

### Prediction of Genome Sequences in Terms of Cellular Automata Expansion of Rule Based Logics

#### Rama Naga Kiran Kumar.K, Ramesh Babu. I



Abstract: This paper proposes a novel concept called "Percentage Nucleotide Concentration of genomes" in terms of cellular automata evolutions of adjoints of Adenine, Thymine, Guanine, and Cytosine. The adjoints of the given a genome sequenceare the characteristic binary string sequences. For example, the adjoint of Adenine of a given genome sequence is a binary string consisting of 0's and 1's where 1's corresponds to the presence of Adenine in the genome sequence. So, one can have four adjoint sequences of Adenine, Thymine, Guanine, and Cytosine corresponding to a given genome sequence. Onedimensional three neighborhood binary value cellular automata rules could be applied to an adjoint sequence and the desired number of evolutions obtained. These rules are defined by linear Boolean functions and one can have 256 such linear Boolean functions. The analysis of genome sequences with predictive analytics gives a scope of getting the inherent properties of the genome. The predictive model suits the Nucleotide concentration and is computed for an adjoint sequence and its variation evaluated for its successive evolutions based on a cellular automaton rule.

Keywords: Cellular Automata, Evolutions of Adjoints, Linear Boolean functions, Nucleotide concentration in a genome.

#### I. INTRODUCTION

The purpose of the research carried out and reported in this paper is that whether it is possible to categorize a set of genomes like the human genome repository. The usage of predictive analytics such as recommender systems provides the possible sequence of genome and leads to the evolution of similar models in the processing of genomes. The concept of "%nucleotide concentration" introduced in this paper seems to show a way to accomplish this task. The genesis of the formulation of this concept originates from chemistry, wherein the quantificational notion of percentage ionic concentration of hydrogen (pH value) is used to categorize solutions into three (i) water, whose pH value is 7, (ii) acidic solutions whose pH values are less than 7 and (iii) alkaline solutions whose pH values are greater than 7. On the same lines, an effort was made to categorize genome sets based on four values (i) % nucleotide concentration of Adenine (pA), (ii) % nucleotide concentration of Thymine (pT), (iii) % nucleotide concentration of Guanine (pG) and (iv) % nucleotide concentration of Cytosine (pC). It is reasonable to surmise that these values, possibly their compositions

Revised Manuscript Received on January 30, 2020. \* Correspondence Author

Rama Naga Kiran Kumar. K\*, Research Scholar, Dept of Computer Science & Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India. E-Mail: karasala\_be@hotmail.com

**Prof. Ramesh Babu. I,** Professor, Dept of Computer Science & Engineering, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India.E-Mail: rinampudi@outlook.com.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <a href="http://creativecommons.org/licenses/by-nc-nd/4.0/">http://creativecommons.org/licenses/by-nc-nd/4.0/</a>

would categorize a given set of genomes. The formulation of the concept is briefly explained below. Section 2 of this paper describes the concept formulation. Section 3 of thispaper describes the fundamental notions of adjoints of a genome and their evolution using one dimensional cellular automata rules defined by linear Boolean functions. Section 4 provides experimental results of a case study pertaining to evaluation of Concentration of Nucleotides in terms of Adjoints of Brucella Suis 1330 Genome Sequence.

#### II. CONCEPT FORMULATION

Analogous to the notion of pH value of a solution, the values of pA, pT, pG and pC of a genome sequence and possibly composition of these values like the proportion pA:pT:pG:pC seems to pave a way to classify and characterize genome sets. The definition of "Percentage Nucleotide Concentration" of a genome sequence is given below.

#### **Definition**

Given a genome sequence, the number of a particular nucleotide, say A, present in that genome sequence is counted and the sum is divided by the total number of nucleotides in that genome sequence. The fraction when multiplied by 100 yields the "**Percentage Concentration of Adenine pA**". Similarly, one can evaluate pT, pG and pC.

# III. ONE-DIMENSIONAL THREE NEIGHBORHOOD CELLULAR AUTOMATA EVOLUTIONS OF ADJOINTS OF A GENOME SEQUENCE

Adjoint of a particular nucleotide in a genome sequence is the binary sequence obtained by substituting the particular nucleotides in the genome sequence by 1's and the others by 0's. For example, let us consider a sample sequence of Brucella Suis 1330 for a case study. The actual length of the genome sequence of Brucella Suis 1330 is 5806. A cellular automaton is an idealized parallel processing system consisting of an array of numbers (1-D, 2-D and more) realized using updating rules based on certain neighborhood. For example, a one dimensional cellular automaton would consist of a finite length array as shown below.

Consider an  $i^{th}$  cell in the array. This cell has a neighbor i-1 on its left and another i+1 on its right. All three put together is called a three neighborhood. One can assign a site (cell) variable  $\xi_{i-1}$ ,  $\xi_{i}$ , and  $\xi_{i+1}$  to the three neighborhood cells. At a particular instant of time, these variables take on numerical values, say either a 0 or a 1. In such a case, the variables are denoted as  $\xi_{i-1}^t$ ,  $\xi_{i}^t$ , and  $\xi_{i+1}^t$ . The value of the  $i^{th}$  cell at the next instant of time is evaluated using an updating rule that involves the present values of the  $i^{th}$ ,  $(i-1)^{th}$  and  $(i+1)^{th}$  cells.

This updating rule is basically a linear Boolean function of three variables.



One can construct 256 linear Boolean functions as updating rules of one dimensional three neighborhood binary valued cellular automata. Each rule defines an automaton by itself. So, one dimensional binary valued three neighborhood cellular automata (123CA) rules could be used to model adjoints of a genome sequence. The first twenty linear Boolean functions of cellular automata 123CA are listed below with their decimal equivalents.

<b>Linear Boolean Function</b>	Decimal
	Equivalent
0	0
$(\bar{\xi}_{i-1}\bar{\xi}_i\bar{\xi}_{i+1})$	1
$(\bar{\xi}_{i-1}\bar{\xi}_i\xi_{i+1})$	2
$(\bar{\xi}_{i-1}\bar{\xi}_i)$	3
$(\bar{\xi}_{i-1}\xi_i\bar{\xi}_{i+1})$	4
$(\bar{\xi}_{i-1}\bar{\xi}_{i+1})$	5
$(\bar{\xi}_{i-1}\xi_i\bar{\xi}_{i+1})+(\bar{\xi}_{i-1}\bar{\xi}_i\xi_{i+1})$	6
$(\bar{\xi}_{i-1}\bar{\xi}_{i+1})_+(\bar{\xi}_{i-1}\bar{\xi}_i)$	7
$(\bar{\xi}_{i-1}\xi_i\xi_{i+1})$	8
$(\bar{\xi}_{i-1}\bar{\xi}_i\bar{\xi}_{i+1}) + (\bar{\xi}_{i-1}\xi_i\xi_{i+1})$	9
$(\bar{\xi}_{i-1}\xi_{i+1})$	10
$(\bar{\xi}_{i-1}\bar{\xi}_i) + (\bar{\xi}_{i-1}\xi_{i+1})$	11
$(\bar{\xi}_{i-1}\xi_i)$	12
$(\bar{\xi}_{i-1}\bar{\xi}_{i+1}) + (\bar{\xi}_{i-1}\xi_i)$	13
$(\bar{\xi}_{i-1}\xi_i) + (\bar{\xi}_{i-1}\xi_{i+1})$	14
$(ar{\xi}_{i-1})$	15
$(\xi_{i-1}\bar{\xi}_i\bar{\xi}_{i+1})$	16
$(\bar{\xi}_i\bar{\xi}_{i+1})$	17
$(\xi_{i-1}\bar{\xi}_i\bar{\xi}_{i+1}) + (\bar{\xi}_{i-1}\bar{\xi}_i\xi_{i+1})$	18
$(\bar{\xi}_i\bar{\xi}_{i+1}) + (\bar{\xi}_{i-1}\bar{\xi}_i)$	19
$(\xi_{i-1}\bar{\xi}_i\bar{\xi}_{i+1}) + (\bar{\xi}_{i-1}\xi_i\bar{\xi}_{i+1})$	20

For the case study rule number 90 is applied to the adjoints of Brucella Suis 1330 genome sequence and 500 evolutions generated. Rule 90 is shown below.

$$(\xi_{i-1}\bar{\xi}_{i+1}) + (\bar{\xi}_{i-1}\xi_{i+1})$$
 90

Since the image of the 500 evolutions of Brucella Suis 1330 is large, a small portion of the images are presented in this paper.

## IV. CONCENTRATION OF NUCLEOTIDES IN ADJOINTS OF BRUCELLA SUIS 1330 GENOME SEQUENCE

The values of pA, pT, pG and pC of the Brucella Suis 1330 genome sequence are computed for the adjoints A(n), T(n), G(n) and C(n) and their 500 evolutions using 123CA rules based one linear Boolean functions. Fig. 1 shows the evolutions of the adjoints of A(n), T(n), G(n) and C(n) using the linear Boolean function rule 90 of 123CA. The values are tabulated and the corresponding graphs shown subsequently. Table 1 shows the pA values of A(n) of Brucella Suis 1330 genome sequence and the 500 generations of A(n) using rule 90 of 123CA. Figs. 2 and 3 shows the graphs of the variations of pA values of all generations. Table 2 shows the pT values of T(n) of Brucella Suis 1330 genome sequence and the 500 generations of T(n) using rule 90 of 123CA. Figs. 4and 5 shows the graph of the variations of pT values of all generations. Table 3 shows the pG values of G(n) of Brucella Suis 1330 genome sequence and the 500 generations of G(n) using rule 90 of 123CA. Fig. 4 shows the graph of variations of pA values of all generations. Table 4 shows the pC values of C(n) of Brucella Suis 1330 genome sequence and 500 generations of C(n) using rule 90 of 123CA. Fig. 5 shows the graph of the variations of pC values of all generations.

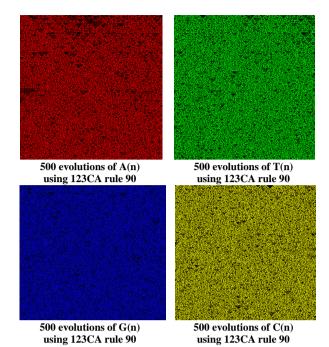


Fig. 1: Evolutions of the adjoints of A(n), T(n). G(n) and C(n).

Table 1: pA values of A(n) and its 500 evolutions

					1					-(,					-					
1	19.18705	21	42.7681	41	42.31829	61	50.08612	81	43.1278		101	48.57379	121	49.03548	141	48.96559	161	43.04168	181	50.32725
2	30.50293	22	49.13882	42	48.93214	62	50.24113	82	48.65656		102	49.52108	122	50.68894	142	50.32725	162	48.72546	182	50.17224
3	30.29625	23	48.67379	43	48.35376	63	52.10127	83	49.3455		103	50.82673	123	50.34447	143	50.36169	163	48.60489	183	49.67275
4	40.80262	24	50.55115	44	50.72339	64	48.7599	84	49.55553		104	50.17224	124	50.56452	144	51.13676	164	50.55115	184	49.46607
5	30.38236	25	42.95556	45	46.60696	65	30.89907	25	49.31106		105	46.77919	125	50.80951	145	42.19773	165	48.50155	185	48.94936
6	41.62935	26	47.96762	46	49.56941	66	41.54323	86	49.63831		106	50.70617	126	51.8257	146	49.63831	166	49.93111	186	49.72442
7	41.52601	27	48.62212	47	50.22391	67	41.78436	87	50.15501		107	49.60386	127	48.77713	147	49.24216	167	50.13779	187	49.74165
8	48.51877	28	50.51671	48	50.51671	68	47.24423	88	49.58998		108	49.58398	128	48.72546	148	49.37828	168	50.55115	188	51.89459
9	31.01963	29	48.69101	49	42.88667	69	41.93937	29	48.31209		109	48.12263	129	31.45022	149	48.8288	169	48.89769	189	50.37892
10	41.02652	30	50.49948	50	48.19153	70	48.39821	90	49.53496		110	50.17224	130	43.0589	150	49.29383	170	51.18843	190	50.13779
11	41.14709	31	49.96555	51	49.39718	71	48.50155	91	50.93007		111	49.98278	131	42.55222	151	50.03445	171	49.29383	191	50.17224
12	48.57044	32	49.39718	52	49.56941	72	49.19049	92	50.15501		112	50,99897	132	49.3455	152	49.50052	172	49.93111	192	50.03445
13	41.49156	33	30.83018	53	48.58767	73	42.30107	93	49.17327		113	45.81364	133	42.57654	153	49.22494	173	50.39614	193	42.15328
14	48.19153	34	41.54323	54	50.27558	74	48.32931	94	50.08612		114	48.96659	134	48.01929	154	50.62005	174	49.44885	194	47.84705
15	47.15811	35	41.6638	55	49,65553	75	49.0527	95	50,20668		115	49.86221	135	49.77609	155	51.18843	175	49.94833	195	48.15708
16	50.06889	36	48.536	56	50.44781	76	50.20668	96	50.39614		116	50.13779	136	49.24216	156	49.87944	176	48.81157	196	50.32725
17	31,34688	37	41,92215	57	49,36273	77	47,53703	97	42,42163		117	48.69101	137	43.16225	157	49.98278	177	49.10437	197	48,05374
18	43.1278	38	46.98588	58	50.17224	78	49.1216	98	48,58767		112	50.6545	138	48.74268	158	49.86221	178	49,29383	192	49.44885
19	43,83396	39	47,48536	59	50,20668	79	48 48433	99	49,00103		119	49.94833	139	48.89769	159	50.72339	179	49.06993	199	49.22494
20			CO DOCAD		10 01100		*******	100	40 00000		120	50.36169	140	50.17224	160	52.06683	180	49.37995	200	49.3455

#### Table 1 continues...

201	48.15708	221	50.01722	241	49.10437	251	43.43782	281	48.94936	301	50.01722	321	42.33551	341	49.44885	361	49.50386	381	51.30899
202	49.51774	222	50.5856	242	50.70617	252	48.19153	282	50.82673	302	50.2928	322	48.7599	342	49.37995	362	49.55219	382	50.12056
203	19.89656	223	50.25835	243	49.39718	253	49.22494	283	49.58996	303	49.15605	323	48.72546	343	49.98278	363	49.27661	383	49.94833
204	50.36169	224	50.60282	244	49.45329	264	48.65656	284	49.79332	304	49.89866	324	49.81054	344	49.86221	364	51.05064	384	49.29383
205	49.25939	225	47.50258	245	51.15398	265	43.04168	285	52.32518	305	48.98381	325	49.29383	345	49.56941	365	19.86221	385	43.19569
206	49.65553	226	50.80951	246	51.30899	256	48.84602	286	49.89666	306	50.51671	326	51.27454	346	50.75784	366	49.63831	386	48.8288
207	50.10334	227	50.79228	247	50.63727	267	48.22597	287	49.91388	307	50.79228	327	49.82776	347	50.67172	367	50.44781	387	48.29487
208	49.51774	228	48.4671	248	50.53393	218	49.7072	288	48.65656	308	50.52005	328	49.24216	348	49.89866	368	49.50386	388	50.9473
209	49.35273	229	50.10334	249	49.39718	269	48.32931	289	42.54554	309	50.27558	329	48.1743	349	50.10334	369	48.50489	389	48.63934
210	48.67379	230	49.74165	250	49.93111	270	49.81054	290	49.27861	310	49.55219	330	49.50052	350	50.91285	370	50.79228	390	50.10334
211	49.68998	231	49.67275	251	49.94833	271	49.84499	291	47.91595	311	49.58998	331	50.53393	351	49.51774	371	48.38099	391	50.32725
212	49.84499	232	50.06889	252	49.63831	272	50.68894	292	49.98278	312	50.56838	332	49.86221	352	50.22391	372	49.29383	392	49.51774
213	50.49948	233	49.27661	253	48.89769	273	42.35274	293	49.27661	313	47.76094	333	49.75887	353	48.57044	373	49.74165	393	48.01929
214	50.75784	234	50.8784	254	49.87944	274	48.86324	294	51.29177	314	49.50052	334	49,89666	354	49.68998	374	50.86118	394	50.25835
215	49.63831	235	50.43059	255	49.1216	275	50.22391	295	51.20565	315	50.43059	335	50.20568	355	50.41837	375	50.22391	395	49.86221
216	50.22391	236	50.55115	256	48.77713	276	50.37892	296	49.22494	316	50.24113	336	50.37892	356	50.70517	376	50.34447	396	50.86118
217	49.79332	237	49.24216	257	30.96796	277	47.8815	297	49.03548	317	48.536	337	48.58767	357	49.58563	377	49.77509	397	50.63727
218	49.19049	238	50	258	41.99104	278	49.56941	298	49.56941	318	49.52108	338	49.01828	358	51.49845	378	49.7072	398	50.51671
219	49.81054	239	50.18946	259	42.52497	279	49.45607	299	49.53831	319	50.49948	339	50.43059	359	48.62212	379	50.25835	399	51.20565
220	51.58457	240	49.82776	260	49.27661	280	50.60282	300	49.81054	320	49.77609	340	50.39514	360	49.22494	380	49.86221	400	49.62108

#### Table 1 continues...

401	48.36376	421	49.67275	441	49.91388	461	50.05167	481	48.55322
402	49.67275	422	49.93111	442	51.56734	462	50.72339	482	50.51671
403	51.22287	423	49.22494	443	50.20668	463	48.89769	483	49.87944
404	50.37292	424	50.44782	444	50.36169	464	50.10224	464	51.05788
405	49.91388	425	49.06993	445	49.86721	465	50.43059	485	50.05167
406	49.37995	426	49.91388	446	50.5856	466	50.06889	485	49.98278
407	49.08715	427	49.58653	447	50.36169	467	50.10334	487	49.79332
408	49.56541	422	50.13779	448	50.20668	468	50.10334	483	49.29383
409	50.68894	429	49.91388	449	48.26042	469	50.18946	489	51.55012
410	49.63831	430	50.10334	450	49.75887	470	50.70617	490	49.48329
411	51.2401	431	49.63998	451	50.05167	471	52.03841	491	49.1216
412	49.29049	432	49.48329	452	50.05889	472	51.42956	492	49.48329
413	50.08612	433	49.89656	453	50.20668	473	49.4144	493	50.89569
414	50.43059	434	50.516/1	454	49.15605	474	49.60386	494	49.91389
415	49.98278	415	49.1216	455	49.10437	475	49.31106	495	50.53393
416	49.75887	436	50.8784	456	48.63934	476	49.81054	496	50.25835
417	49.72442	437	49.13882	457	51.05064	477	48.98381	497	49.44885
418	49.5349E	438	50.13779	453	50.82673	478	49.96555	493	49.00103
419	50.22391	439	49.56941	459	50.80951	479	50.17224	499	49.79332
420	49.51774	460	50.05357	460	49.37995	480	50,9473	500	50.84395

Rule number 90 is applied to A(n) and its 500 generations. It is observed that the pA value becomes minimum at regular intervals of 1, 2, 4, 8, 16, 32, 64, 128 and 256. This indicates a fractal behavior of the evolution. Min(A(n))=30.2965 and Max(A(n))=31.4502. The deviation is 1.15.

The Assessment of the control of the

Fig. 2: pA values of A(n) and of its evolutions





A <sub>e</sub> (n)	pA	31.6
e = 1	30.50293	31.4
e = 2	30.29625	312
e = 4	30.38236	3
e = 8	31.01963	328
e = 16	31.34688	336
e = 32	30.83018	33.4
e = 64	30.89907	322
e = 128	31.45022	30
e = 256	30.96796	235
		236

Fig. 3: Minimum pA values of A(n) and of its evolutions

#### Table 2: pT values of T(n) and its 500 evolutions

1	20.25491	21	43.73062	41	44.09232	61	51.42956	81	43.28281	101	48,4671	121	49.15615	141	48.12263	161	42,4733	131	50.34447
2	30.45126	22	49.22494	42	48.69101	62	49.62108	82	49.05993	102	49.32829	122	49.7072	142	50.51671	162	49.17327	182	50.4133
3	37.15639	23	49.62108	43	48.94935	63	51.53674	83	48.70873	103	49.75887	123	50.89553	143	49.46507	163	48.59101	183	50.8267
4	41.74991	24	49.50052	44	51.10231	64	49.27661	84	50.5856	104	49.24216	124	49.43152	144	50.99597	164	51.1712	184	49.9311
5	31.94971	25	43.38615	45	48.84602	65	31.82914	85	48.65656	105	49.7072	125	51.51567	145	42.97279	165	49.1216	185	49.0354
6	41.68102	26	49.44885	46	49.65553	66	41.83603	86	49.53496	106	50,03445	126	50.82673	145	49.36273	166	49.32776	136	49.68933
7	42.7661	27	48.34654	47	49.58663	67	43.59283	87	50.84395	107	49.93111	127	50.93037	147	49.15505	167	49.89666	187	50.08613
2	48.65656	28	49.83054	42	49.25939	68	47.22701	28	49.74165	108	49,62108	128	50.44781	148	49.25939	168	50,03445	188	49.50053
9	32.3803	29	50.24113	49	43.76507	69	43.52394	89	49.98278	109	49.03548	129	31.82914	149	49.0527	169	48.7599	189	58.1894
10	43.11058	30	50.48226	50	49.75887	70	47.70927	90	49.62108	110	50.08612	130	43.09335	150	49.3455	170	51.10231	190	50.1894
11	43,30003	31	49.48329	52	48.31209	71	48.79435	91	49.46607	111	49.84499	131	43.78229	151	50.91285	171	50.08612	191	49.414
12	48.94936	32	49.24216	52	51.22287	72	50.18946	92	49.86221	112	49.91389	132	48.69101	152	50.9473	172	48.94936	192	50.4478
13	43.36893	33	32.19084	53	48.8268	73	42.93834	93	50.27558	113	50.55115	133	43.16225	153	48,70823	173	50.25835	193	42.5594
14	47.9504	34	41.97382	54	48.81157	74	48.50155	94	49.51774	114	50.75784	134	48.74258	154	50,77506	174	49.4E607	124	49.345
15	47.07199	35	42.62832	55	50.08612	75	49.74165	95	49.00103	115	50	135	48.98381	155	48.81157	175	50.55115	195	48.6393
16	51.27454	36	47.67482	56	49.20777	76	49.37878	96	50.77339	116	50,05167	136	51.05054	155	49.29349	176	49.03548	136	49.82771
17	32.65587	37	43.33448	57	48.62212	72	48.8288	97	43.00723	117	51.20565	137	49.40339	157	49.20772	177	48.55322	197	49.0023
18	42.16328	38	48.69101	58	50.74061	78	50.08612	98	47.02032	118	50.18946	138	48.79435	158	50.55115	178	50.15501	198	50.947
19	42.99001	39	48.69101	59	49.91388	79	50.08612	93	48.79435	119	50,43053	139	48.96881	153	49.86221	179	49,74165	139	49.2360
20	49 55222	40	consace	60	\$1 24244	90	AG 50622	100	40 73C4G	120	49,68999	140	50.15501	160	49.87944	180	50.08612	210	49.79333

#### Table 2 continues...

202	50.41337	222	50.40548	242	50.17224	262	47.8648	282	49.77609	302	49.13882	322	47,69204	142	49.55553	362	50.05167	382	50,25835
203	49.82776	223	48.88047	243	49.48329	263	48.63934	283	49.10437	313	48.95322	323	47,6576	343	90,70617	363	49.46607	383	50.15301
204	49.1216	224	49.17327	244	50.98174	264	50	284	49.75887	304	51.11953	324	48.4571	344	50.15501	364	50.70617	384	51.44578
205	50.27558	225	47.9504	245	49.60386	265	43.64451	285	50.17224	315	49.46607	325	49.19049	345	49.58831	365	49.32828	385	43,47277
206	49.58663	226	49.39718	246	49.56941	265	48.07096	286	50.63727	316	49.29383	326	50.31002	346	49.19049	366	50.01722	386	48.94936
207	50.53393	227	50.31002	247	49.81054	267	49.48329	287	50.93007	397	49.24216	327	48,86324	347	50.27558	367	49.91398	387	49,62108
208	49.24216	228	49.01826	248	50.79228	268	50.2928	288	49.48329	318	50.05157	328	48.8288	348	49,4144	368	49.72442	388	49.44855
209	49.00103	225	50.74051	249	50.8784	263	49.03548	289	43.23114	319	49.87914	329	48.84602	349	50.18945	369	49.65553	389	49.24216
210	49.91388	230	49.20437	250	49.67275	270	49.84499	290	48.81157	310	50.13779	330	50.10334	350	49.82775	370	50.49948	390	50,05157
211	49.82776	231	50.34447	251	50.24113	271	49.1216	291	47.33035	311	50.41337	331	50	951	49.89665	371	50.62005	391	50.68894
212	49.67275	232	49.60386	252	50.72339	272	49.82776	292	50.32725	312	49,72442	332	49,87944	352	50.25835	372	50,77506	392	48,72546
213	50.2928	233	50.34447	253	50.55115	273	43.47227	293	48.39821	313	48.96324	333	50.51671	253	48.99214	372	58.32725	393	49.24499
214	49.25939	234	49.13882	254	50.62005	274	48.62212	294	51.25732	314	49,58683	334	49,51774	354	50.20668	374	49,75887	394	49,77519
215	48.59101	235	49.89656	255	48.77713	275	48.19153	295	50.72339	215	50.17224	325	30.13779	355	49.94833	375	A0.70017	395	50.32725
216	50.6545	236	49.63831	256	49.31106	276	49.17327	296	50.82673	316	51,20565	336	50,6545	156	49,57175	376	50.89563	396	50.72339
217	50.15501	237	50.43059	257	33.06924	277	48,7599	297	48.57044	317	48.95381	337	43,03548	157	53	377	49.24216	397	49,22494
218	50.43059	238	49.22494	258	41.92215	278	50.10334	298	49.20772	313	49,1216	338	50,75784	358	49,94833	378	50.10334	398	50,9472
219	49.13192	239	50.05267	259	44.17844	279	50.08445	299	49.43162	319	50.63727	339	50,85118	359	19,55553	379	51.27454	399	50.25835
220	50.06889	240	50.06889	260	48.63934	280	49.22194	300	49.20772							1			

#### Table 2 continues...

401	49.37995	421	50.27558	441	50.51671	461	49.55219	481	48.98383
402	50	422	49.36273	442	50.20568	462	50.20668	482	49.7588
403	49.84199	423	49.62108	443	51.01619	463	50.27558	483	50.15500
404	48,79135	424	49.36273	444	49.44885	464	49.74165	484	50.17224
403	50.24113	425	50,67172	445	51.06786	465	50.08612	485	49.86223
406	49,46607	426	49.44885	446	49.43167	466	49.00103	486	49.0699
407	50.25885	427	50.43059	447	50.G3727	457	49.98278	487	48.6737
408	50.75784	428	50.34447	448	50.71339	468	49,44895	488	49.6210
409	50.43059	429	50.34447	449	48.05374	459	49.61054	489	49.8622
410	51.7568	430	50.55115	450	52.37685	470	51,464	490	50.3616
411	50.68894	431	51.03341	451	49.29383	471	50.55838	491	49.8622
412	48,94936	432	49.7072	452	49.20772	472	48.96659	492	49.6383
413	50.68894	433	49.29383	453	49.27661	473	49,4144	493	49.9655
414	49.89666	434	49.20772	454	49.74165	474	50.8784	494	50.5511
415	49.51774	435	49.87944	455	48.4671	475	50.36169	495	49.5886
416	51.8257	436	50.45225	456	50.12056	476	49.58663	496	51.3951
417	47.55425	437	50.37892	457	50.2928	477	50.49948	497	50.3200
418	49.94833	438	48.27764	458	49.1216	478	49.91388	498	49.68390
419	48.27517	429	49.9185N	459	57.45797	479	49.63831	499	41.7244
420	48,62212	440	49.93111	450	50.72339	480	49.08715	500	49,6555

Rule number 90 is applied to T(n) and its 500 generations. It is observed that the pT value becomes minimum at regular intervals of 1, 2, 4, 8, 16, 32, 64, 128 and 256. This indicates a fractal behavior of the evolution. Min(A(n))=30.45126;

Max(A(n))=33.06924.The deviation is 2.61.

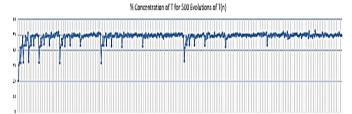


Fig. 4: pT values of T(n) and of its evolutions



Fig. 5: Minimum pT values of T(n) and of its evolutions

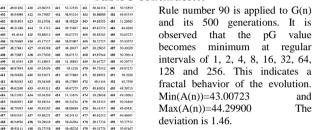
#### Table 3: pG values of G(n) and its 500 evolutions

2	43.00723	22	49.24216	42	51.1712	62	50:24113	82	50.44781	102	49.82776	122	50.53393	142	50.6545	162	50.03445	182	48.91492
3	43.97175	23	49.75187	43	49.96555	63	49.86221	83	49.89666	103	49,56941	123	51.34244	143	49.82776	163	49.1216	183	50.18946
4	49.22494	24	50.74061	44	49.08715	64	50.68894	84	49.87944	104	49.65553	124	50.41337	144	49.15605	164	49.68998	184	50.24113
5	43.86841	25	49.20772	45	49.27661	65	43.74785	85	49.3455	105	50.24113	125	50.22391	145	50.03445	165	49.25939	185	51.18843
6	49,00103	26	49,37995	45	49.87944	66	50.05167	86	50.43059	106	49.96555	126	50.06889	146	50.82673	166	50.01722	186	48.81157
7	48.50155	27	50.18946	47	49.91388	67	49.72442	87	49.0527	107	49.93111	127	50.25835	147	51,37788	167	49,46607	187	50.56838
8	50.22391	28	49.36273	48	49.65998	68	50.01722	88	50.36169	108	49.01826	128	49.39718	148	50.34447	168	48.96659	188	49.82776
9	43.7134	29	50.36169	49	48.31209	69	48.89769	89	50.34447	109	49.65553	129	43.57561	149	50.46504	169	49.32828	189	50.05167
10	48.10541	30	50.06189	50	49.86221	70	49.94833	50	49.82776	110	51.18843	130	49.20772	150	50.39614	170	50.03445	190	50.91285
11	48.2432	31	48.10541	51	49.51774	71	50.48226	91	49.89666	111	49.86221	131	49.15605	151	50.01722	171	49.24216	191	49.65553
12	50,77506	32	49.22494	52	50.60282	72	51.08508	92	48.88047	112	48.98381	132	59.6545	152	48.51877	172	50	192	50.84393
13	49.36273	33	42,835	53	49.50052	73	49.00103	93	50.13779	113	50,44781	133	47,03755	153	50.17224	173	48,70323	193	49.84499
14	49.82776	34	48,86324	54	49.95555	74	50.60282	94	5L36068	114	49.91388	134	50.06589	154	50.12056	174	50.65394	194	49.86221
15	49.7072	35	50.20668	55	49.05993	75	51.03341	95	49.44885	115	50.74061	135	50.8784	155	49.94833	175	50.86118	195	50.03445
16	49,44885	36	51.25732	56	49.27661	76	49.63831	96	50.43059	116	49.60386	136	49.29383	156	50.15501	176	51.10231	196	49.84499
17	43,75952	37	48.74268	57	49.91388	27	50	97	48.70823	117	49,03548	137	49.01826	157	50	177	49,84499	197	49.43162
18	48.41543	38	50.12856	58	50.44781	78	49.81054	98	49.19049	118	50,48226	138	49.25929	158	50.10334	178	49.55219	198	49.19049
19	48.20875	39	50.48126	59	50.65894	79	51.05064	99	49.93111	119	50,08612	139	50.32725	159	50.39614	179	50.10334	199	50.18946

#### Table 3 continues...

202	50,70617	222	50.41337	242	50,67172	262	49.98278	282	51.01619	302	49.96555	322	50.91285	342	50.84395	362	49.87944	382	49.60386
203	49,84499	223	49.89666	243	50.6545	263	50.13779	283	50.70617	303	50.18946	323	49.56941	343	49.91318	363	50.2928	383	49.62108
2014	49,60386	224	49.53111	244	49.39718	264	49.84499	284	48.89769	304	49.31106	324	50.80951	344	50.08612	364	48.88047	384	50.60282
205	50,12016	225	50.37892	245	49.67275	265	42.3455	285	50.49948	305	49.24216	325	50.55115	345	50,46504	365	49.65553	385	50.20668
206	50,37892	226	50.63727	246	50	266	50.84395	286	49.79332	306	51.05064	326	49.65553	346	49.87944	366	49.89666	386	49.48329
287	51,05064	227	49.0527	247	50.56838	267	49.65553	287	49.15685	307	49.20772	327	50.49948	347	49.8527	367	50.51671	387	48.96659
208	49,44885	228	49,72442	248	50.24113	268	49.61938	288	30.56838	308	49.06993	328	49.86221	348	49.79332	368	49.68998	388	51.06786
2119	50.51671	229	49.37995	249	51.2401	269	50.44781	289	49.36273	309	49.84499	329	51.49845	349	49,44815	369	90.6345	389	48.62212
210	50.68194	238	50.24113	250	49.62108	278	50.39614	290	51.03341	310	30,55115	330	50.43039	330	58.13946	370	49.84499	390	50.70617
211	51.20565	231	49.03548	251	49,48329	271	50.15501	291	50.155EL	341	50.44781	331	49.27661	351	50.34447	371	49.17327	391	49.44885
212	49.86221	232	49.51774	252	49.75887	272	10.60282	292	49.65553	312	50.68194	232	50.55115	352	49.91318	372	58,77506	392	50.43059
213	50.08612	233	50.6545	253	50.86118	273	50,49948	293	50.9473	313	50.51671	333	49,44885	353	50.01722	373	49.79332	393	50.55115
214	50.25835	234	50	254	49.36273	274	49.15605	294	49.48329	344	50.06889	334	51.22287	354	49.37995	374	49.82776	394	50.03445
215	49.00103	235	49.89666	255	47.9504	275	50.75784	295	49.82776	315	49.84499	335	49.75887	355	50	375	50.8784	395	50.46504
216	49.087L5	236	50.98174	256	50.03445	276	50.05389	296	49.89666	316	50.01722	336	49.03548	356	49.84499	376	49.63831	396	48.89769
217	49.63831	237	49.32828	257	44.299	277	50.82673	297	48.65656	317	49.01826	337	49.87944	357	49.03548	377	49.65553	397	50.44781
218	50.10334	238	50.10234	258	49.39718	278	50.2928	298	50.20668	318	50.68894	338	50.15501	358	51.32621	378	51.01619	398	49.77609
219	49.22494	239	50.08612	259	48.89769	279	50.39614	299	49.72442	319	50.13779	339	49.48329	359	51.27454	319	48.58767	399	50.37892
220	49.74165	240	50.91285	269	50.31002	250	49 1216	300	51.76513	320	49.89666	340	50.12056	360	49.93111	380	50.9473	400	49.72442

#### Table 3 continues...





and

Fig. 6: pG values of G(n) and of its evolutions

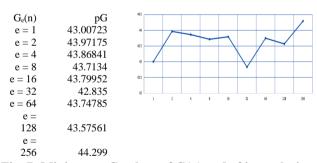


Fig. 7: Minimum pG values of G(n) and of its evolutions

#### Table 4: pC values of C(n) and its 500 evolutions



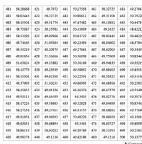
Table 4 continues...



#### Prediction of Genome Sequences in Terms of Cellular Automata Expansion of Rule Based Logics

20 L	49.39718	221	49.3455	241	48.88047	261	47.82983	281	48.86324	301	49.23494	321	48.62212	341	49.27661	361	49.98278	381	50.56838
202	49.84499	122	51.25732	242	50.93007	262	50.67172	282	49.31106	302	49.56941	322	50.05167	342	49.56941	362	49.65553	382	48:01402
203	49.7072	223	50.06889	243	50.80951	263	49.72442	283	50.9473	303	50.20668	323	50.67172	343	50.77506	363	49.96555	383	49.96555
204	49.50052	224	48.81157	244	50.20568	264	50.43059	284	51.1712	304	50,46504	324	50,98174	344	50.44781	364	49,67275	384	49,55219
205	50.68894	225	50.27558	245	51.67069	265	48.36376	285	49.82776	305	48,7500	325	49.86221	345	49.56941	365	59.46504	385	47.10644
206	50,44781	226	49.08715	246	50,70517	266	51.2401	286	51.11953	306	49.72142	326	50.24113	346	50,5856	366	49.48329	336	50.01722
207	49.00103	227	49.7072	247	50.8784	267	50.36169	287	49.68998	307	49.86221	327	50,9473	347	50.86118	367	49.53496	387	49.81054
208	48.93214	228	50.10334	248	50.18946	268	49.48329	288	50.22391	308	51.15398	328	49.82776	348	50.13779	368	48.98381	388	51.05064
289	43.91492	229	50.80951	249	50.79228	269	50.2928	289	46.53806	309	49.00103	329	50.24113	349	49.08715	369	59.01722	389	49.36273
210	49.86221	230	50.36169	250	48,77713	270	50.25835	290	50.01722	310	49.37995	330	49,96555	350	50.17224	370	49.44885	390	49,79332
211	50.20568	231	49.7072	251	50,70617	271	49.53496	291	50.80051	311	49.96555	331	50,32725	351	51.28565	371	49.31106	391	50.67172
212	50.6545	232	49.56941	252	49.81054	272	50.39614	292	51.68791	312	49.87944	332	50.9473	352	50.25835	372	49.84499	392	49.74165
213	49.53496	233	49.62108	253	48,58767	273	47.76094	293	48.31209	313	50.13779	333	50.6545	353	50.5856	373	50.37892	393	49,4144
214	49.72442	234	49.15605	254	50	274	49.7072	294	50.03445	314	49.29383	334	50.36169	354	48.79435	374	49.65553	394	20.99897
215	49.43162	235	50.53393	255	50.46504	275	49.68598	295	50.44781	315	49.86221	335	49.31106	355	50.18946	375	49,7072	395	50.75784
216	49.55219	236	50	256	50.18946	276	49.51774	196	49.98278	316	49.63831	336	50.6545	356	50.15501	376	51.87737	396	50
217	50.03445	237	49.65553	257	40.95763	277	49.98178	197	49.46607	317	49.94833	337	48,39821	357	50.03445	377	50.49948	397	49.96555
218	49.51774	238	49.53496	258	48.36376	278	49.72442	298	49.87944	318	51.5329	338	49.65998	358	50.01722	378	49.37995	398	51.37788
219	50,75784	239	50.56838	259	48.31209	279	50.82673	299	49.31106	319	50.17224	339	50.08612	359	49.10137	379	50.80951	399	50.63727
220	49.98278	2:10	49.94833	260	49.32528	280	49.74165	300	50.6545	320	50,74061	340	50.06889	368	50.91285	380	50.25835	400	49.17327

#### Table 4 continues...



Rule number 90 is applied to C(n) and its 500 generations. It is observed that the pC value becomes minimum at regular intervals of 1, 2, 4, 8, 16, 32, 64, 128 and 256. This indicates a fractal behavior of the evolution.Min(A(n))=40.71650 and Max(A(n))=40.95763. The deviation is 1.99



Fig. 8: pC values of C(n) and of its evolutions

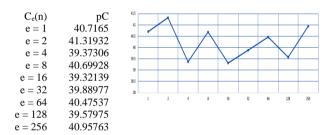


Fig. 9: Minimum pC values of C(n) and of its evolutions

#### V. CONCLUSIONS

This paper proposes a novel concept called "Percentage Nucleotide Concentration of genomes" in terms of cellular automata evolutions of adjoints of Adenine, Thymine, Guanine, and Cytosine. The research carried out and reported in this paper exhibits the possibility to categorize a set of genomes like the human genome repository the predictive method allows us to guess the pattern of sequences which helps the formulation of genome sequences. In short, the concept of "Percentage Nucleotide Concentration (PNC)" introduced in this paper seems to show a way to accomplish this task.

#### REFERENCES

- Carey, Nessa (2015). Junk DNA: A Journey Through the Dark Matter of the Genome. Columbia University Press. ISBN 9780231170840.
- Chris M. Rands, Stephen Meader, Chris P. Ponting and Gerton Lunter (2014). "8.2% of the Human Genome Is Constrained: Variation in Rates of Turnover across Functional Element Classes in the Human Lineage". PLoS Genet 10 (7): e1004525. doi:10.1371 /journal.pgen.1004525. PMC 4109858. PMID 25057982.
- Costa, Fabrico (2012). "7 Non-coding RNAs, Epigenomics, and Complexity in Human Cells". In Morris, Kevin V. Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection. Caister Academic Press. ISBN 1904455948.
- Dan Graur, Yichen Zheng, Nicholas Price, Ricardo B. R. Azevedo1, Rebecca A. Zufall and Eran Elhaik (2013). "On the immortality of television sets: "function" in the human genome according to the evolution-free gospel of ENCODE" (PDF). Genome Biology and Evolution5 (3): 578–90. doi:10.1093/gbe/evt028. PMC 3622293. PMID 23431001.

- Doolittle, W. Ford (2013). "Is junk DNA bunk? A critique of ENCODE". Proc Natl Acad Sci USA110 (14): 5294–5300. Bibcode:2013PNAS..110.5294D. doi:10.1073/pnas.1221376110. PMC 3619371. PMID 23479647.
- Elgar G, Vavouri T; Vavouri (July 2008). "Tuning in to the signals: noncoding sequence conservation in vertebrate genomes". Trends Genet. 24 (7): 344–52. doi:10.1016/j.tig.2008.04.005. PMID 18514361.
- Ehret CF, De Haller G; De Haller (1963). "Origin, development, and maturation of organelles and organelle systems of the cell surface in Paramecium". Journal of Ultrastructure Research. 9 Supplement 1: 1, 3–42. doi:10.1016/S0022-5320(63)80088-X. PMID 14073743.
- Gregory TR, Hebert PD; Hebert (April 1999). "The modulation of DNA content: proximate causes and ultimate consequences". Genome Res. 9 (4): 317–24. doi:10.1101/gr.9.4.317 (inactive 2015-02-01). PMID 10207154.
- Hawkins, JS.; Kim, H.; Nason, JD.; Wing, RA.; Wendel, JF. (Oct 2006). "Differential lineage-specific amplification of transposable elements is responsible for genome size variation in Gossypium". Genome Res 16 (10): 1252–61. doi:10.1101/gr.5282906. PMC 1581434. PMID 16954538.
- International Human Genome Sequencing Consortium (February 2001). "Initial sequencing and analysis of the human genome". Nature 409 (6822): 879–888.
- Bibcode:2001Natur.409..860L. doi:10.1038/35057062.
   PMID 11237011.

#### **AUTHORS PROFILE**



**K Rama Naga Kiran Kumar** completed his B.E(CSE), M.Tech(CSE) and pursuing Ph.D. (CSE) from Acharya Nagarjuna University. His area of interest is in AI, Data Mining, Machine Learning, Software Engineering.



**Dr. Ramesh Babu I,** done his B.E (ECE), M.E (Computer Engg), Ph.D. (CSE) in ANU. His area of interest is in Image Processing, Computer Graphics, AI, Data Mining, Machine Learning, Software Engineering

