MA513

Parallel Computing

Assignment-9 Report

Paranjay Bagga
160101050

Problem Statement :-

Write a parallel program that takes, a **Deterministic Finite Automata-DFA**(*M*) and a string *w*, as input, and returns true if M accepts w, and false otherwise. Report speedup with respect to the no. of processors and the size of string used.

System Specifications :-

All the experiments have been performed on **Dell Inspiron 5559** laptop with **Intel i5-6200U dual core processor** and **8 GB RAM**. Each core has **2 threads.** Also, it was made insured that no other applications were running in the background during each experiment which could have incurred biased readings.

Implementation Details :-

1. The parallel DFA-run algorithm used in the program is referred from the paper **"Implementation of Deterministic Finite Automata on Parallel Computers" - Jan Holub - 2009**. The whole string is divided equally among each processor/thread, giving each one of them a substring of length |w|(=*N*). Each processor runs the basic sequential DFA algorithm for their own substring part. But, the main problem is that all the threads, other than the first thread that knows the initial state to start running the DFA with, do not

know where to start running their DFA algorithm. Hence, all these threads consider all the possible states in the DFA, as their initial state, sequentially, and store the corresponding info of the mapping from the initial state, that each of them started, to the last active state they reached upon running the DFA algorithm. Now, these mappings are reduced with both sequential and binary reduction in the program so that the final result (about the last state reached) can be made available to the first thread. This last state is further compared to the final state so as to deduce the acceptance of the string.

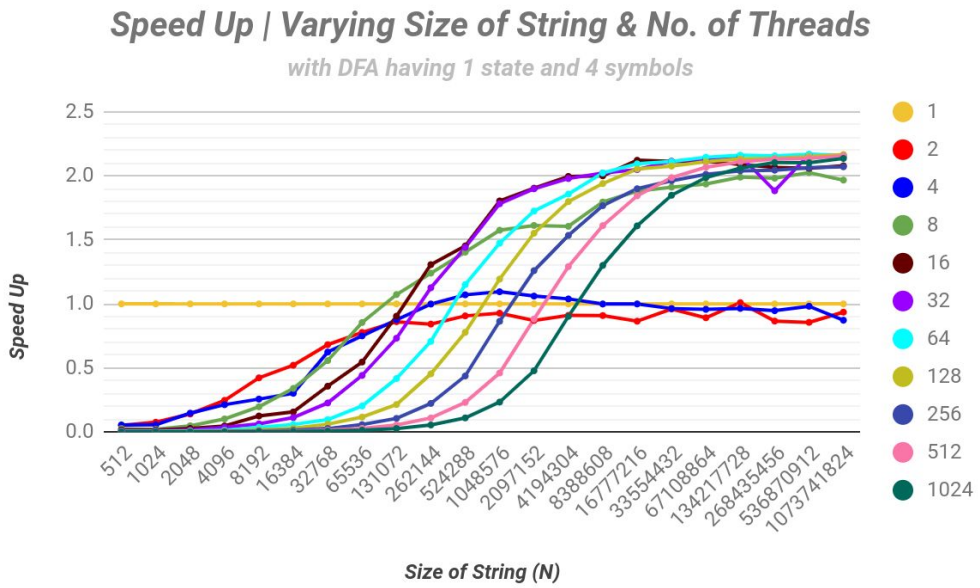
2. The input to the program are- length of string (*N*), no. of producer threads (*P*), and, for easability, the custom File No. (*fileNum*) (having details regarding the particular no. of states & the Transition Table used).
3. The time complexity for the sequential DFA run is ***O(N)***. Also, as mentioned in the above mentioned paper, the time complexity for the parallel version of DFA run (on Symmetric shared-memory multiprocessors - SMP) is :-
    a. **O((N*Q)/P + log(P) + Q*log(P))** - using Binary Reduction
    b. **O((N*Q)/P + log(P) + P)** - using Sequential Reduction
4. Thus, in the program, using the above facts, the binary reduction is considered only when the factor of **Q*log(P)** is strictly less than **P**, else sequential reduction is considered always.
5. The program has been implemented using **Pthreads**.

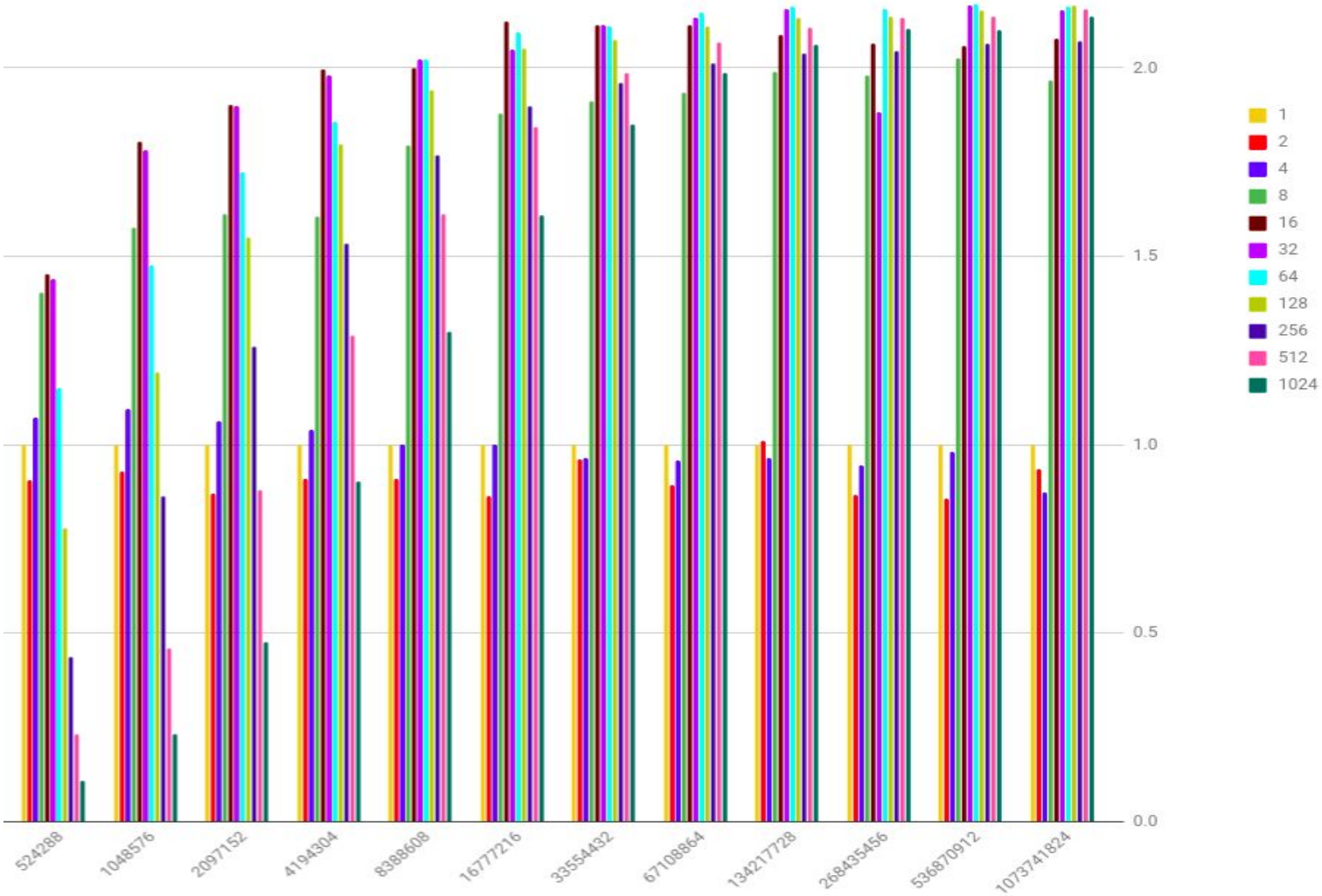
Experiments :-

The no. of threads used is taken from **2^(0)=1 to 2^(10)=1024**.
The size of the string is taken from **2^(9)=512 to 2^(30)=1073741824**.
The total no. of states(in various files) are taken as **1,2,4,5** & **8**.
Experiment results are shown in the tables and graphs(with comparison) below.
The values in the table are average values of 8 (or 4 or 2 for larger N) executions to avoid any unexpected behaviour introduced as a result of thread switching.

1. Speed Up | Varying Size of String & No. of Threads | with DFA having 1 state and 4 symbols :-

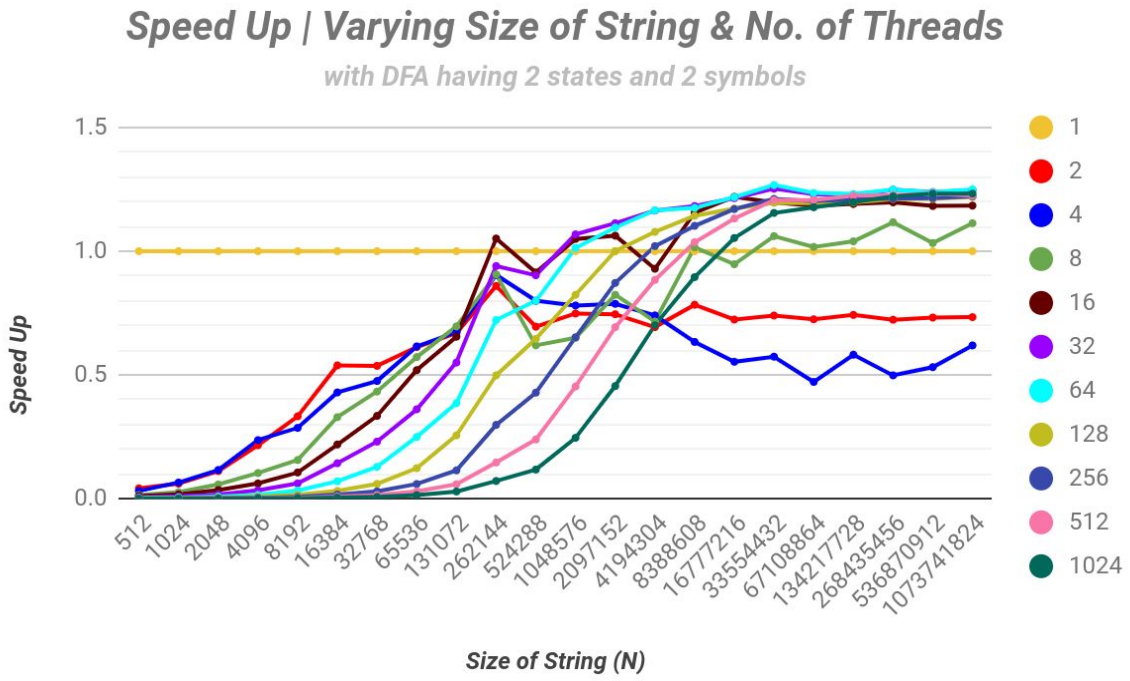
Size of String (N)	No. of Threads (P)										
	1	2	4	8	16	32	64	128	256	512	1024
512	1	0.0515	0.0551	0.0266	0.0154	0.0085	0.0044	0.0021	0.0011	0.0005	0.0003
1024	1	0.0762	0.0562	0.0209	0.0137	0.0083	0.0039	0.0021	0.0010	0.0005	0.0003
2048	1	0.1400	0.1467	0.0482	0.0279	0.0135	0.0075	0.0039	0.0022	0.0010	0.0005
4096	1	0.2467	0.2131	0.1017	0.0469	0.0362	0.0131	0.0075	0.0038	0.0018	0.0009
8192	1	0.4226	0.2567	0.1967	0.1249	0.0633	0.0364	0.0159	0.0081	0.0043	0.0023
16384	1	0.5210	0.3021	0.3400	0.1563	0.1122	0.0587	0.0289	0.0140	0.0069	0.0036
32768	1	0.6816	0.6240	0.5573	0.3573	0.2265	0.0958	0.0604	0.0269	0.0148	0.0070
65536	1	0.7771	0.7504	0.8549	0.5462	0.4414	0.2030	0.1161	0.0574	0.0285	0.0142
131072	1	0.8608	0.8753	1.0729	0.9040	0.7323	0.4173	0.2143	0.1059	0.0543	0.0268
262144	1	0.8425	0.9988	1.2401	1.3069	1.1263	0.7079	0.4537	0.2228	0.1091	0.0545
524288	1	0.9065	1.0712	1.4043	1.4528	1.4410	1.1506	0.7779	0.4372	0.2309	0.1098
1048576	1	0.9276	1.0949	1.5754	1.8044	1.7820	1.4745	1.1936	0.8640	0.4603	0.2334
2097152	1	0.8709	1.0611	1.6129	1.9038	1.8988	1.7247	1.5522	1.2599	0.8810	0.4779
4194304	1	0.9102	1.0391	1.6061	1.9958	1.9792	1.8579	1.7992	1.5352	1.2915	0.9029
8388608	1	0.9090	0.9992	1.7959	2.0004	2.0215	2.0228	1.9399	1.7672	1.6115	1.3008
16777216	1	0.8650	0.9998	1.8804	2.1220	2.0489	2.0940	2.0536	1.8997	1.8444	1.6096
33554432	1	0.9608	0.9636	1.9120	2.1134	2.1137	2.1120	2.0762	1.9620	1.9865	1.8486
67108864	1	0.8918	0.9574	1.9363	2.1139	2.1328	2.1455	2.1116	2.0120	2.0674	1.9866
134217728	1	1.0098	0.9649	1.9901	2.0893	2.1565	2.1616	2.1322	2.0393	2.1088	2.0607
268435456	1	0.8658	0.9469	1.9815	2.0660	1.8837	2.1565	2.1368	2.0444	2.1318	2.1037
536870912	1	0.8561	0.9817	2.0260	2.0593	2.1670	2.1686	2.1516	2.0638	2.1360	2.1023
1073741824	1	0.9358	0.8726	1.9669	2.0786	2.1523	2.1612	2.1665	2.0709	2.1567	2.1361

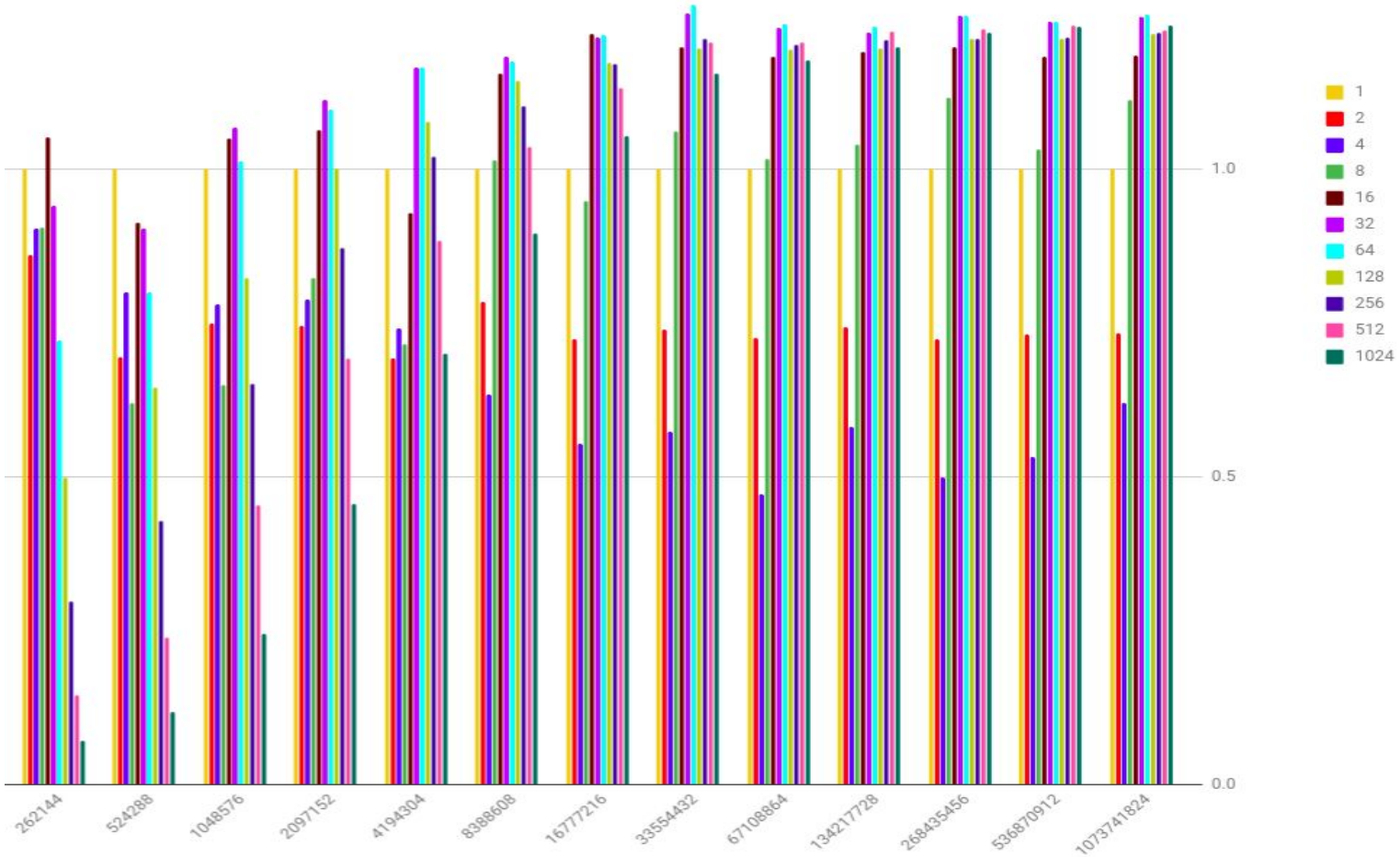

Speed Up | Varying Size of String & No. of Threads
with DFA having 1 state and 4 symbols
Speed Up
2.5
2.0
1.5
1.0
0.5
0.0
512
1024
2048
4096
8192
16384
32768
65536
131072
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
Size of String (N)
1
2
4
8
16
32
64
128
256
512
1024


2.0
1.5
1.0
0.5
0.0
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
1
2
4
8
16
32
64
128
256
512
1024

2. Speed Up | Varying Size of String & No. of Threads | with DFA having 2 states and 2 symbols :-

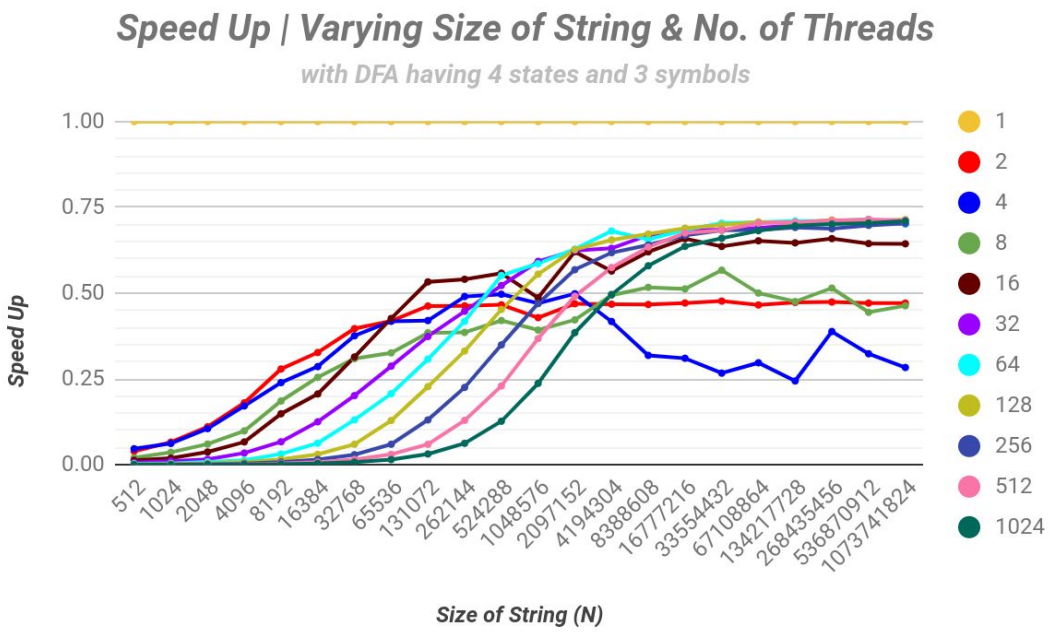
Size of String (N)	No. of Threads (P)										
	1	2	4	8	16	32	64	128	256	512	1024
512	1	0.0418	0.0313	0.0148	0.0108	0.0058	0.0031	0.0015	0.0008	0.0004	0.0002
1024	1	0.0608	0.0653	0.0262	0.0173	0.0091	0.0048	0.0024	0.0012	0.0006	0.0003
2048	1	0.1118	0.1160	0.0577	0.0350	0.0167	0.0088	0.0041	0.0022	0.0011	0.0005
4096	1	0.2160	0.2366	0.1037	0.0617	0.0338	0.0153	0.0084	0.0039	0.0021	0.0010
8192	1	0.3328	0.2862	0.1561	0.1059	0.0617	0.0329	0.0158	0.0075	0.0038	0.0019
16384	1	0.5386	0.4292	0.3294	0.2190	0.1432	0.0704	0.0314	0.0168	0.0085	0.0042
32768	1	0.5367	0.4755	0.4327	0.3346	0.2306	0.1289	0.0599	0.0294	0.0147	0.0072
65536	1	0.6129	0.6153	0.5722	0.5193	0.3613	0.2495	0.1232	0.0594	0.0296	0.0148
131072	1	0.6691	0.6729	0.6960	0.6544	0.5502	0.3862	0.2559	0.1145	0.0584	0.0291
262144	1	0.8597	0.9040	0.9059	1.0510	0.9400	0.7218	0.4984	0.2982	0.1462	0.0715
524288	1	0.6948	0.7995	0.6195	0.9138	0.9026	0.7999	0.6455	0.4285	0.2392	0.1177
1048576	1	0.7483	0.7805	0.6499	1.0495	1.0679	1.0129	0.8236	0.6520	0.4532	0.2459
2097152	1	0.7451	0.7873	0.8238	1.0626	1.1129	1.0956	1.0000	0.8717	0.6930	0.4549
4194304	1	0.6931	0.7406	0.7158	0.9289	1.1641	1.1650	1.0778	1.0205	0.8839	0.7007
8388608	1	0.7836	0.6334	1.0153	1.1554	1.1823	1.1746	1.1430	1.1023	1.0360	0.8951
16777216	1	0.7238	0.5531	0.9471	1.2194	1.2141	1.2181	1.1723	1.1700	1.1320	1.0532
33554432	1	0.7400	0.5738	1.0609	1.1974	1.2523	1.2670	1.1956	1.2112	1.2060	1.1545
67108864	1	0.7248	0.4717	1.0171	1.1829	1.2300	1.2353	1.1943	1.2023	1.2064	1.1773
134217728	1	0.7425	0.5817	1.0398	1.1907	1.2213	1.2309	1.1967	1.2089	1.2241	1.1981
268435456	1	0.7231	0.4982	1.1169	1.1972	1.2484	1.2489	1.2107	1.2118	1.2267	1.2213
536870912	1	0.7319	0.5314	1.0327	1.1826	1.2387	1.2383	1.2108	1.2140	1.2337	1.2321
1073741824	1	0.7335	0.6192	1.1130	1.1840	1.2466	1.2498	1.2191	1.2217	1.2246	1.2327

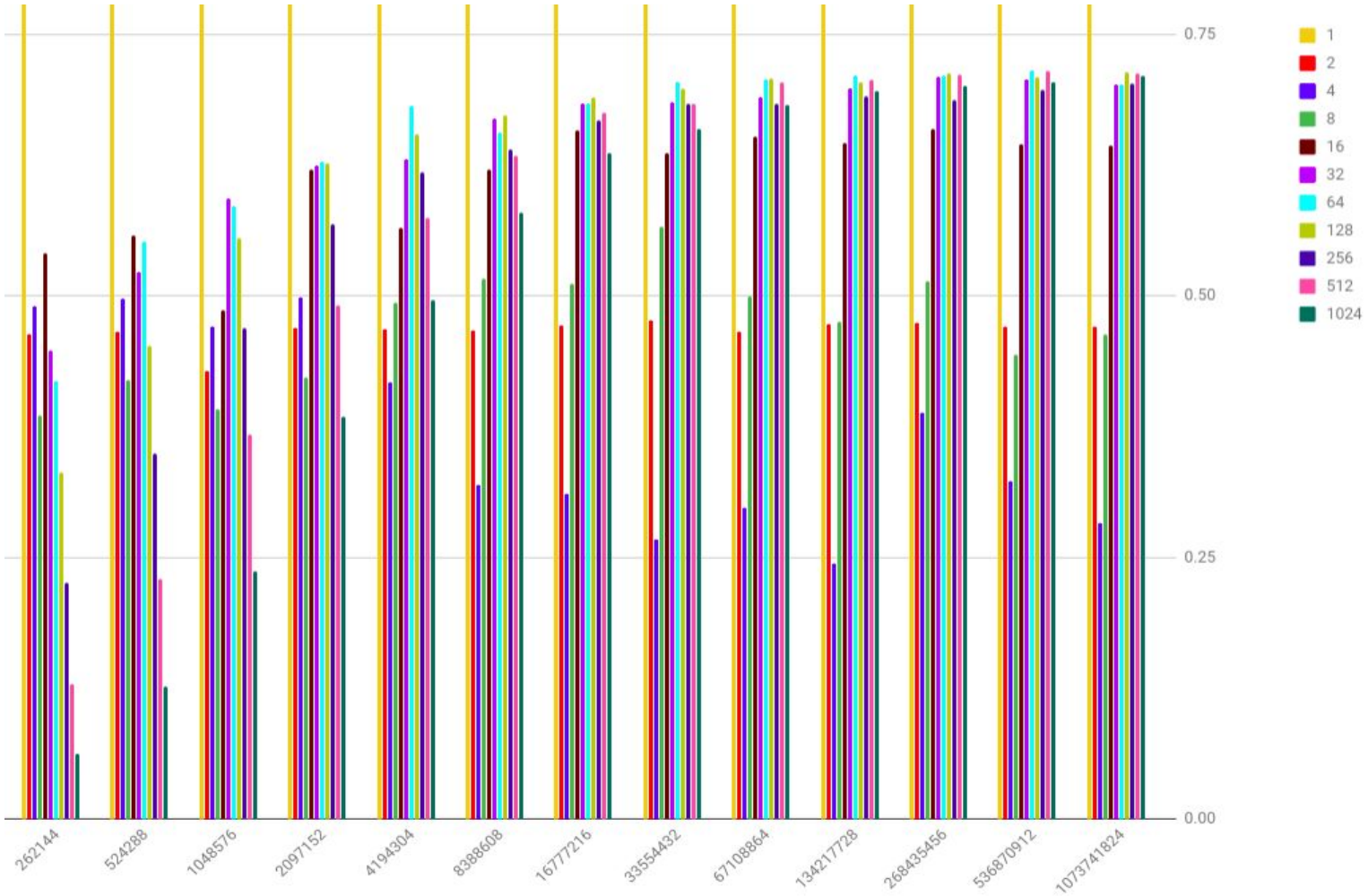

Speed Up | Varying Size of String & No. of Threads
with DFA having 2 states and 2 symbols
Speed Up
1.5
1.0
0.5
0.0
512
1024
2048
4096
8192
16384
32768
65536
131072
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
Size of String (N)
1
2
4
8
16
32
64
128
256
512
1024


1.0
0.5
0.0
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
1
2
4
8
16
32
64
128
256
512
1024

3. Speed Up | Varying Size of String & No. of Threads | with DFA having 4 states and 3 symbols :-

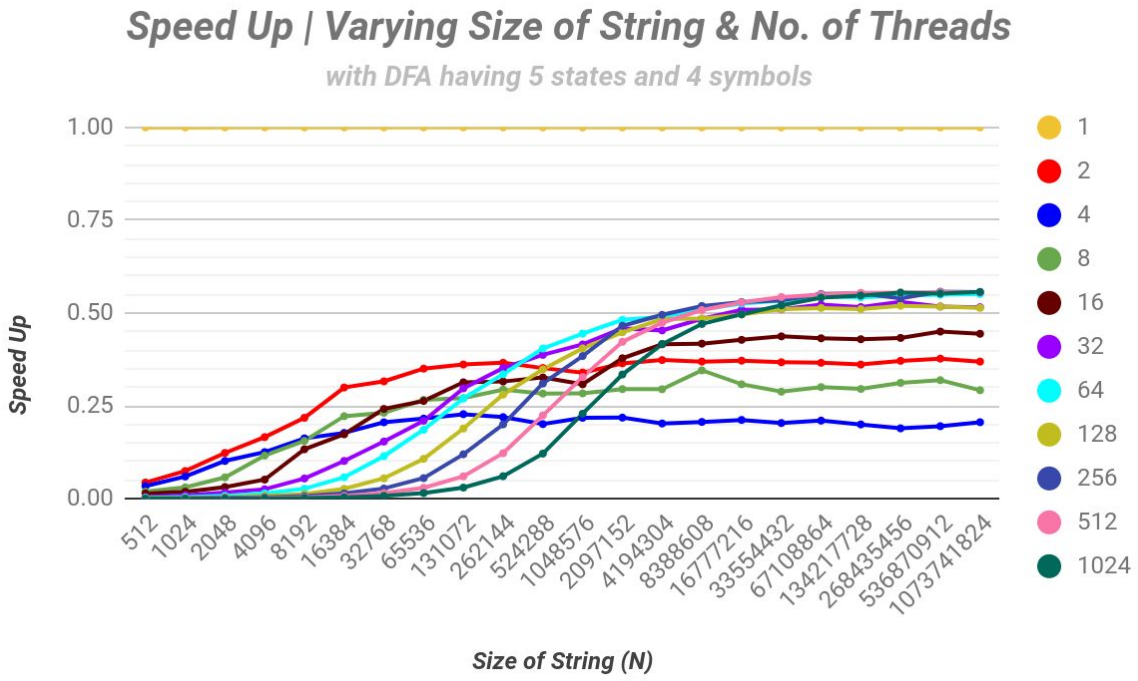
Size of String (N)	No. of Threads (P)										
	1	2	4	8	16	32	64	128	256	512	1024
512	1	0.0392	0.0466	0.0203	0.0130	0.0063	0.0032	0.0015	0.0008	0.0004	0.0002
1024	1	0.0658	0.0621	0.0363	0.0189	0.0107	0.0040	0.0024	0.0012	0.0006	0.0003
2048	1	0.1105	0.1053	0.0603	0.0372	0.0158	0.0084	0.0042	0.0020	0.0010	0.0005
4096	1	0.1804	0.1715	0.0987	0.0665	0.0342	0.0133	0.0074	0.0039	0.0020	0.0010
8192	1	0.2791	0.2394	0.1857	0.1486	0.0669	0.0315	0.0158	0.0078	0.0040	0.0020
16384	1	0.3272	0.2860	0.2547	0.2063	0.1252	0.0628	0.0304	0.0151	0.0076	0.0039
32768	1	0.3966	0.3758	0.3091	0.3140	0.2014	0.1308	0.0598	0.0294	0.0154	0.0077
65536	1	0.4189	0.4181	0.3260	0.4262	0.2874	0.2070	0.1290	0.0597	0.0303	0.0151
131072	1	0.4626	0.4201	0.3847	0.5329	0.3738	0.3076	0.2275	0.1306	0.0595	0.0314
262144	1	0.4631	0.4902	0.3855	0.5405	0.4472	0.4180	0.3315	0.2254	0.1293	0.0625
524288	1	0.4661	0.4969	0.4202	0.5583	0.5222	0.5516	0.4526	0.3492	0.2297	0.1269
1048576	1	0.4287	0.4704	0.3925	0.4866	0.5927	0.5860	0.5554	0.4697	0.3675	0.2371
2097152	1	0.4692	0.4983	0.4221	0.6206	0.6241	0.6281	0.6271	0.5683	0.4910	0.3850
4194304	1	0.4678	0.4175	0.4940	0.5645	0.6307	0.6811	0.6551	0.6180	0.5748	0.4961
8388608	1	0.4670	0.3187	0.5169	0.6205	0.6689	0.6563	0.6723	0.6402	0.6343	0.5798
16777216	1	0.4713	0.3103	0.5124	0.6588	0.6838	0.6843	0.6894	0.6682	0.6753	0.6364
33554432	1	0.4769	0.2670	0.5669	0.6362	0.6846	0.7038	0.6989	0.6835	0.6833	0.6600
67108864	1	0.4656	0.2972	0.5003	0.6526	0.6898	0.7069	0.7077	0.6839	0.7042	0.6820
134217728	1	0.4731	0.2442	0.4757	0.6466	0.6980	0.7101	0.7041	0.6915	0.7070	0.6964
268435456	1	0.4741	0.3883	0.5148	0.6590	0.7093	0.7106	0.7134	0.6877	0.7115	0.7011
536870912	1	0.4710	0.3234	0.4442	0.6446	0.7070	0.7150	0.7096	0.6976	0.7150	0.7043
1073741824	1	0.4711	0.2834	0.4634	0.6439	0.7018	0.7020	0.7146	0.7035	0.7124	0.7101

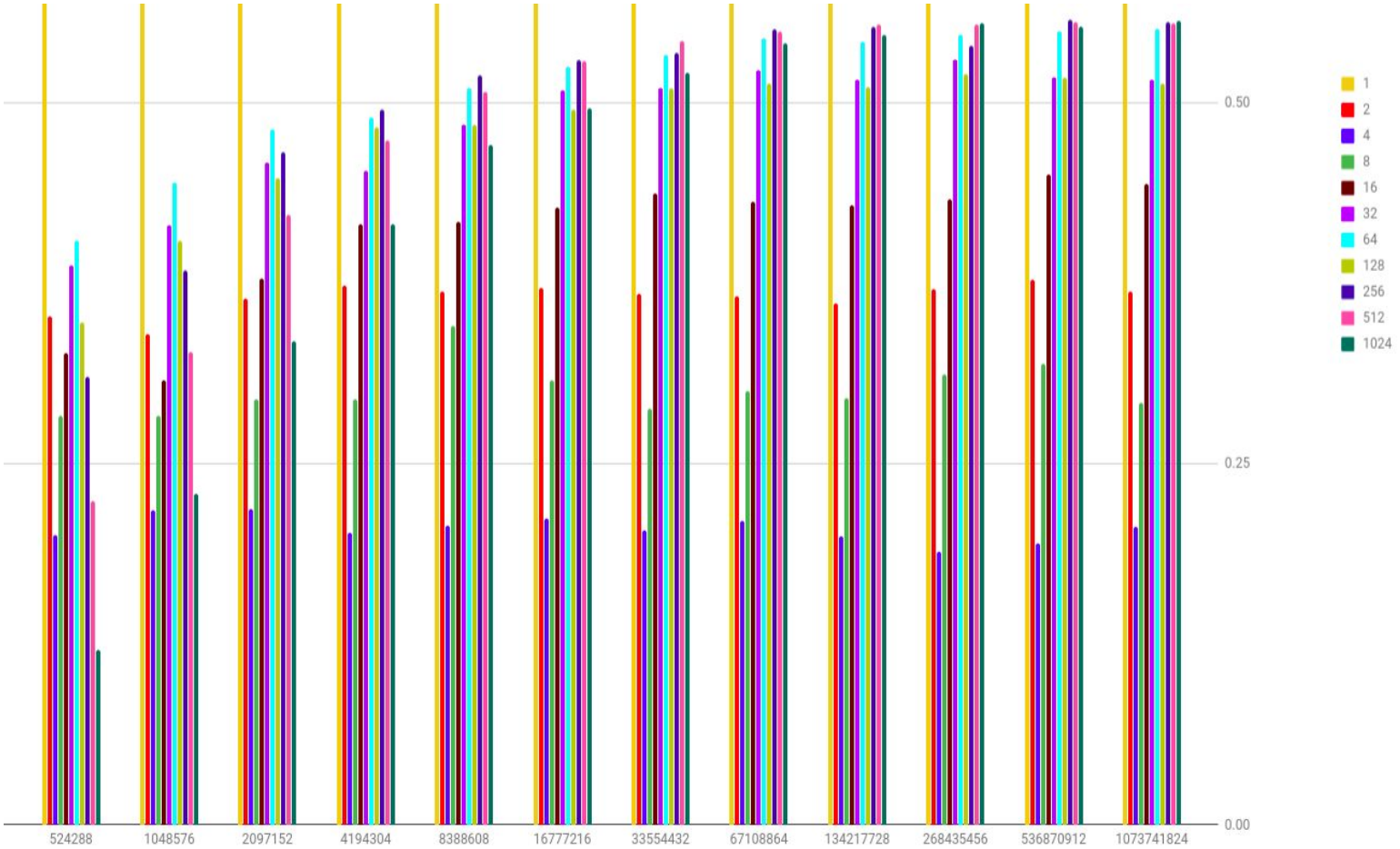

Speed Up | Varying Size of String & No. of Threads
with DFA having 4 states and 3 symbols
Speed Up
1.00
0.75
0.50
0.25
0.00
512
1024
2048
4096
8192
16384
32768
65536
131072
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
Size of String (N)
1
2
4
8
16
32
64
128
256
512
1024


0.75
0.50
0.25
0.00
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
1
2
4
8
16
32
64
128
256
512
1024

4. Speed Up | Varying Size of String & No. of Threads | with DFA having 5 states and 4 symbols :-

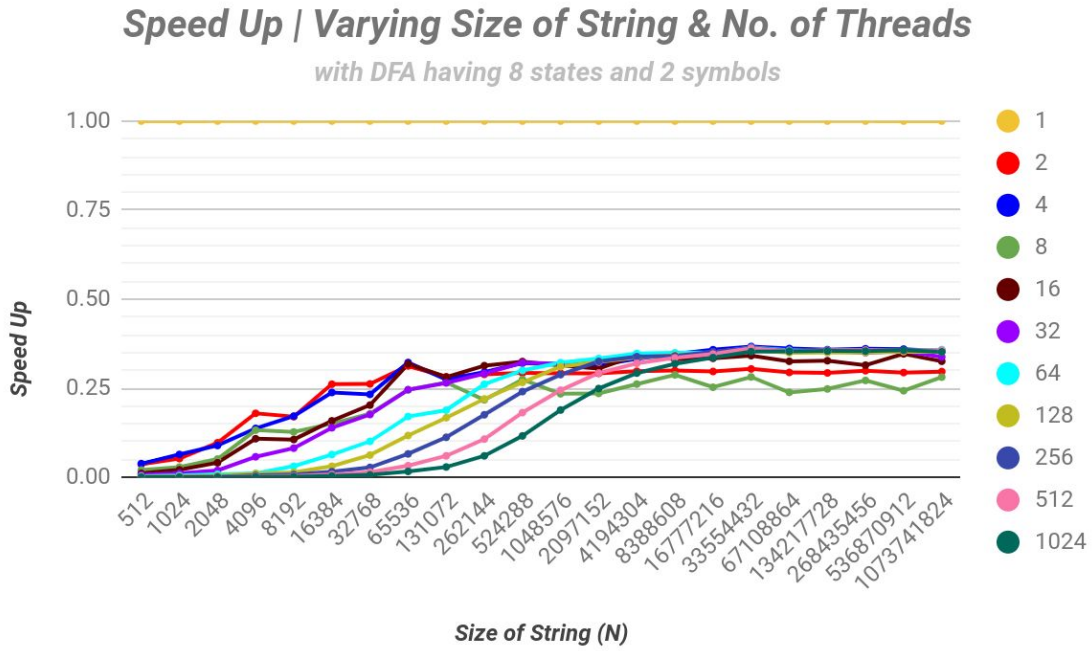
Size of String (N)	No. of Threads (P)										
	1	2	4	8	16	32	64	128	256	512	1024
512	1	0.0436	0.0340	0.0191	0.0140	0.0041	0.0030	0.0016	0.0007	0.0004	0.0002
1024	1	0.0745	0.0599	0.0307	0.0186	0.0089	0.0044	0.0020	0.0011	0.0006	0.0003
2048	1	0.1233	0.1017	0.0574	0.0314	0.0164	0.0078	0.0038	0.0022	0.0012	0.0006
4096	1	0.1659	0.1255	0.1161	0.0514	0.0255	0.0143	0.0075	0.0039	0.0021	0.0010
8192	1	0.2179	0.1625	0.1556	0.1332	0.0544	0.0268	0.0125	0.0068	0.0039	0.0019
16384	1	0.2995	0.1772	0.2222	0.1735	0.1017	0.0576	0.0265	0.0142	0.0076	0.0038
32768	1	0.3160	0.2058	0.2314	0.2417	0.1539	0.1147	0.0548	0.0275	0.0152	0.0076
65536	1	0.3503	0.2158	0.2663	0.2632	0.2099	0.1857	0.1073	0.0554	0.0293	0.0151
131072	1	0.3614	0.2276	0.2709	0.3134	0.2973	0.2700	0.1889	0.1196	0.0602	0.0299
262144	1	0.3660	0.2197	0.2939	0.3156	0.3520	0.3352	0.2804	0.2001	0.1226	0.0606
524288	1	0.3519	0.2009	0.2832	0.3265	0.3873	0.4043	0.3482	0.3102	0.2244	0.1215
1048576	1	0.3398	0.2179	0.2837	0.3080	0.4154	0.4447	0.4049	0.3842	0.3277	0.2289
2097152	1	0.3646	0.2184	0.2952	0.3781	0.4586	0.4818	0.4481	0.4657	0.4225	0.3346
4194304	1	0.3734	0.2026	0.2948	0.4160	0.4533	0.4898	0.4835	0.4951	0.4739	0.4164
8388608	1	0.3693	0.2068	0.3457	0.4176	0.4847	0.5101	0.4850	0.5191	0.5081	0.4706
16777216	1	0.3717	0.2121	0.3082	0.4278	0.5085	0.5253	0.4959	0.5298	0.5290	0.4966
33554432	1	0.3674	0.2038	0.2882	0.4372	0.5100	0.5335	0.5105	0.5345	0.5432	0.5208
67108864	1	0.3660	0.2103	0.3005	0.4320	0.5226	0.5450	0.5136	0.5514	0.5499	0.5417
134217728	1	0.3613	0.2000	0.2955	0.4295	0.5161	0.5424	0.5107	0.5529	0.5545	0.5472
268435456	1	0.3713	0.1895	0.3121	0.4330	0.5300	0.5467	0.5197	0.5396	0.5548	0.5551
536870912	1	0.3772	0.1953	0.3192	0.4501	0.5177	0.5493	0.5174	0.5574	0.5562	0.5526
1073741824	1	0.3693	0.2060	0.2922	0.4443	0.5157	0.5513	0.5139	0.5562	0.5557	0.5568

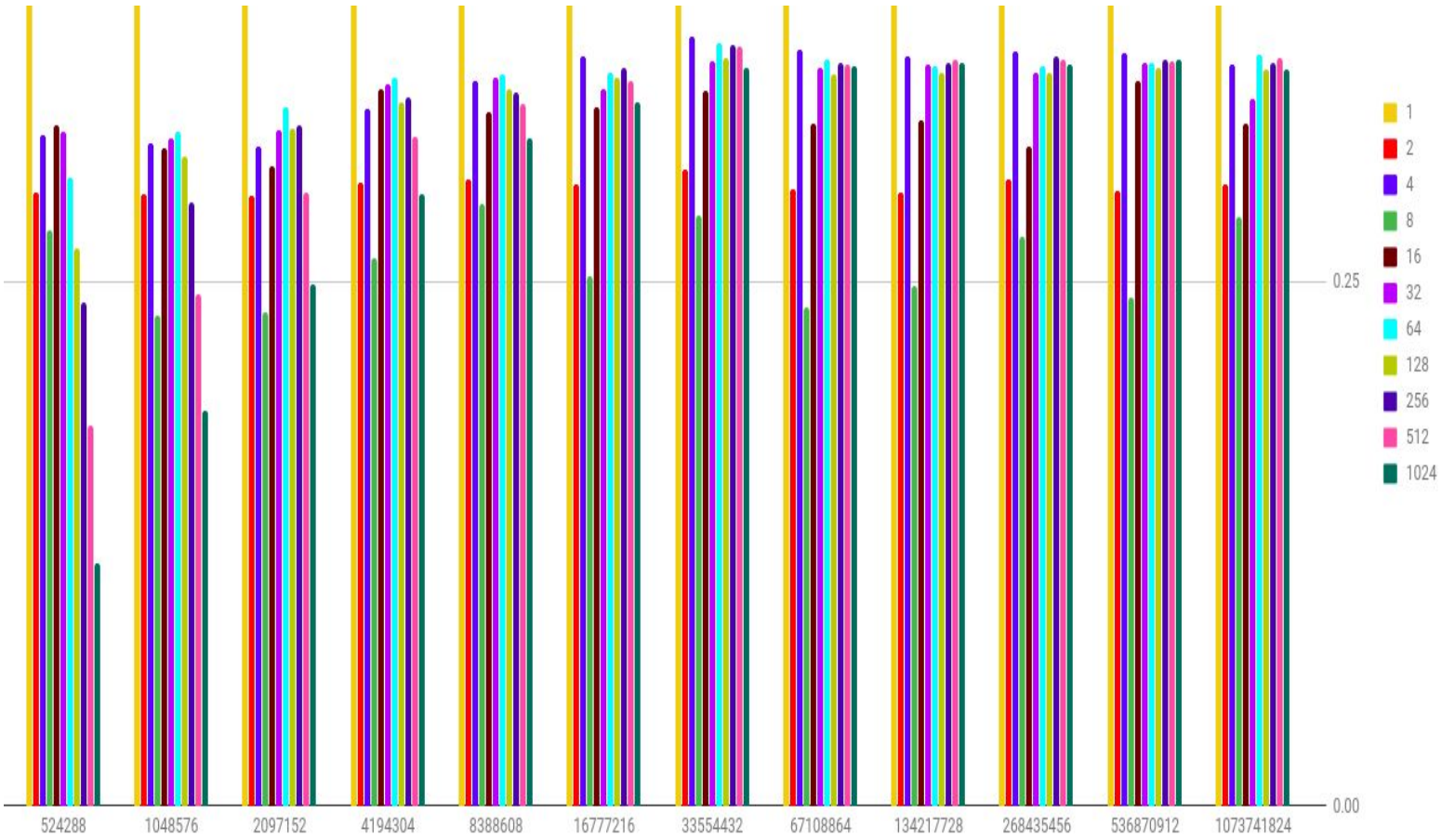

Speed Up | Varying Size of String & No. of Threads
with DFA having 5 states and 4 symbols
Speed Up
1.00
0.75
0.50
0.25
0.00
512
1024
2048
4096
8192
16384
32768
65536
131072
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
Size of String (N)
1
2
4
8
16
32
64
128
256
512
1024


0.50
0.25
0.00
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
1
2
4
8
16
32
64
128
256
512
1024

5. Speed Up | Varying Size of String & No. of Threads | with DFA having 8 states and 2 symbols :-

Size of String (N)	No. of Threads (P)										
	1	2	4	8	16	32	64	128	256	512	1024
512	1	0.0362	0.0384	0.0205	0.0110	0.0072	0.0029	0.0016	0.0007	0.0004	0.0002
1024	1	0.0525	0.0641	0.0288	0.0209	0.0109	0.0049	0.0024	0.0012	0.0006	0.0003
2048	1	0.0970	0.0892	0.0513	0.0408	0.0192	0.0087	0.0044	0.0017	0.0011	0.0005
4096	1	0.1796	0.1373	0.1323	0.1083	0.0576	0.0108	0.0106	0.0055	0.0028	0.0014
8192	1	0.1701	0.1718	0.1273	0.1060	0.0818	0.0310	0.0138	0.0077	0.0037	0.0019
16384	1	0.2613	0.2381	0.1503	0.1585	0.1391	0.0640	0.0310	0.0159	0.0077	0.0039
32768	1	0.2621	0.2328	0.1782	0.2027	0.1762	0.1009	0.0624	0.0278	0.0143	0.0068
65536	1	0.3126	0.3222	0.2455	0.3182	0.2459	0.1707	0.1173	0.0657	0.0327	0.0165
131072	1	0.2823	0.2746	0.2663	0.2810	0.2650	0.1884	0.1673	0.1123	0.0602	0.0289
262144	1	0.2889	0.2952	0.2174	0.3134	0.2910	0.2611	0.2200	0.1755	0.1072	0.0602
524288	1	0.2928	0.3204	0.2746	0.3249	0.3217	0.2999	0.2664	0.2406	0.1817	0.1162
1048576	1	0.2919	0.3163	0.2344	0.3142	0.3187	0.3218	0.3105	0.2882	0.2445	0.1888
2097152	1	0.2917	0.3146	0.2355	0.3056	0.3227	0.3334	0.3239	0.3252	0.2927	0.2491
4194304	1	0.2977	0.3326	0.2617	0.3427	0.3450	0.3476	0.3361	0.3387	0.3198	0.2925
8388608	1	0.2996	0.3461	0.2875	0.3317	0.3475	0.3496	0.3427	0.3407	0.3352	0.3188
16777216	1	0.2970	0.3581	0.2528	0.3340	0.3420	0.3499	0.3479	0.3524	0.3459	0.3361
33554432	1	0.3042	0.3672	0.2817	0.3415	0.3560	0.3644	0.3571	0.3636	0.3624	0.3528
67108864	1	0.2944	0.3615	0.2384	0.3256	0.3524	0.3566	0.3497	0.3547	0.3544	0.3536
134217728	1	0.2930	0.3583	0.2480	0.3273	0.3542	0.3533	0.3501	0.3548	0.3566	0.3551
268435456	1	0.2991	0.3607	0.2719	0.3149	0.3498	0.3536	0.3501	0.3576	0.3568	0.3542
536870912	1	0.2940	0.3599	0.2429	0.3465	0.3545	0.3548	0.3525	0.3564	0.3556	0.3566
1073741824	1	0.2967	0.3543	0.2812	0.3257	0.3379	0.3585	0.3521	0.3552	0.3570	0.3520


Speed Up | Varying Size of String & No. of Threads
with DFA having 8 states and 2 symbols
Speed Up
1.00
0.75
0.50
0.25
0.00
512
1024
2048
4096
8192
16384
32768
65536
131072
262144
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
Size of String (N)
1
2
4
8
16
32
64
128
256
512
1024


0.25
0.00
524288
1048576
2097152
4194304
8388608
16777216
33554432
67108864
134217728
268435456
536870912
1073741824
1
2
4
8
16
32
64
128
256
512
1024

Observations *:-*

1. From the above graphs, it can be easily observed that the *SpeedUp* increases as the size of string (*N*) increases, irrespective of the value of no. of threads (*P*) or no. of states (*Q*).
2. Also, it can be seen that, for every value of thread taken, the *SpeedUp* increases upto a certain max value (till some particular *N*) and then remains nearly the same for further increase in *N*. This particular value of *N*, for which the max value of *SpeedUp* is reached, increases with the value of *P* taken.
3. However, as it can be clearly observed, the max value does not necessarily increase with increase in *P*. The reason behind this may be attributed to the fact that the system on which the experiment is performed has a total of 4 threads, and hence, upon further increasing value of threads, the pseudo-parallelism too comes into picture, thus causing possible delays because of thread-switching.
4. Also, one other main reason observed behind such behaviour is the effect of cache size (because of large size of string) and the possible delay, because of regular cache misses, caused as a result of varied, diverse and extremely fast memory accesses by each and every thread and hence, the time lag and race conditions associated because of it.
5. Thus, in view of the above 3 and 4 points, and observing all the graphs above, one can clearly state that max speed up is observed when no. of threads taken were ***64***. Thus, one should not waste their resources by increasing the no. of threads more than this value since there will be only minimal benefit associated then.
6. Also, it can be deduced that the max speed up reached ***inversely*** varies approximately **linearly** with no. of states used.
   { **1**: ≃2.17, **2**: ≃1.27, **4**: ≃0.72, **5**: ≃0.56, **8**: ≃0.36 }