_{1 \le p,q \le m}$ then $b^i_{pq} = \begin{cases} 1 & if\ \ next(v_i, u_p) = u_q \\ 0 & otherwise \end{cases}$.

We assume that for the matrix Q_X that *exactly* characterizes the original input sequence, we construct an *approximate* matrix Q'_X which preserves up to an error ε $(0 < \varepsilon < 1)$ the entries in the original matrix. The following results hold for the FSM in Fig.2.

Lemma 1. If matrix Q_X is preserved up to an $\varepsilon > 0$, then matrix Q_{XS} is also preserved up to the same ε.

Theorem 2. If matrix Q_X is preserved up to an $\varepsilon > 0$, then steady-state probabilities for the primary inputs and state lines are also preserved up to a polynomial function of ε $(0(\varepsilon)$, read as '*zero of epsilon*') having the form $k_1\varepsilon + k_2\varepsilon^2 + \ldots$.

Corollary 1. If matrix Q_X is preserved up to an $\varepsilon > 0$, then transition probabilities for the primary inputs and state lines for two consecutive time steps are asymptotically close to the original ones, that is $|p(x_n s_n x_{n-1} s_{n-1}) - p'(x_n s_n x_{n-1} s_{n-1})| = 0(\varepsilon)$ where p and p' represent the probabilities that correspond to matrices Q and Q', respectively.

Corollary 2. If the matrix Q_X is preserved up to an $\varepsilon > 0$, the probability distribution for the state lines (that is the probability of being in any state of the FSM) is asymptotically close to the initial one.

Thus, we may conclude that the probability distribution for the state lines is asymptotically close to the correct one only if the temporal correlations on the primary inputs are correctly taken into account. We will estimate in the next section the effect of ignoring the temporal effects on the primary inputs upon the accuracy of estimation.

IV. Experimental Results

In this section we provide our experimental results for some common benchmark circuits. In particular, we compare the probability distribution for the states where the order of the input sequence is assumed to be either zero or one, to the case where the actual order of the source is extracted from the input sequence.

In Tables 1 and 2 we provide the experimental results for different initial sequences having the total length of 10,000 vectors; these sequences were produced using a second order information source based on Fibonacci series. We give in Table 1 the number of reached states and their probability distribution when the input sequence is correctly modeled as having order two. In Table 2, we provide a comparison between the results obtained

[1] That is, the elements of the matrices B_i are only 0 or 1.

Table 2: Zeroth- and first-order model vs. second-order model

	Zeroth-order model					First-order model				
Circuit	# of reached states	MAX%	MEAN%	MAX	MEAN	# of reached states	MAX%	MEAN%	MAX	MEAN
bbara	10	46.63	36.89	0.2304	0.0784	10	49.40	35.91	0.2469	0.0786
bbsse	8	59.62	22.29	0.0211	0.0079	6	22.12	13.76	0.0458	0.0167
bbtas	6	64.93	24.07	0.1083	0.0429	6	55.70	24.28	0.0929	0.0406
cse	8	52.88	20.12	0.0078	0.0046	6	19.23	7.48	0.0070	0.0024
dk17	8	47.22	35.44	0.3147	0.0984	7	37.48	29.96	0.2498	0.0927
donfile	24	84.70	69.80	0.2514	0.0589	6	52.70	26.76	0.0912	0.0592
ex1	15	192.31	35.97	0.0308	0.0089	13	50.00	18.03	0.0269	0.084
planet	48	431.95	48.85	0.1879	0.0123	48	513.79	67.14	0.2235	0.0153
sand	32	103.57	25.32	0.0196	0.0047	32	103.57	24.49	0.0247	0.0040

for the state probabilities assuming a zeroth-order model (i.e. temporal independence on the primary inputs) or a first-order model (i.e. considering temporal correlations only between two consecutive clock cycles) against the true second-order model.

For each of these two cases we report the total number of reached states when a zeroth- or first-order model (columns 2 and 7) is used. The difference between these values and the corresponding ones from Table 1 represent the number of states that are wrongly predicted as having non-zero occurrence probability. This actually shows a fundamental limitation for both zeroth- and first-order models compared to the true second-order one. In the next five columns, we present the classical measures for reporting the error (maximum percentage error for the reached states MAX%, mean percentage error MEAN%, maximum absolute error MAX, and the mean absolute error MEAN) in estimating the steady state probabilities. They also provide a quantitative measure of the impact of using an inappropriate temporal correlation model for the primary inputs. For example, in the case of benchmark circuit *dk17*, a correct temporal model of order two would provide for this sequence a number of two reached states (out of the total number of $2^3 = 8$ states). However, assuming temporal independence on the primary inputs, the number of reached states increases to 8 and thus 6 new states are predicted as having a non-zero probability of occurrence. For the first-order model, we still have 5 extra states that are wrongly estimated. We can see that in all cases a lower order model covers all the states from the original one, but it may also introduce a significant number of extra states and, furthermore, the quality of the results in estimating the steady-state probabilities is seriously impaired.

V. Conclusion

In this paper we investigated from a probabilistic point of view the effect of finite-order statistics of the input sequence on FSMs behavior. In particular, the effect of temporal correlations longer that one clock-cycle was analyzed for steady-state and transition probabilities calculations.

The results presented in this paper can be used in synthesis and verification and represent an important step towards understanding the FSM behavior from a probabilistic point of view.

References

[1] K. Parker and E.J. McCluskey, ‘Probabilistic Treatment of General Combinational Networks’, in *IEEE Trans. on Computers*, vol. C-24, June 1975.

[2] G. Hachtel, E. Macii, A. Pardo, and F. Somenzi, ‘Probabilistic Analysis of Large Finite State Machines’, in *Proc. ACM/IEEE Design Automation Conference*, pp. 270-275, June 1994.

[3] J. Savir, G.S. Ditlow, and P.H. Bardell, ‘Random Pattern Testability’, in *IEEE Trans. on Computers*, vol. C-33, Jan. 1984.

[4] M. Pedram, ‘Power Minimization in IC Design: Principles and Applications’, in *ACM Transactions on Design Automation of Electronic Systems,* vol.1, no.1, pp.1-54, Jan.1996.

[5] G.D. Hachtel, E. Macci, A. Pardo, and F. Somenzi, ‘Markovian Analysis of Large Finite State Machines’, in *IEEE Trans. on CAD*, vol.15, no.12, Dec. 1996.

[6] C.-Y. Tsui, J. Monteiro, M. Pedram, S. Devadas, A.M. Despain, and B. Lin, ‘Power Estimation Methods for Sequential Logic Circuits’, in *IEEE Trans. on VLSI Systems*, vol.3, no.3, Sept. 1995.

[7] R. Marculescu, D. Marculescu, and M. Pedram, ‘Adaptive Models for Input Data Compaction for Power Simulators’, in *Proc. IEEE Asian-South Pacific Design Automation Conference*, Japan, Jan. 1997.

[8] A. Papoulis, ‘Probability, Random Variables, and Stochastic Processes’, McGraw-Hill Co., 1984.

[9] T.M. Cover, and J.A. Thomas, ‘Elements of Information Theory’, John Wiley & Sons, Inc., 1991.

[10] J.G. Kemeny, J.L. Snell, and A.W. Knapp, ‘D2enumerable Markov Chains’, D.Van Nostrand Company, Inc., 1966.

[11] J.H. Wilkinson, ‘The Algebraic Eigenvalue Problem’, Clarendon Press, 1988.