An Improved Genetic Algorithm for the Inference of Finite State Machine

Nattee Niparnan
Computer Engineering Department
Chulalongkorn University,
Bangkok, Thailand.
nattee.n@student.chula.ac.th

Prabhas Chongstitvatana
Computer Engineering Department
Chulalongkorn University,
Bangkok, Thailand
prabhas@chula.ac.th

Abstract

This work proposes a new method for the use of genetic algorithms to synthesize a finite state machine consistent with a given input/output sequence set. The new approach improves the method previously used by many researchers by omitting the evolution of output function of a FSM. The output will later be inferred from the input/output sequence set. The results show that our method is able to solve a larger problem size than the previous approach.

1 INTRODUCTION

This work proposes the use of genetic algorithms to solve the problem of synthesizing a finite state machine consistent with a given input/output sequence set. Our motivation is to mimic the target machine by synthesizing a compact machine that is consistent with partial input/output sequences observed from the target machine.

It is well known that finding a minimum size deterministic finite automaton consistent with a set of given samples is NP-complete. Furthermore, Pitt and Warmuth (1993) show that even finding a DFA whose size is polynomial on the size of minimum solution is also NP-complete. Oliveira and Silva (2001) propose a search method which is known to be the current state of the art on finding a minimum size FSM that is consistent with a given input/output sequence set.

The new method improves the former method by reducing the search space of the problem and introduces a new evaluation method that is more effective in guiding the search. The former method evolves a transition function and an output function. Our method omits the evolution of the output function. To infer the output, the method feeds the input sequence to the evolving machine. For each transition of the evolving machine, the frequency of each output that is mapped to the transition is counted. The most frequent output is chosen to be the output for each transition. If every transition have only one output and its frequency is more than zero then that FSM is consistent with the target machine. Fitness function is calculated from the frequency of each output on every transition. The fitness value is the sum of all maximum frequency outputs of all transitions.

Series of experiments were set up. The new method is compared with a previous method (Manovit *et al*, 1998). 800 machines were randomly generated. The size of the machine varies from 2 to 25 states and the output size varies from 1 to 3 bits. An input/output sequence set was randomly sampling from each machine. The new method shows much improvement both on the time and the number of problem solved

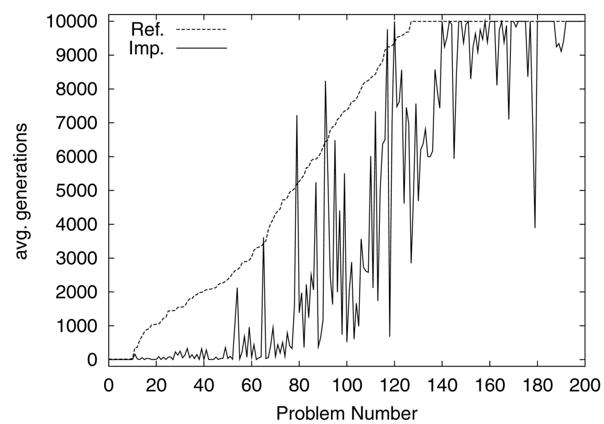

Figure 1 : Average Generation

Figure 1 shows the number of average generation used by our method (labeled Imp.) and the former method (labeled Ref.) on problems with 3 bit output. The data were sorted by the value of former method. The new method shows clearly the improvement over the former method.

2 REFERENCES

L. Pitt and M. Warmuth (1993). The minimum consistent DFA problem cannot be approximated within any polynomial. Journal of ACM, 40(1) pp.95–142.

A. L. Oliveira and J. P. M. Silva (2001). Efficient Algorithms for the Inference of Minimum Size DFAs, Machine Learning 44 pp.93-119

C. Manovit, C. Aporntewan, and P. Chongstitvatana, (1990). Synthesis of Synchronous Sequential Logic Circuits from Partial Input/Output Sequences. Proc. of 2nd Int. Conf. on Evolvable Systems (ICES98), pp. 98-105.