Measuring the Uncomputable
Entropy Analysis of Automatic Sequences Revisited

Dr. Kostas Karamanos

Kolmogorov-Chaitin Complexity I

- Descriptive Complexity in 1D
- Turing machines as caricatures of Computation
- Algorithmic compressibility issues
- Encompassing noetic and linguistic processes
- Connections with DNA, music, natural languages
- Connections with Symbolic Dynamics
- No generalization in Higher Dimensions
- Uncomputable in the general case even in 1D

Kolmogorov-Chaitin Complexity II Technical Aspects

- Periodic Strings: K(n)~logn
- Automatic Strings: K(n)~logn
- Random Strings: K(n)~n
- Sporadic Strings (conjectural): K(n)~intermediate
- Normal Sequences (almost all !)
- Champernownes' artificial decimal normal number: 0.12356789101111213...
- Pi: Borwein-Bailey-Plouffe (BBP) numbers
- Pi: Is Pi 2-automatic ? (M. Waldschmidt)

Symbolic Dynamical Systems Before Newton

Substitutive sequences (real time generable sequences)

Periodic strings

Fibonacci word: 0->01, 1->0 (sturmian sequences)

Chacon sequence: 0->0010, 1->1
Weakly mixing but not Strongly mixing

Automatic sequences
(Computation of individual digits possible)

Cryptoautomatic sequences

Possible Quantization ?

Combinatorial Characterization of Automatic Sequences by Alan Cobham

- A. Cobham, Math. Syst. Theory **6,** 164 **(1972)**
- Set of constant length m substitution rules and a final letter-to-letter projection = equivalent to a finite automaton with m-states

Finite automata can generate only fractions or transcendental numbers

- **On the Complexity of Algebraic Numbers I. Expansions in Integer Bases**
- Boris Adamczewski and Yann Bugeaud
- *Annals of Mathematics*
- Second Series, Vol. 165, No. 2 (Mar., 2007), pp. 547-565

Topics I

- Fatou and Julia Theorem

(G.Julia, Journal de Mathematiques, Liouville, 4, 1918)

- Metropolis –Stein-Stein (MSS) Algorithm

(Journal of Combinatorial Theory A 15, 1973)

- Grassberger: Symbolic dynamics at Feigenbaum point

(International Journal of Theoretical Physics, 25, 1986)

- Ebeling and G. Nicolis: Entropy Analysis at the Feigenbaum point

(Chaos Solitons and Fractals, 2, 1992)

Topics II

- Review on the Entropy analysis of symbolic sequences

(G. Nicolis and P. Gaspard, Chaos Solitons and Fractals, 4, 1994)

- Universal Symbolic dynamics at Feigenbaum point: rigorous results

(K. Karamanos and G. Nicolis, Chaos Solitons and Fractals, 10, 1999)

- Automaticity at the Feigenbaum point

(J.P. Allouche and M. Cosnard, CRAS, 296, 1983)

- Structure of the correlation function at the Feigenbaum point at the Logistic map

Metropolis –Stein-Stein (MSS) Algorithm

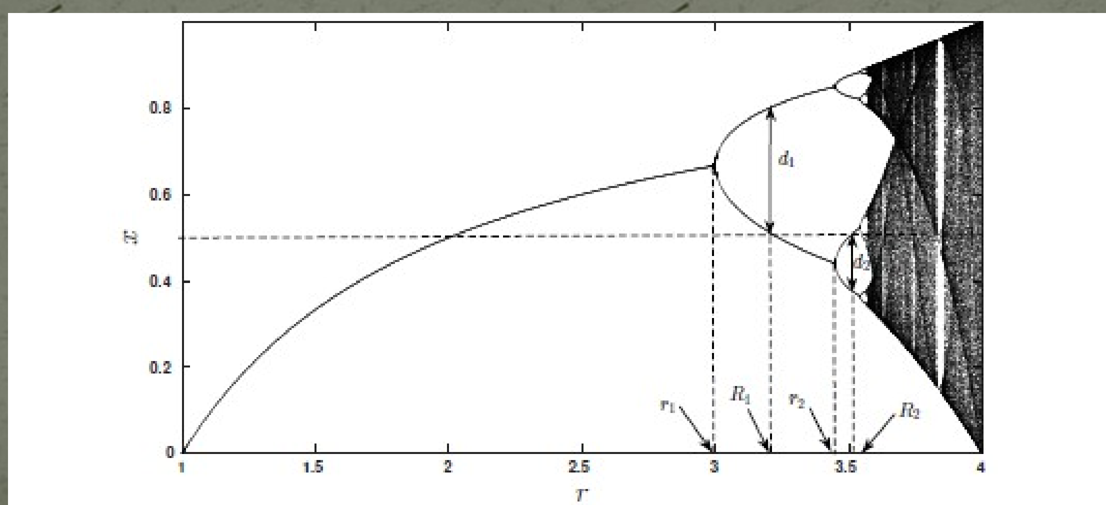
- Unimodal map: f_λ: [0,1] → [0,1]

(a) **λ**=control parameter , real number

(b) $\boldsymbol{f_\lambda}$ is continuous and piecewise differentiable in [0,1]

(c) $\boldsymbol{f_\lambda}$ is convex and has a unique maximum at "c

Logistic map: $x_{n+1}=\lambda\ x_n.(1-\ x_n)$



"**Minimum distinguish information**" in a sequence of iterates x_n.

Pattern: x_n<c -> L

x_n>c -> R

x_0=c, if we start from c and return to c -> **superstable orbit**

Example: -superstable symbolic orbit period 2

cRcRcR... -> cR (periodic) Pattern: P2=R

-superstable symbolic orbit period 3

cRLcRLcRL... -> cRL (periodic) Pattern: P3=RL

Universal ordering, U-sequence (Metropolis)

- Harmonic operator: $\widehat{\boldsymbol{H}}(\mathbf{P}) = \boldsymbol{P\mu P}$

$\mu = \mathrm{L}$, if P contains an odd number of R′s

$\mu = \mathrm{R}$, if P contains an even number of R′s

- List of the first few harmonics associated with the $(2)^k$ cycles:

P_2=R **(period 2)**

$\widehat{\boldsymbol{H}}(P_2) = \widehat{\boldsymbol{H}}(\boldsymbol{R})$ =**RLR**=P_4 **(period 4)**

$\widehat{\boldsymbol{H}}^2(P_2) = \widehat{\boldsymbol{H}}(P_4) = \widehat{\boldsymbol{H}}(\boldsymbol{RLR})$ =**RLRRRLR**=P_4 **(period 4)**

...

$\widehat{\boldsymbol{H}}^\infty(P_2) = \widehat{\boldsymbol{H}}^\infty(\boldsymbol{R})$ = **RLRRRLR LRLRRRLR...**

(accumulation point)

3. $(2)^k$cycles : P_3=RL **(period 3)**

$\widehat{\boldsymbol{H}}(P_3) = \widehat{\boldsymbol{H}}(\boldsymbol{RL})$ =**RLLRL**=P_6 **(period 6)**

...

Self-similariy and lumping

(K. Karamanos and G. Nicolis, Chaos solitons and fractals, 10, 1999)

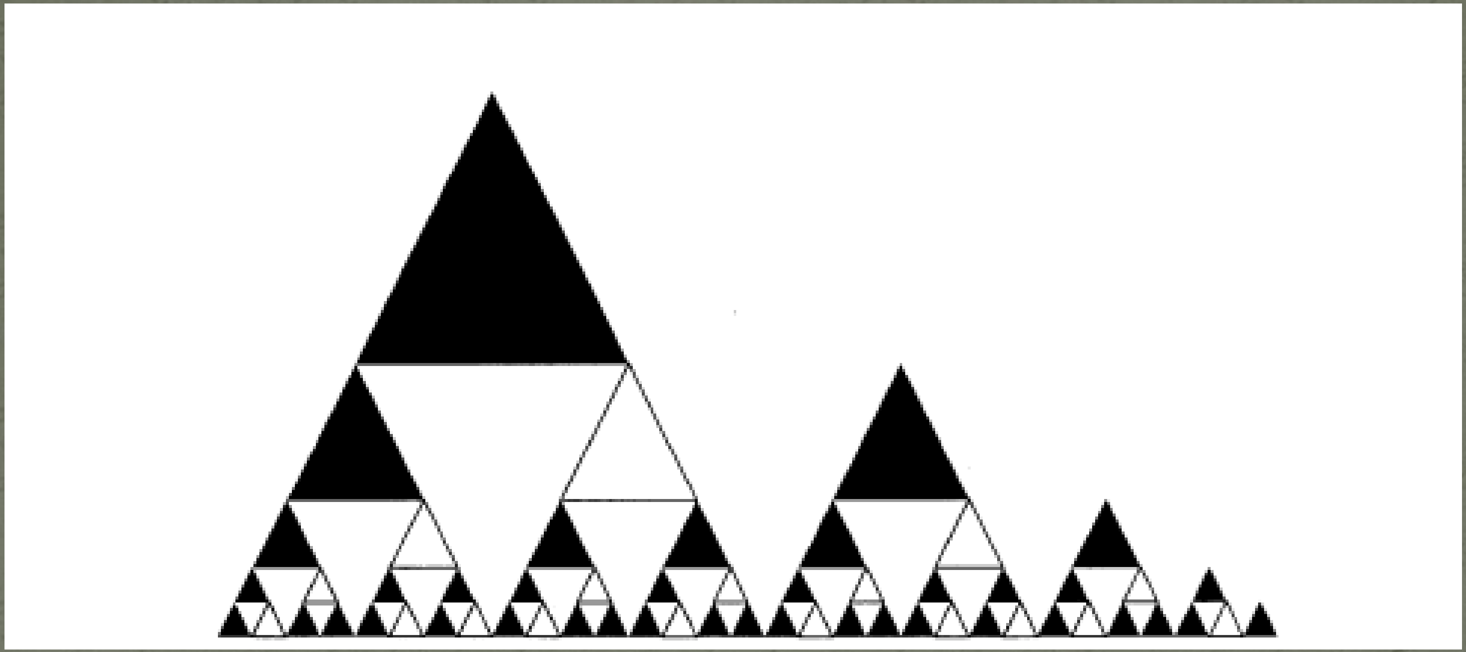
- Operator $\widehat{\boldsymbol{K}}$ in the space of the $(2)^k$ sequences
- i) cut the last R of the pattern P
- ii) in the remaining part of the pattern perform the lumpings: RR → L RL →R

- Choosing $\boldsymbol{P}=\widehat{\boldsymbol{H}}^{\boldsymbol{m}+\mathbf{1}}(\boldsymbol{R})$ one can show that:

$$\widehat{\boldsymbol{K}}\left(\widehat{\boldsymbol{H}}^{\boldsymbol{m}+\mathbf{1}}(\boldsymbol{R})\right)=\widehat{\boldsymbol{H}}^{\boldsymbol{m}}(\boldsymbol{R})$$

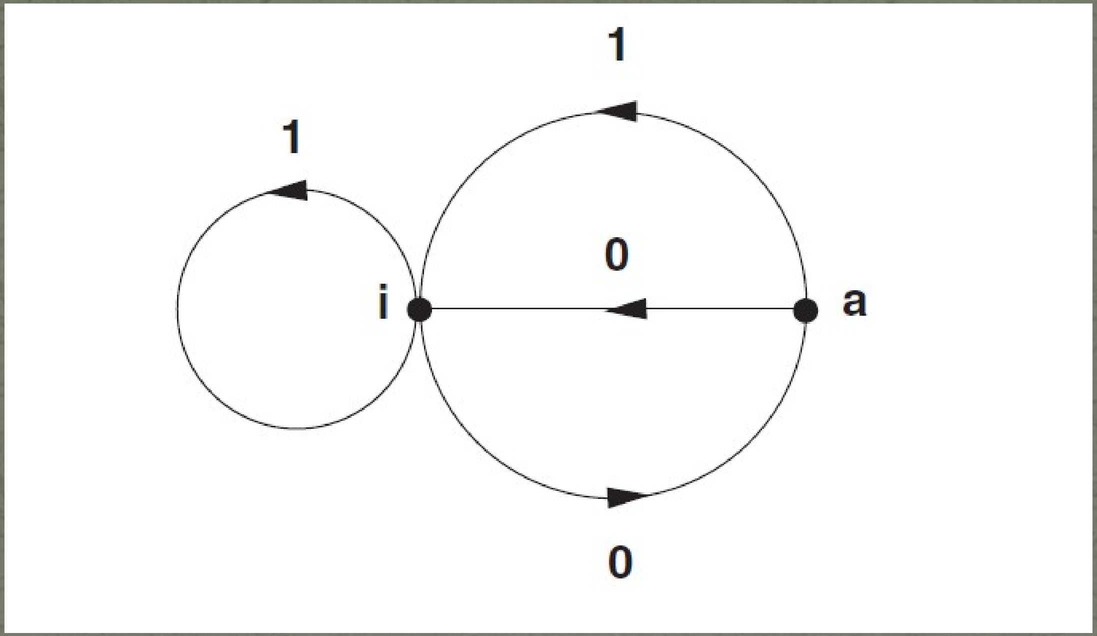
Symbolic Dynamics and lumping

Taking the limit m->∞ we have: $\hat{K}\left(\hat{H}^{\infty}(R)\right) = \hat{H}^{\infty}(R)$



RR → L RL →R

Automaticity at the Feigenbaum point
(J.P. Allouche and M. Cosnard, CRAS, 296, 1983)



$F(i) = R$ $F(a) = L$

Possible applications of these ideas I

Connection between entropy analysis with finite automata

- Boltzmann Entropy
- Shannon Entropy
- Block-Entropy by gliding
- Influence of the way of reading
- Block-Entropy by lumping – Invariance Property
- H(m) = H(m*m) for a finite automaton with m-s
- Necessary Condition – Automaton Reconstructi
- Breaking of Crypto-automaticity

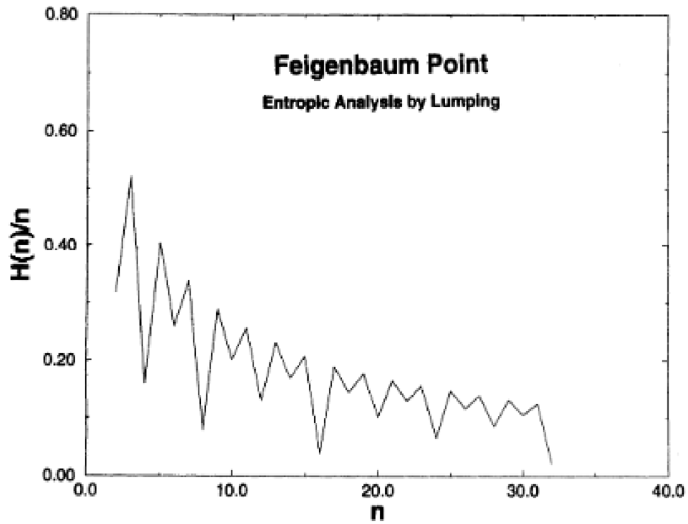


Fig. 1. Entropy per letter, $h^{(n)}$, as a function of n, obtained numericaly by 10^4 iterations of the logistic map, $x_{n+1} = rx_n(1-x_n)$, at the 2^∞ point, $r = 3.56994567$. We observe a non-monotonic decay—a natural consequence of the invariance properties (equation (27)equation (28) and equation (29)). We also notice that the envelope tends to zero as n increases. This is in agreement with equation (5), since the attractors at the Feigenbaum points have zero Lyapounov exponent and, hence, zero Kolmogorov–Sinai entropy.

Possible applications of these ideas II

Connection between entropy analysis with finite automata

Automaticity means “algorithmic compressibity”

Towards a formal definition of “meaning”

Most of the sequences are not automatic

The violation of the invariance property is a clear Signal of non-automaticity

Numerical treatment – Massive production of new theorems

Plausibility arguments in the sense of J.M. Borwein

Experimental and Constructive Mathematics

Pi and the Feigenbaum constants are not 2-automatic

Possible applications of these ideas III

Connection between entropy analysis with finite automata

Other ideas: Beyond Yes/No Answers
Departure from automaticity to develop Automaticity me
Unpublished Material

Developing tools for Practical cases

Cantorian Stochastic Automata, new decimation scheme
H(m^k) = k. H(m)
Unpublished Material

DNA Coding/Non-coding Regions are Cantorian Sets
Conjecture by Provata and Almirantis
First Numerical Verification, Phys. Rev. E (2010)

TOPICS III:
Entropy and Automaticity

K. Karamanos, Lect. Not. Phys. **550,** 357 (2000)
M. Planat (Ed.), CNRS Thematic School
on Number Theory and Physics

K. Karamanos, AIP Conf. Proc. **573,** 278 (2001)
D. Dubois (Ed.), CASYS 2000, CHAOS Inst. Liege

K. Karamanos, J. Phys. A 34, 9231 (2001)

K. Karamanos, Kybernetes 38, 1025 (2009)

Topics IV: Applications

Number Theory

J.M. Borwein and K. Karamanos, Nonconvex Optim. Appl. **77,** 3 (2005)

K. Karamanos and I. Kotsireas, J. Franklin Inst. **342,** 329 (2005)

K. Karamanos and I. Kotsireas, J. Franklin Inst. 343, 759 (2006)

Ergodic Theory

K. Karamanos and I. Kotsireas, Kybernetes **31,** 1409 (2002)

Topics V: Applications

Large Scale Entropy Computations (TSD)
K. Karamanos and I. Kotsireas, AIP Conf. Proc. **718,** 385 (2004)
D. Dubois (Ed.), CASYS'2003, CHAOS Inst., Liege

Noisy systems ?

Chacon sequences ?

DNA Structure
Works with I. Kotsireas and Y. Almirantis
Special thanks to A. Provata for discussions about the DNA structure

Entropy Analysis – DNA Sequences

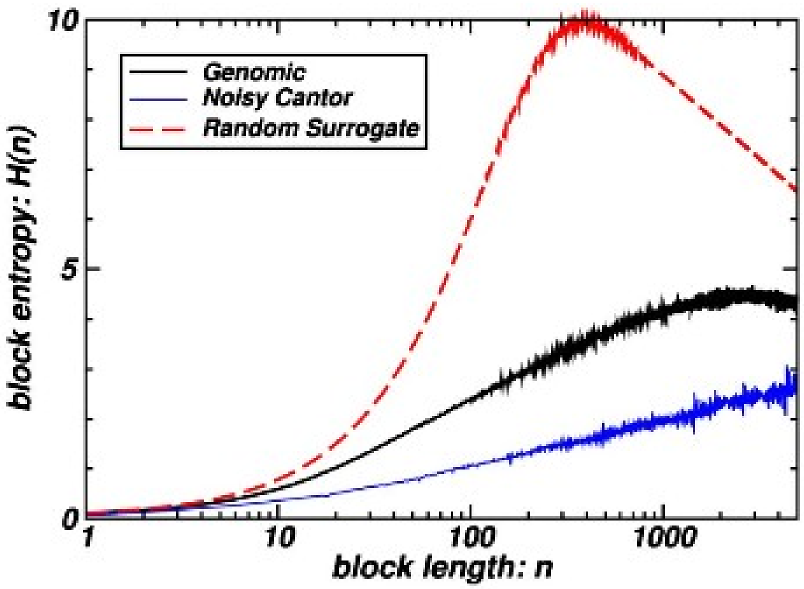
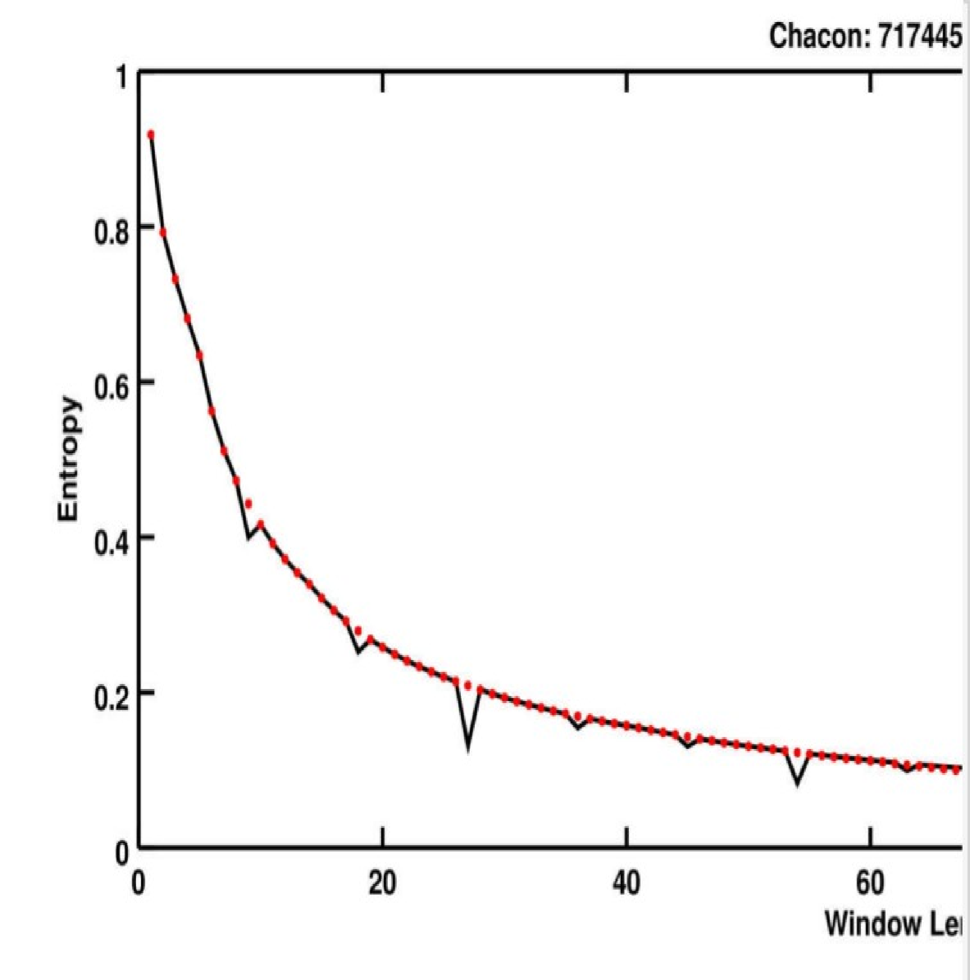


FIG. 3. (Color online) Block entropy $H(n)$ is plotted in semi-logarithmic scale as a function of the word length n for human chromosome 21 with s.f.=100 (see in the text), alongside with a deterministic noisy Cantor-like sequence, with 1% indel occurrences, of (almost) equal length and number of "coding segments." Also, a common random surrogate is included.

7-2

Entropy Analysis – Chacon's sequence



We built a Computable Complexity Theory in the sense of Engineering

Thank you