Chapter 4
Experiments and Results

This chapter illustrates the concept of experiments on computing Unique Input Output Sequences (UIOs) for conformance testing of the finite state machines. The experiments will observe in several scenarios for the simple FSM and the real-world FSM. All results of these experiments based on the hypothesis of this work which is “The Self-Adaptive Evolutionary Algorithm should be faster in the performance of computing UIO sequences of finite state machines than the Genetic Algorithm”. Two programs will run three experiments for observing the different results of the fitness values of UIOs. There is the analysis of the distribution of the sample data between two samples by using the Wilcoxon Rank Sum Test. Three experiments for observing and analysing the performance between the GA and Self-Adaptive EA as follows.

Experiment 1: Consideration in computing UIOs of the small FSM (4 states) and large FSM (55 states) in some different lengths of input sequences. Some lengths are identified at 3, 5, 10, 15, 20, 30 and 50 for observing the fitness values.

Experiment 2: Consideration in computing UIOs of some FSMs in the different number of states in the FSMs with the same lengths of the input sequences. The states’ number will be supposed at 4, 10, 20 and 40 for observing the fitness values.

Experiment 3: Consideration in computing UIOs of the real FSMs. The real FSMs are ‘train4’, ‘bbtas’, ‘lion9’, ‘train11’ and ‘dk512’.

4.1 Experiment 1: Observation of the GA and the Self-Adaptive EA when considering the different lengths of input sequences.

The programs will be set to compute UIOs of the small FSM which contains four states. In this experiment, the programs will run with the different length of the input sequences (individuals) namely 3, 5, 10, 15, 20, 30 and 50. Such input sequences (individuals) will be generated at random. The population contains 100 individuals and is run at 1000 iterations. The GA and Self-Adaptive EA will run at 250 times. Two programs will compute the fitness values and also calculate the average of such fitness values. The results of both programs will be plotted on the graph.

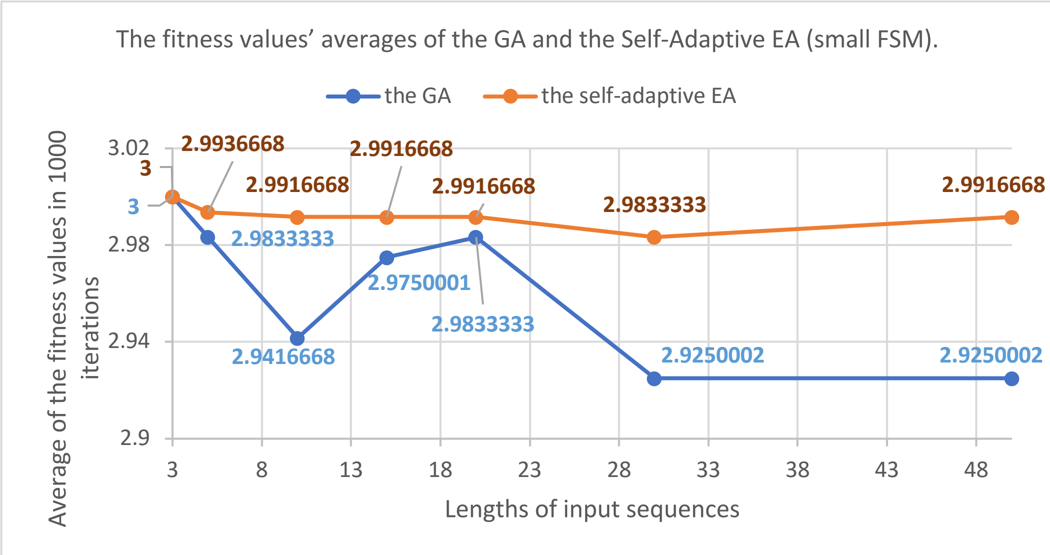
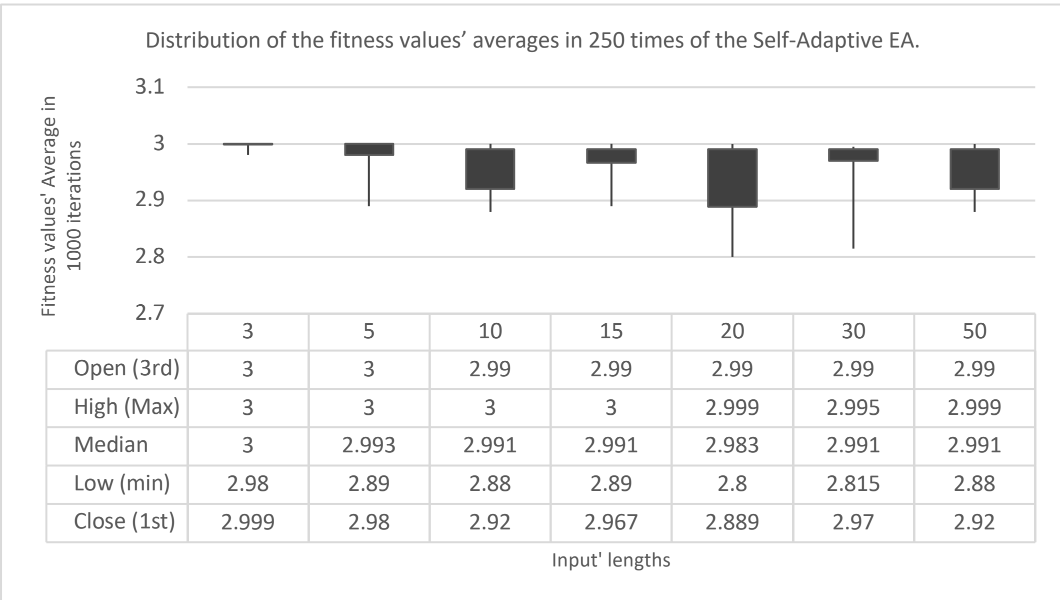


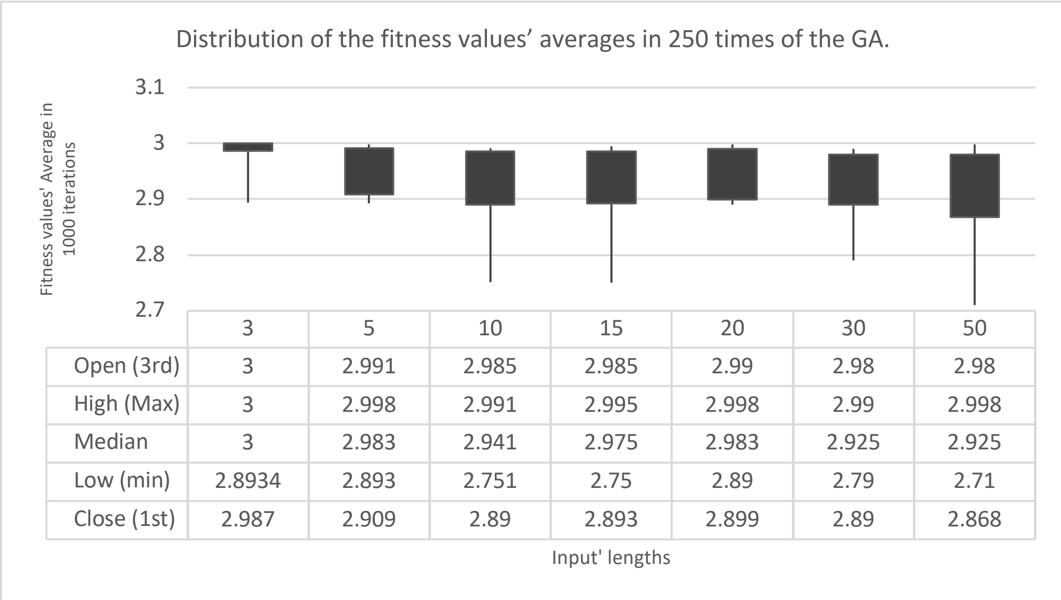
Figure 4.1 the fitness values' averages of the GA and the Self-Adaptive EA (Simple FSM).

The graph in Figure 4.1 shows the difference of the fitness values' averages when considering in different length of several input sequences in the simple FSM of the GA and the Self-Adaptive EA. The X-axis of this graph presents the lengths of the input sequences (individuals) and the Y-axis of this graph displays the averages of the fitness values. The programs run by changing the input's length at 3, 5, 10, 15, 20, 30 and 50 respectively. Obviously, in the simple finite state machine, the fitness values' averages of the GA are between 3.00 and 2.925. In the Self-Adaptive EA, the fitness values' averages are between 3.00 and 2.983. The fitness values' averages of the GA and Self-Adaptive EA decreased steadily.



	3	5	10	15	20	30	50
Open (3rd)	3	3	2.99	2.99	2.99	2.99	2.99
High (Max)	3	3	3	3	2.999	2.995	2.999
Median	3	2.993	2.991	2.991	2.983	2.991	2.991
Low (min)	2.98	2.89	2.88	2.89	2.8	2.815	2.88
Close (1st)	2.999	2.98	2.92	2.967	2.889	2.97	2.92

Figure 4.2 the distribution of the fitness values' averages in 250 running times of the Self-Adaptive EA.



	3	5	10	15	20	30	50
Open (3rd)	3	2.991	2.985	2.985	2.99	2.98	2.98
High (Max)	3	2.998	2.991	2.995	2.998	2.99	2.998
Median	3	2.983	2.941	2.975	2.983	2.925	2.925
Low (min)	2.8934	2.893	2.751	2.75	2.89	2.79	2.71
Close (1st)	2.987	2.909	2.89	2.893	2.899	2.89	2.868

Figure 4.3 the distribution of the fitness values' averages in 250 running times of the GA.

The distribution of the fitness values' averages of the Self-Adaptive EA and the GA is shown in the box plots in Figure 4.2 and 4.3 shows the results of each program such as highest value, lowest value and median. It can be seen that the Self-Adaptive EA is quite stable than the GA.

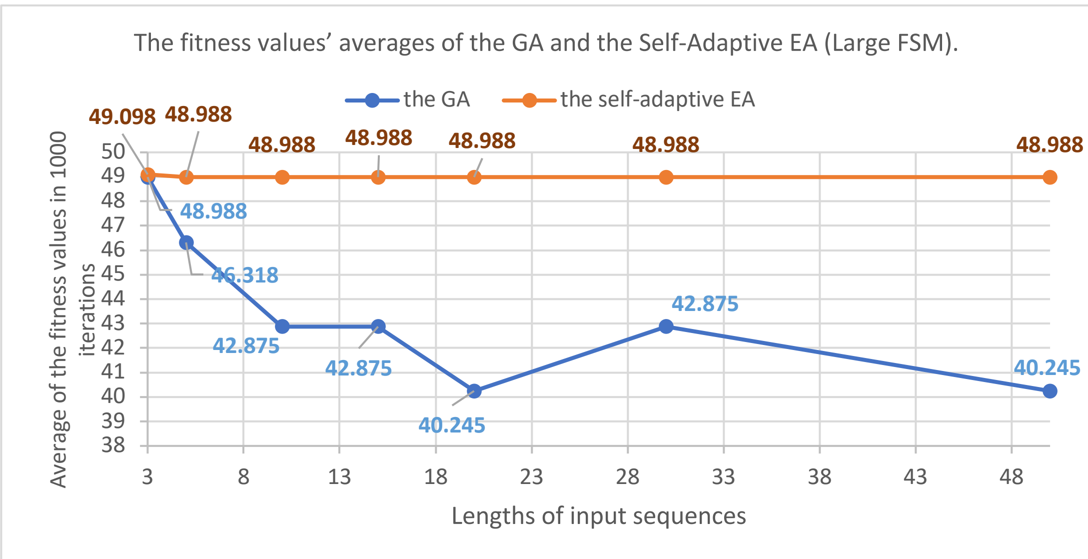
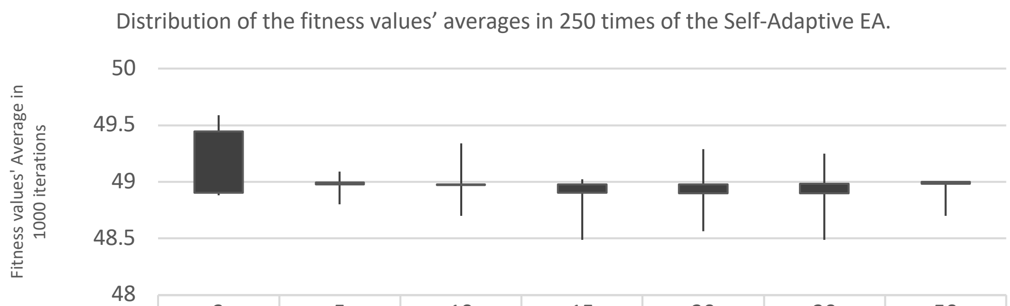


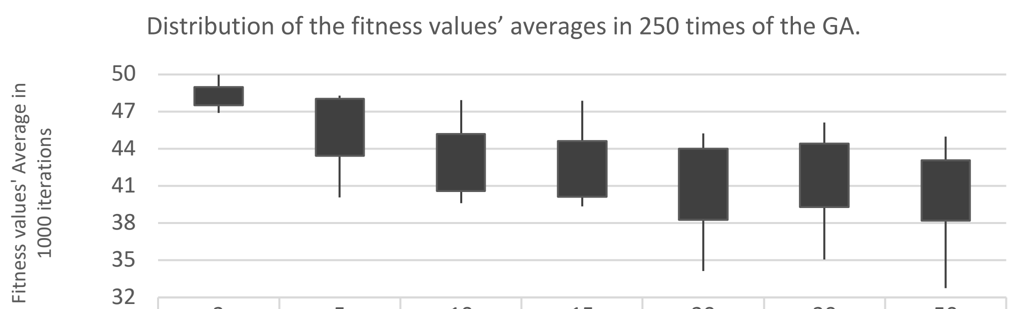
Figure 4.4 The fitness values' averages of the GA and the Self-Adaptive EA (Large FSM).

Furthermore, there is a consideration of the large FSM in this experiment. The graph in Figure 4.2 shows the difference of the fitness values' averages when considering in different length of several input sequences in the large FSM (55 states) of the GA and the Self-Adaptive EA. The X-axis of this graph presents the lengths of the input sequences (individuals) and the Y-axis of this graph displays the averages of the fitness values. The programs run by changing the input's length at 3, 5, 10, 15, 20, 30 and 50 respectively. Obviously, in the simple finite state machine, the fitness values' averages of the GA are between 48.988 and 40.245. In the Self-Adaptive EA, the fitness values' averages are between 49.098 and 48.988. The fitness values' averages of the GA and Self-Adaptive EA decreased steadily.



	3	5	10	15	20	30	50
Open (3rd)	49.445	48.995	48.98	48.98	48.98	48.981	48.981
High (Max)	49.588	49.09	49.34	49.023	49.288	49.25	49
Median	49.098	48.98	48.98	48.98	48.98	48.98	48.98
Low (min)	48.883	48.8	48.7	48.488	48.566	48.488	48.701
Close (1st)	48.901	48.98	48.98	48.901	48.899	48.899	48.998

Figure 4.5 the distribution of the fitness values' averages in 250 running times of the Self-Adaptive EA.



	3	5	10	15	20	30	50
Open (3rd)	48.98	48.05	45.17	44.644	44.025	44.432	43.084
High (Max)	49.98	48.295	47.908	47.889	45.229	46.145	44.995
Median	48.988	46.328	42.475	42.875	40.245	42.875	40.245
Low (min)	46.885	40.088	39.588	39.35	34.135	35.067	32.75
Close (1st)	47.494	43.45	40.58	40.135	38.27	39.295	38.2

Figure 4.6 the distribution of the fitness values' averages in 250 running times of the GA.

The distribution of the fitness values' averages of the Self-Adaptive EA and the GA is shown in the box plots in Figure 4.5 and 4.6 shows the results of each program such as highest value, lowest value and median. Obviously, it can be seen that the Self-Adaptive EA can present the results better and more stable than the GA.

Overall, the different lengths of various input sequences can affect to decrease slightly the fitness values' averages in the Self-Adaptive EA when we consider in both small and large FSMs. While the fitness values' averages went down gradually and fluctuated in the GA.

4.2 Experiment 2: Consideration of the GA and the Self-Adaptive EA when observing the different artificial FSMs.

In this experiment, we suppose the artificial FSMs which identify some states and transitions. Each artificial FSM will contain the many states and several transitions. The programs will be set to compute UIOs to find the fitness values of the artificial FSMs which have 4, 10, 20 and 40 states from the same length of input sequences (input's length = 10). We suppose the length of the input sequences at 10 which will be generated at random. The population will be generated at 100 individuals and is run at 1000 iterations in each time. The GA and Self-Adaptive EA will run at 250 times. Two programs will compute the fitness values and also calculate the average of such fitness values.

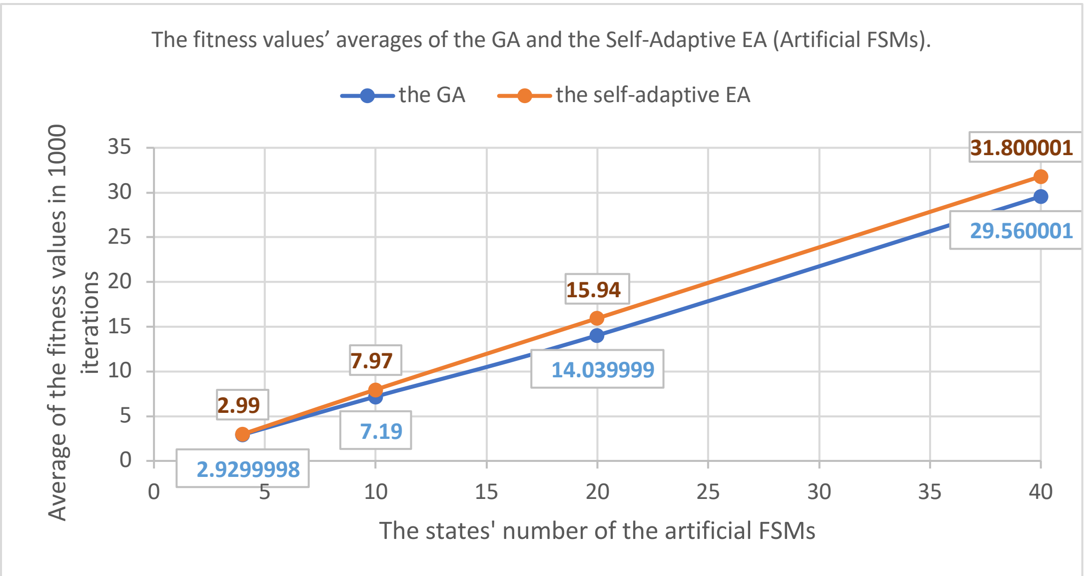
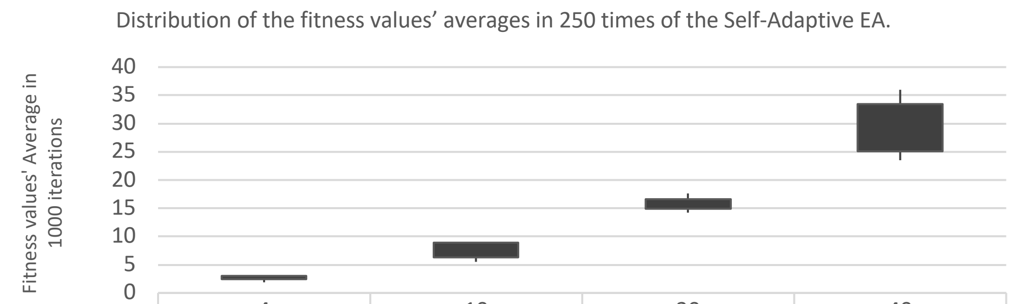


Figure 4.7 The fitness values' averages of the GA and the Self-Adaptive EA (Artificial FSMs).

In Figure 4.3, the graph displays the averages of the fitness values when computing UIOs of the several artificial FSMs with the same input length between the GA and the Self-Adaptive

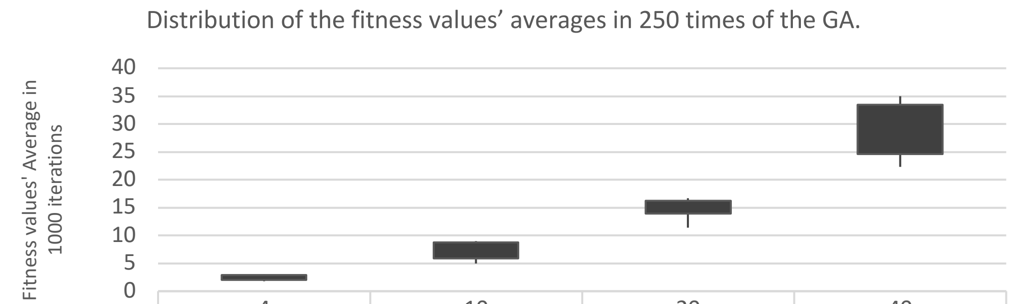
EA. The X-axis presents the states' number of different artificial FSMs and the Y-axis displays the averages of the fitness values. The averages of the fitness values rose in both programs. However, the Self-Adaptive EA can get a little higher fitness values' averages than the GA. Hence, it can be noted that the number of the states in the artificial FSMs can affect to increase the fitness values.



	4	10	20	40
Open (3rd)	2.985	8.892	16.65	33.454
High (Max)	2.98	8.95	17.665	35.989
Median	2.998	7.97	15.94	31.8
Low (min)	1.889	5.54	14.186	23.559
Close (1st)	2.439	6.357	14.937	25.065

The number of states

Figure 4.8 the distribution of the fitness values' averages in 250 running times of the Self-Adaptive EA.



	4	10	20	40
Open (3rd)	2.952	8.764	16.231	33.45
High (Max)	2.965	8.989	16.65	34.985
Median	2.922	7.19	14.039	29.56
Low (min)	1.788	4.95	11.454	22.35
Close (1st)	2.028	5.883	13.899	24.608

The number of states

Figure 4.9 the distribution of the fitness values' averages in 250 running times of the GA.

The distribution of the fitness values' averages of the Self-Adaptive EA and the GA is shown in the box plots in Figure 4.8 and 4.9 shows the results of each program such as highest value, lowest value and median. Obviously, it can be seen that the Self-Adaptive EA can present the results better and more stable than the GA when we consider the artificial FSMs.

From this experiment, there is the description of the fitness values' averages on computing UIOs of the different cases of FSMs the artificial FSMs. There is the Wilcoxon Rank Sum Test which computes a p-value for presenting the different of the fitness' averages between the GA and the Self-Adaptive EA when observing in the different artificial FSMs. This test can be used to explore the different data of two programs. Table 4.1 displays the p-values of each artificial FSMs.

Artificial FSMs' Name	Number of States	Number of input	Number of output	Average of the GA in 250 times	Average of the Self-Adaptive EA in 250 times	p-value of each FSM of 250 times
FSM_1	4	1	1	2.929	2.998	0.3947456
FSM_2	10	1	1	7.19	7.97	0.2725045
FSM_3	20	1	1	14.039	15.94	0.2886283
FSM_4	40	1	1	29.56	31.8	0.1871255

Table 4.1 the description and p-value of each artificial FSM

4.3 Experiment 3: Consideration of the GA and the Self-Adaptive EA when observing in different states of the real FSMs.

This experiment presents the distribution of the fitness values' averages between the GA and the Self-Adaptive EA when observing on computing UIOs of the real FSMs with the same length of input sequences (individuals). In this experiment, we suppose the input's length as 10. The population contains 100 individuals and each time is run at 1000 iterations. The GA and Self-Adaptive EA will run at 250 times. The graph which represents the results of this experiment shows in Figure 4.3. In this graph, the X-axis presents the states' number of different artificial FSMs and the Y-axis displays the averages of the fitness values. The real FSMs is namely 'train4', 'bbtas', 'lion9', 'train11' and 'dk512' (Yang, 1991).

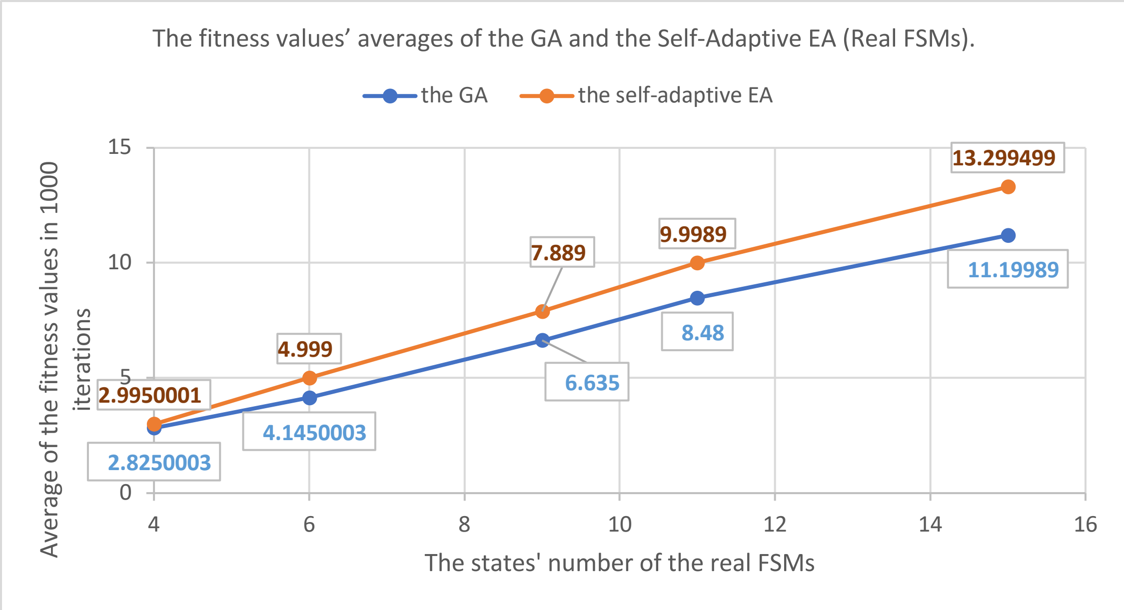
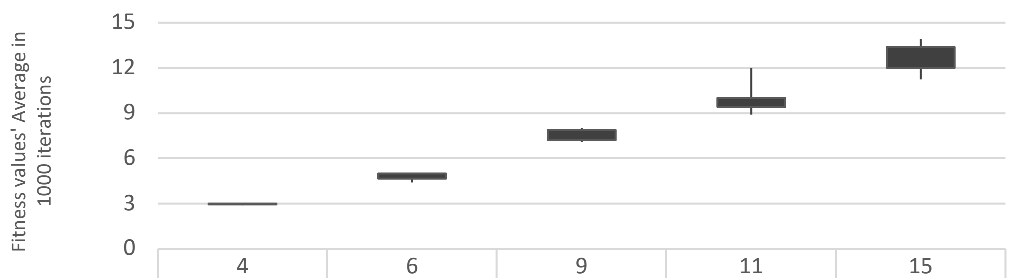


Figure 4.10 The fitness values' averages of the GA and the Self-Adaptive EA (Real FSMs).

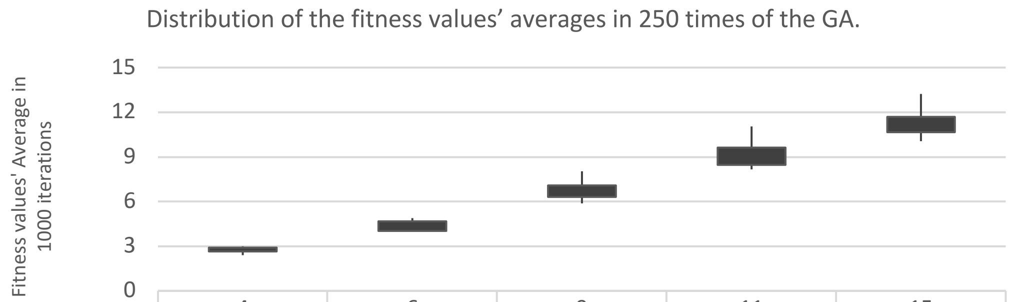
In Figure 4.10, the graph displays the averages of the fitness values when computing UIOs of the several real FSMs with the same input length as 10 bit-strings between the GA and the Self-Adaptive EA. Overall, the Self-Adaptive EA can get higher fitness values than the GA. The number of the states can affect to increase the fitness values' averages. Hence, it can be noted that the Self-Adaptive EA can compute UIOs for the real FSMs better than the GA. In addition, in the real FSMs, the input sequences are sometimes read more than one bit-string to obtain some output sequences. This constraint can affect the computation of UIOs of the real FSMs.



	4	6	9	11	15
Open (3rd)	2.995	4.999	7.899	9.998	13.385
High (Max)	2.999	4.999	8.019	11.99	13.899
Median	2.995	4.999	7.899	9.998	13.299
Low (min)	2.889	4.399	7.088	8.915	11.228
Close (1st)	2.948	4.648	7.199	9.407	11.997

Input' lengths

Figure 4.11 the distribution of the fitness values' averages in 250 running times the Self-Adaptive EA.



	4	6	9	11	15
Open (3rd)	2.921	4.675	7.099	9.6355	11.689
High (Max)	2.995	4.899	8.022	11.028	13.228
Median	2.825	4.145	6.635	8.48	11.199
Low (min)	2.399	4.012	5.89	8.165	10.035
Close (1st)	2.648	4.0189	6.2995	8.48	10.646

Figure 4.12 the distribution of the fitness values’ averages in 250 running times the GA.

The distribution of the fitness values’ averages of the Self-Adaptive EA and the GA is shown in a type of the box plot in Figure 4.11 and 4.12 shows the results of each program such as highest value, lowest value and median. Obviously, it can be seen that the Self-Adaptive EA can present the results better and more stable than the GA when we consider the artificial FSMs.

From this experiment, there is the description of the fitness values’ averages on computing UIOs of the different cases of FSMs the artificial FSMs. There is the Wilcoxon Rank Sum Test which computes a p-value for presenting the different of the fitness’ averages between the GA and the Self-Adaptive EA when observing in the different artificial FSMs. This test can be used to explore the different data of two programs. Table 4.2 displays the p-values of each artificial FSMs.

FSMs’ Name	Number of States	Number of input	Number of output	Average of the GA in 250 times	Average of the Self-Adaptive EA in 250 times	p-value of each real FSM
train4	4	2	1	2.8250003	2.9950001	0.22185304
bbtas	6	2	2	4.1450003	4.999	0.16544372
lion9	9	2	1	6.635	7.889	0.09936437
Train11	11	2	1	8.48	9.9989	0.11732495
dk512	15	1	3	11.19989	13.299499	0.08355062

Table 4.2 the description and p-value of each FSM

From such experiments, they display the fitness values' averages on computing UIOs of the different cases of FSMs namely the small and large FSMs, the artificial FSMs and the real FSMs. The Wilcoxon Rank Sum Test computes a p-value for presenting the different of the fitness' averages between the GA and the Self-Adaptive EA. This test can be used to explore the different data of two programs.

For the condition of the p-value, if the p-value is as 1.0, it means that there is no different data between two programs. If the p-value is more than 0.05, it means that there are different data between two programs but it is not different too much for desiring that the higher one is really worth enough to say as the better one. On the other hand, If the p-value is less than 0.05, it means that the data of two programs are different enough to assure that the higher one is really better than another one. Table 4.3 illustrates the significant p-value of three experiments.

Experiments	Scenarios' description	p-values
1	Observation of the GA and the Self-Adaptive EA when considering on computing UIOs of the small FSM in different lengths of input sequences.	0.0093240095
	Observation of the GA and the Self-Adaptive EA when considering on computing UIOs of the large FSM in different lengths of input sequences.	0.003933
2	Consideration of the GA and the Self-Adaptive EA when observing on computing UIOs of the different artificial FSMs. (Average p-value of Table 4.1)	0.28574985
3	Consideration of the GA and the Self-Adaptive EA when observing computing UIOs of the real FSMs. (Average p-value of Table 4.2)	0.13750734

Table 4.3 the p-value of three experiment