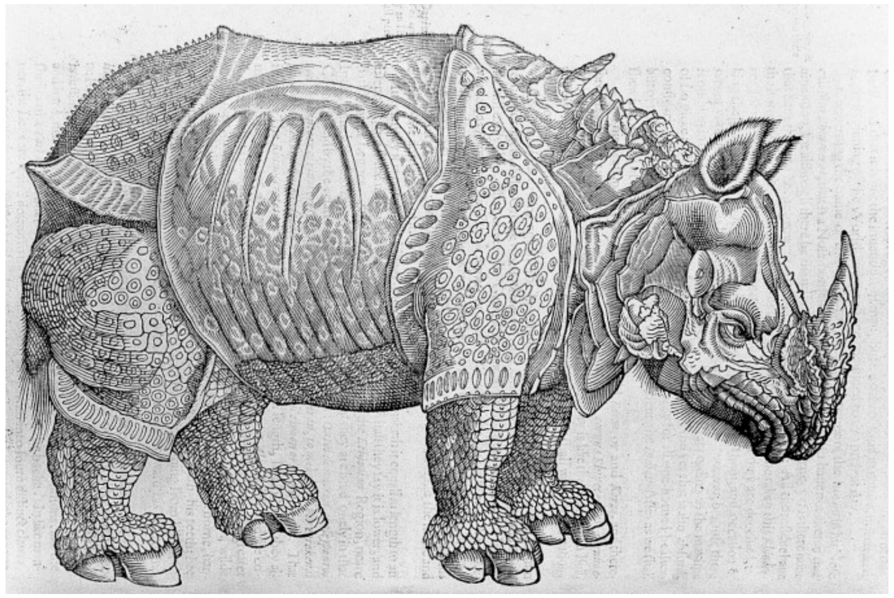
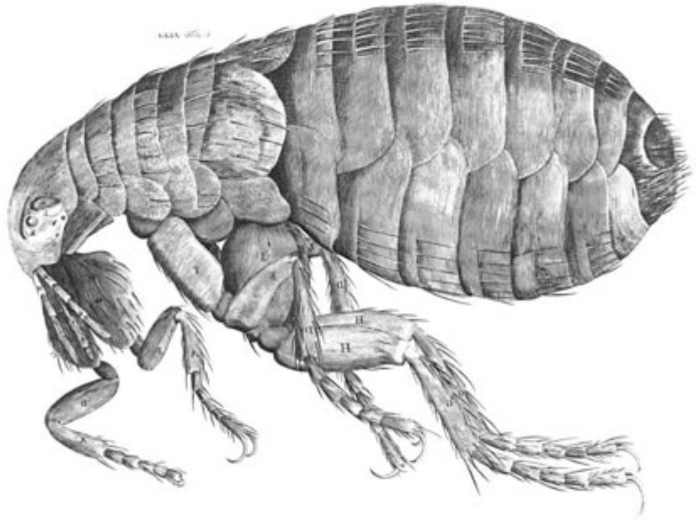
Evolutionary Dynamics
& its Tendencies





David Krakauer, Santa Fe Institute.

A Talk in 2 Parts

- *Part 1: What is Evolution, What has it generated & What are its limits?*
- *Part II: The Evolutionary dynamics of Minimal Forms - microbes*

Evolutionary questions?

- Why is there life-like dynamics on earth?
- Why are organisms so diverse?
- Why are organisms so complex?
- What is the relationship of genetic information to learned information?
- When does cultural evolution outpace genetic evolution?
- Is "self-awareness" inevitable?

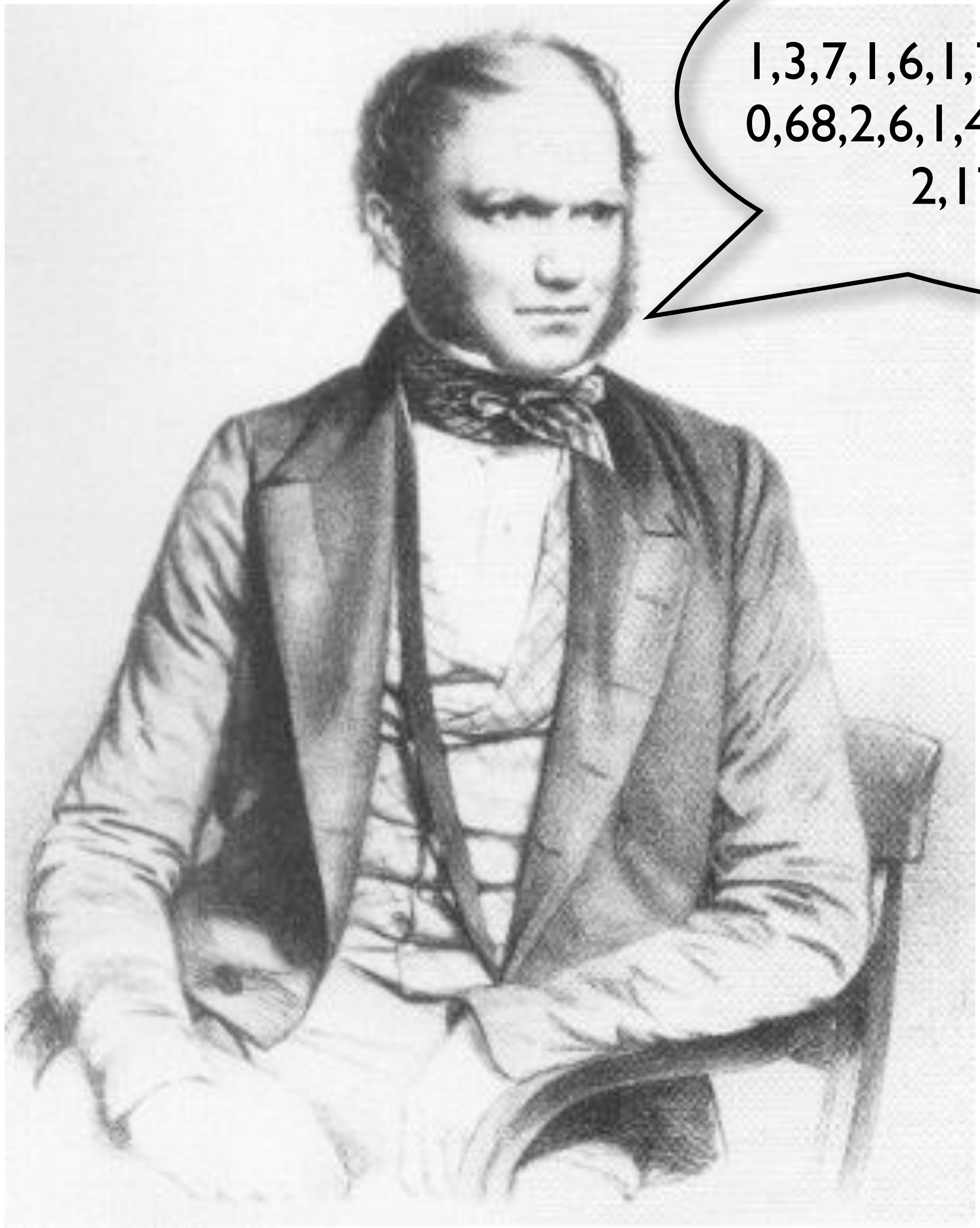

ASTOUNDING SUPER-SCIENCE FANTASIES!
STRANGE WORLDS
10c
No.7
space-gods of planetoid 50
sabotage on space station 1
the man who fought the world


BENJAMIN BRATT
ERIC McCORMACK
CHRISTA MILLER
DANIEL DAE KIM
VIOLA DAVIS
RICKY SCHRODER
ANDRE BRAUGHER
RIDLEY SCOTT AND TONY SCOTT PRESENT
THE ANDROMEDA STRAIN
It's a bad day to be human.
TWO-NIGHT TELEVISION EVENT
A&E
PREMIERES MEMORIAL DAY 9/8C

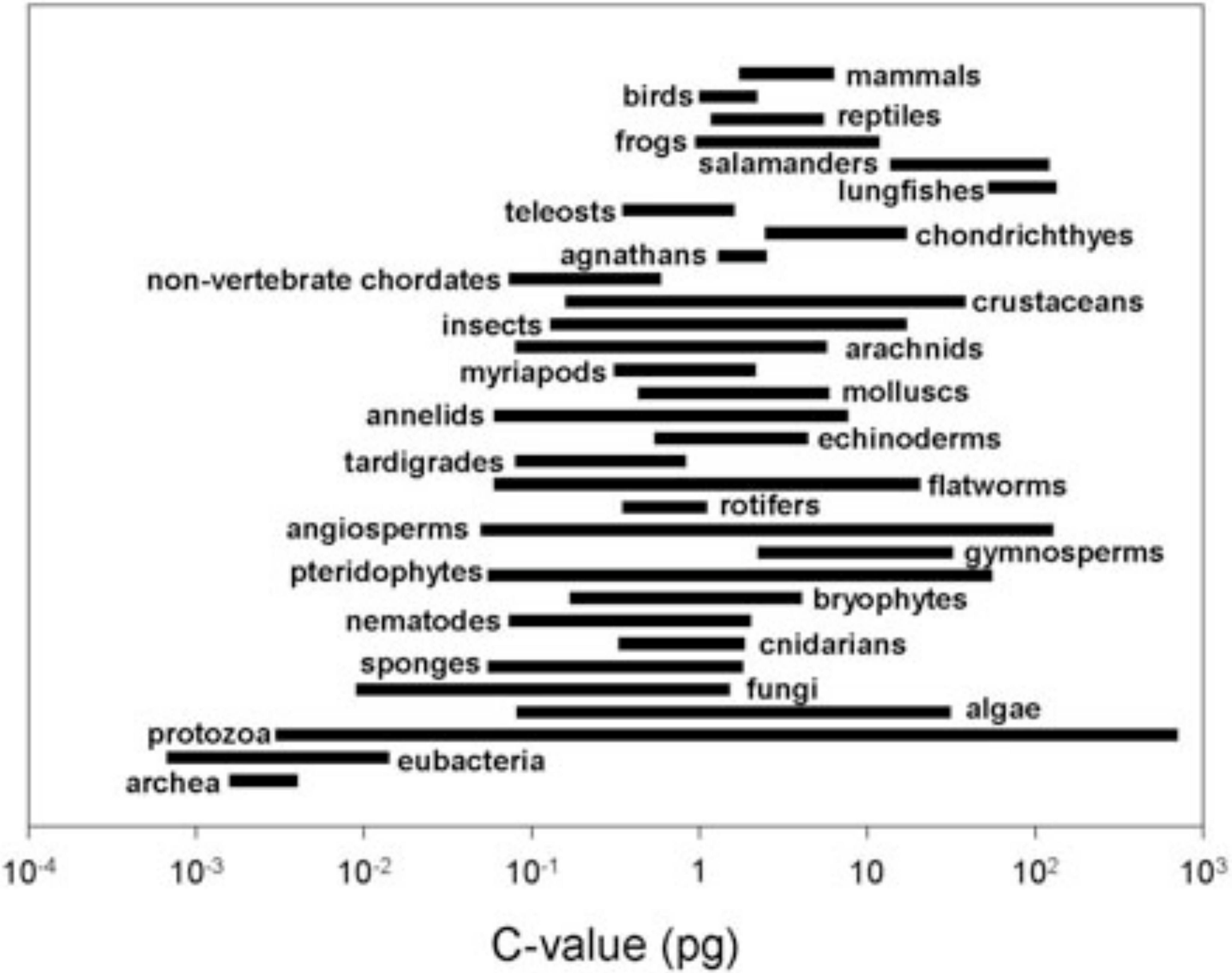
What is Evolutionary Theory?

- A Physics like theory searching for Laws?
- A Statistical/Inferential Theory like Bayesian learning or approximate dynamic programming?
- An algorithmic/computational theory?

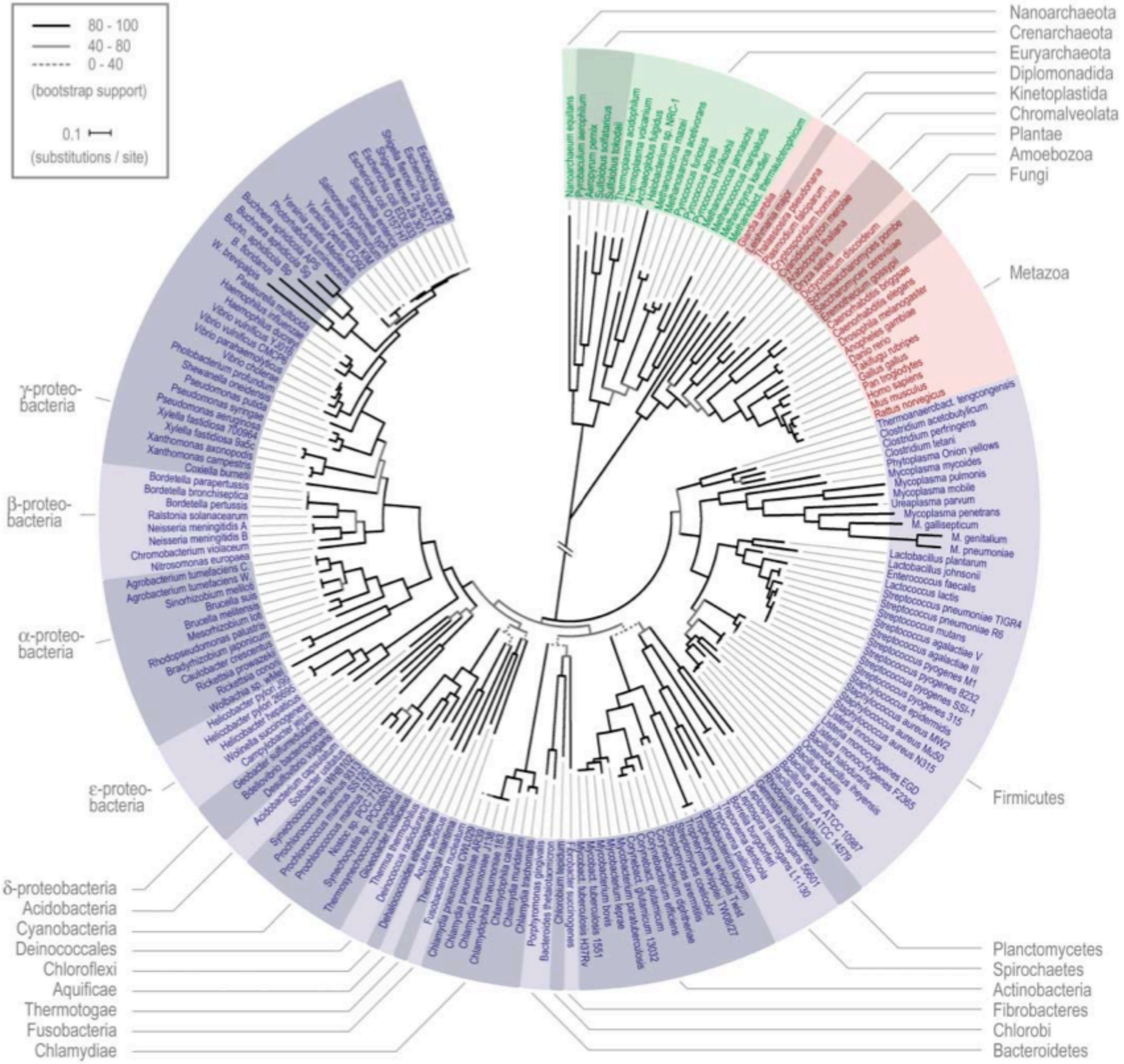
The Darwinian Polynucleotide Machine



1,3,7,1,6,1,7,1,1,5,5,4,3,15,7,1,10,1,15,8,19,10,68,2,6,1,4,5,3,1,6,9,2,4,3,1,2,6,5,2,6,20,1,3,2,176,2,1,1,2 (in MegaBases)



http://www.genomesize.com/statistics.php



Ciccarelli et al. Science, 311, 1283-1287 2006

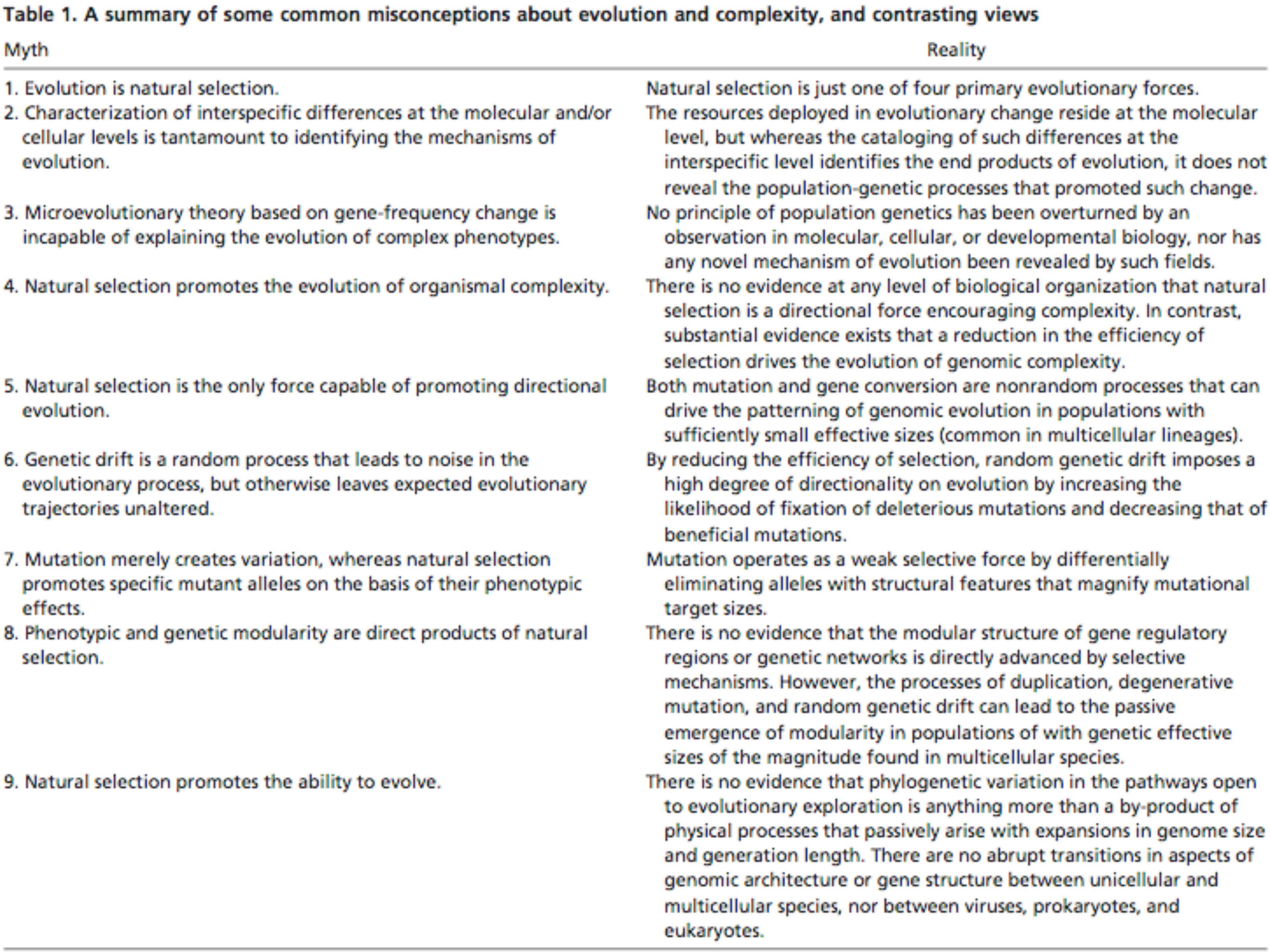
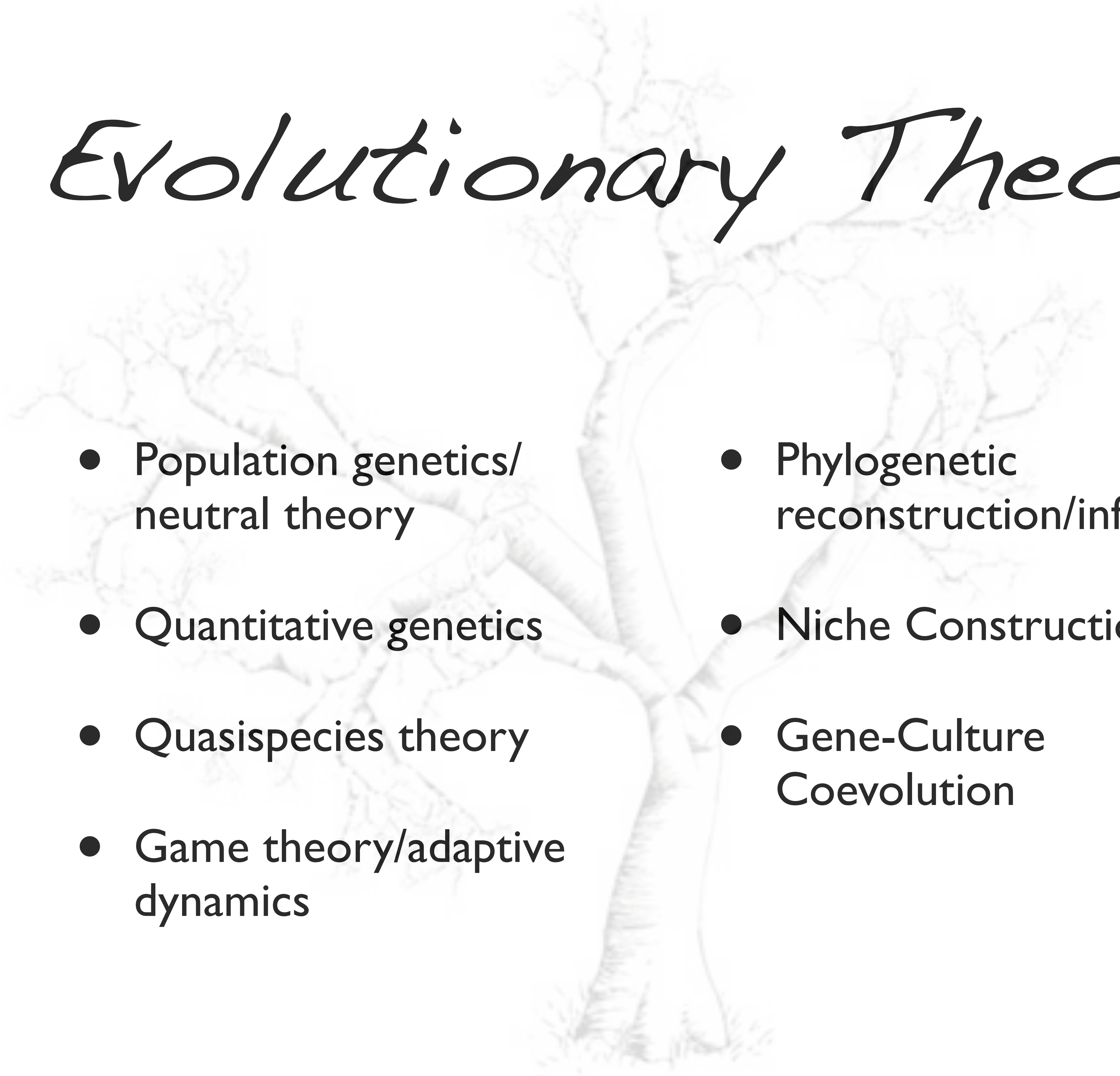


Table 1. A summary of some common misconceptions about evolution and complexity, and contrasting views

Myth	Reality
1. Evolution is natural selection.	Natural selection is just one of four primary evolutionary forces.
2. Characterization of interspecific differences at the molecular and/or cellular levels is tantamount to identifying the mechanisms of evolution.	The resources deployed in evolutionary change reside at the molecular level, but whereas the cataloging of such differences at the interspecific level identifies the end products of evolution, it does not reveal the population-genetic processes that promoted such change.
3. Microevolutionary theory based on gene-frequency change is incapable of explaining the evolution of complex phenotypes.	No principle of population genetics has been overturned by an observation in molecular, cellular, or developmental biology, nor has any novel mechanism of evolution been revealed by such fields.
4. Natural selection promotes the evolution of organismal complexity.	There is no evidence at any level of biological organization that natural selection is a directional force encouraging complexity. In contrast, substantial evidence exists that a reduction in the efficiency of selection drives the evolution of genomic complexity.
5. Natural selection is the only force capable of promoting directional evolution.	Both mutation and gene conversion are nonrandom processes that can drive the patterning of genomic evolution in populations with sufficiently small effective sizes (common in multicellular lineages).
6. Genetic drift is a random process that leads to noise in the evolutionary process, but otherwise leaves expected evolutionary trajectories unaltered.	By reducing the efficiency of selection, random genetic drift imposes a high degree of directionality on evolution by increasing the likelihood of fixation of deleterious mutations and decreasing that of beneficial mutations.
7. Mutation merely creates variation, whereas natural selection promotes specific mutant alleles on the basis of their phenotypic effects.	Mutation operates as a weak selective force by differentially eliminating alleles with structural features that magnify mutational target sizes.
8. Phenotypic and genetic modularity are direct products of natural selection.	There is no evidence that the modular structure of gene regulatory regions or genetic networks is directly advanced by selective mechanisms. However, the processes of duplication, degenerative mutation, and random genetic drift can lead to the passive emergence of modularity in populations of with genetic effective sizes of the magnitude found in multicellular species.
9. Natural selection promotes the ability to evolve.	There is no evidence that phylogenetic variation in the pathways open to evolutionary exploration is anything more than a by-product of physical processes that passively arise with expansions in genome size and generation length. There are no abrupt transitions in aspects of genomic architecture or gene structure between unicellular and multicellular species, nor between viruses, prokaryotes, and eukaryotes.

Lynch, PNAS, 104, 8597-8604 2007

Evolutionary Theory



- Population genetics/ neutral theory
- Quantitative genetics
- Quasispecies theory
- Game theory/adaptive dynamics
- Phylogenetic reconstruction/inference
- Niche Construction
- Gene-Culture Coevolution

Evolutionary Stoichiometry

replication

$$g_i \xrightarrow[\text{Energy + Resources}]{r_i} 2g_i$$

Evolutionary Stoichiometry

competition

$$g_i + g_j \xrightarrow{c_{ij}} g_j$$

Evolutionary Stoichiometry

mutation

$$g_i \xrightarrow[\text{Radiation}]{m_{ij}} g_j$$

$$m_{ij} = \mu^{H(i,j)}(1-\mu)^{L-H(i,j)}$$

Evolutionary Stoichiometry

recombination

$$g_j + g_l \xrightarrow{b_{ijl}} g_i$$

$$b_{ijl} = 1, \ \ \text{if} \ \ i = j = l$$

$$b_{ijl} = \left(\frac{1}{2}\right)(1-c) + c\left(\frac{1}{2}\right)^{H(j,l)} \quad \text{if} \quad i = j \qquad \text{or} \quad i = l$$

$$b_{ijl} = c\left(\frac{1}{2}\right)^{H(j,l)} \quad \text{if} \quad H(i,j) + H(i,l) = H(j,l)$$

Replicator Equation

$$g_i \xrightarrow{r_i} 2g_i$$

$$g_i + g_j \xrightarrow{c_{ij}} g_j$$

$$n \quad \text{genomes}$$

$$\dot{g}_i = g_i(r_i - \bar{f})$$

$$\text{where} \quad \bar{f} = \sum_i^n r_i g_i \quad \text{and} \quad c_{ij} = 1$$

Evolutionary Game Theory: Frequency dependent Replicator Equation

$$g_i \xrightarrow{r_i(\mathbf{g})} 2g_i$$

$$g_i + g_j \xrightarrow{c_{ij}} g_j$$

n genomes

$$\dot{g}_i = g_i(r_i(\mathbf{g}) - \bar{f})$$

where $\bar{f} = \sum_{i}^{n} r_i(\mathbf{g}) g_i$ and $c_{ij} = 1$

Evolutionary Game Theory:
Frequency dependent Replicator Equation

$$\dot{g}_i = g_i(r_i(\mathbf{g}) - \bar{f})$$

$$\text{Payoff Matrix } \; P = [p_{ij}]$$

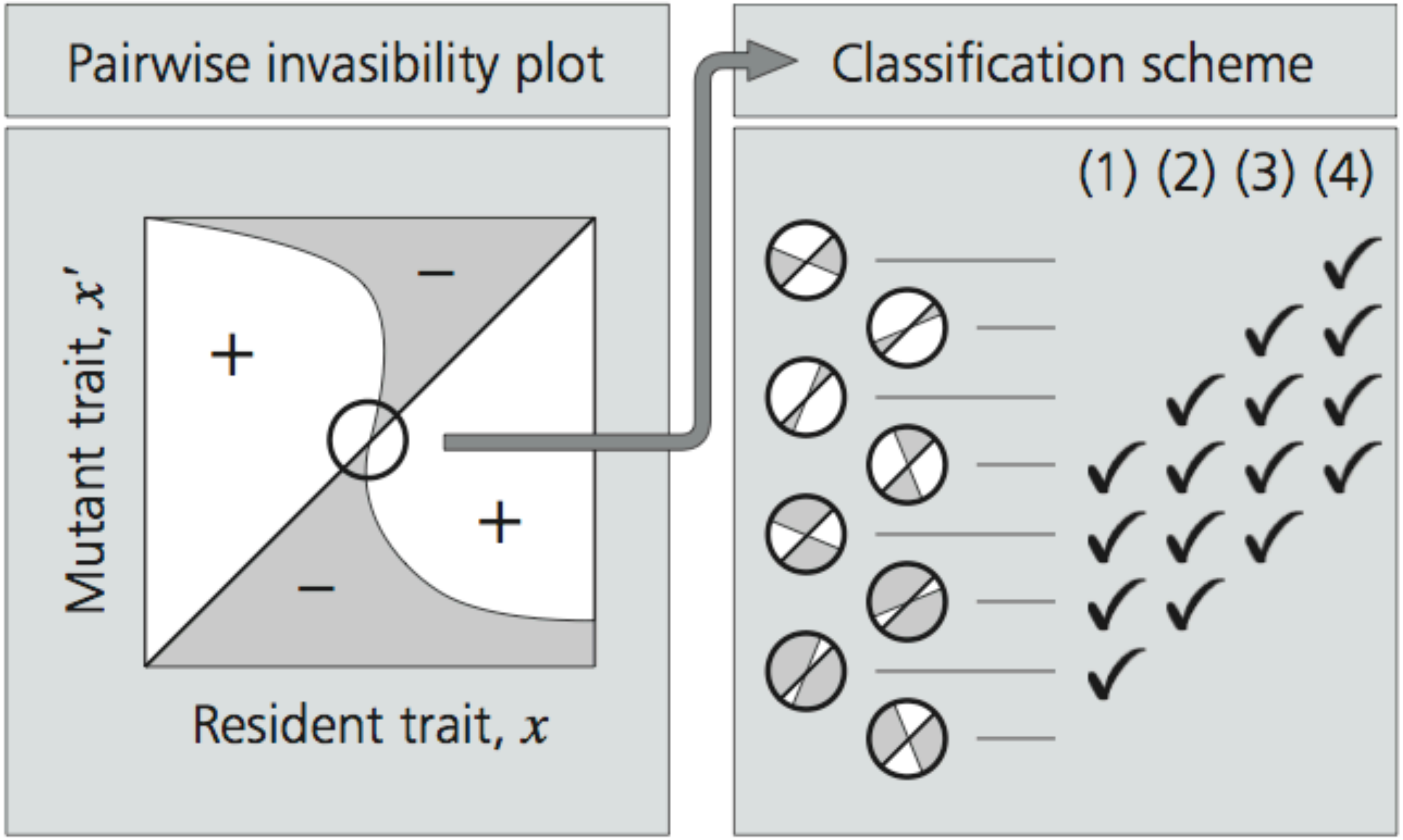
with linear payoffs:

$$r_i(\mathbf{g}) = \sum_{j}^{n} g_j p_{ij}$$

Evolutionary Game Theory: Frequency dependent Replicator Equation

$$\dot{g}_i = g_i(\sum_j^n g_j p_{ij} - \sum_j^n g_j \sum_k^n g_k p_{jk})$$

Evolutionary Game Theory: Adaptive Dynamics for Continuous Traits



$$\frac{dx}{dt} = \frac{1}{2}\mu\sigma^2\bar{N}(x)\frac{\delta f(x',x)}{\delta x'}\Bigg|_{x'=x}$$

Freq-dep Replicator Equation & Bayesian Inference

An Insight by Cosma Shalizi

$$\frac{\Delta g_i(t)}{\Delta t} = g_i(t-1)(r_i(\mathbf{g}) - \bar{f})$$

$$P(X|Y) = P(X)\frac{P(Y|X)}{P(Y)}$$

$$P(X|Y) = P(X)\frac{L_X}{\bar{L}}$$

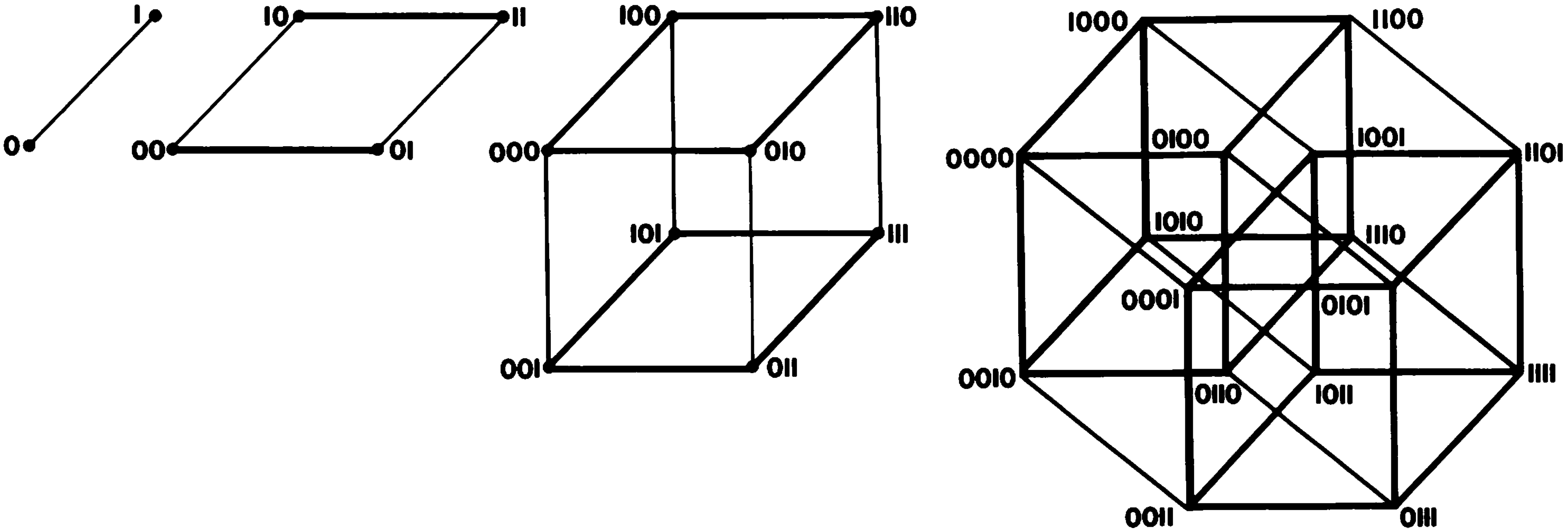
$$\bar{L} = P(Y) = \sum_{x \in \omega} P(Y|X)P(X)$$

$$P_X(t) = P_X(t-1)\frac{L_X}{\bar{L}}$$

$$\Delta P_X(t) = P_X(t-1)(\frac{L_X}{\bar{L}} - 1) = P_X(t-1)\frac{1}{\bar{L}}(L_X - \bar{L})$$

$$\Delta P_X(t) = P_X(t-1)(f_t - \bar{f}), \quad \text{where} \quad f_t = L_X/\bar{L}$$

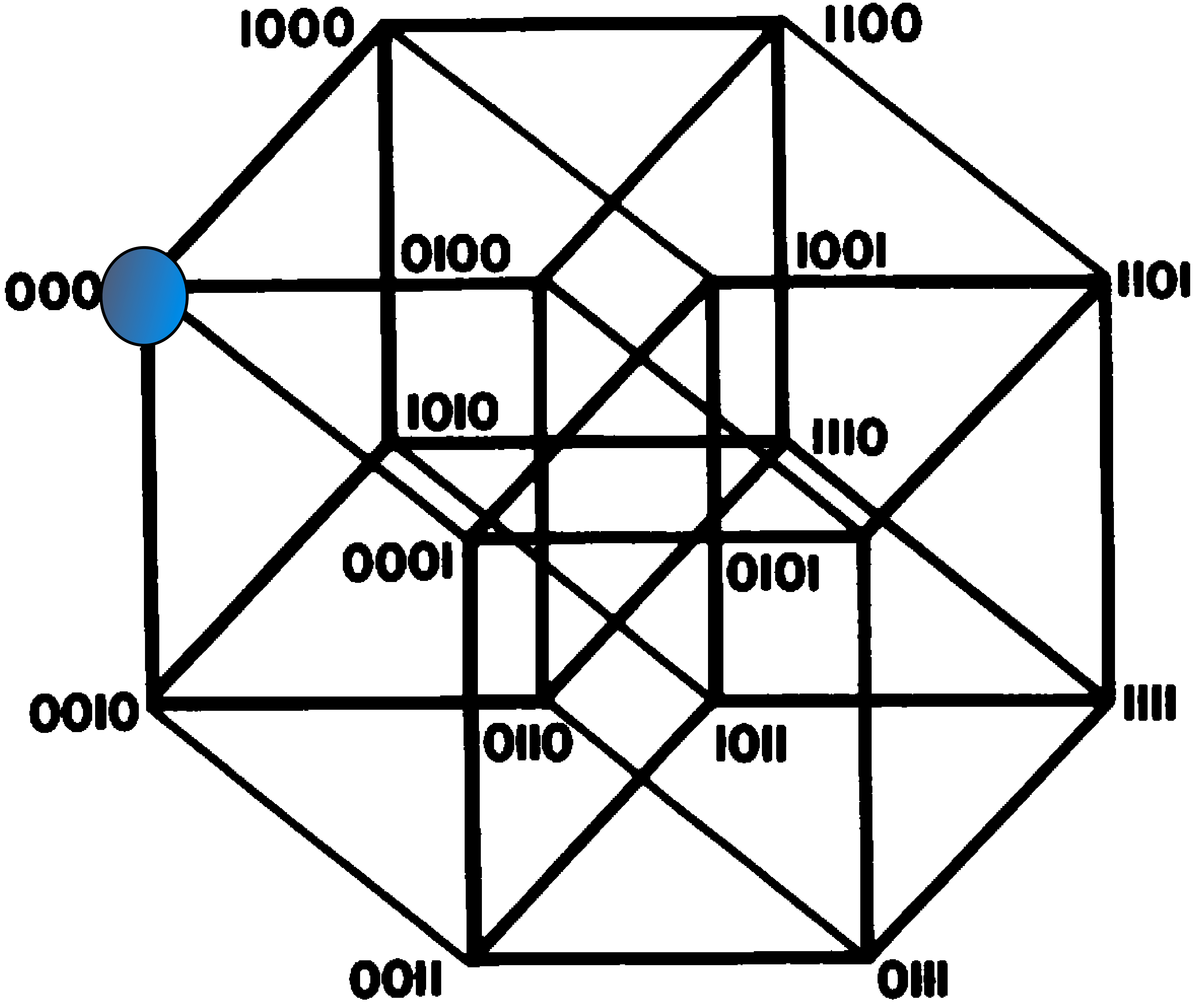
Sequence Space
& Limits to Evolution

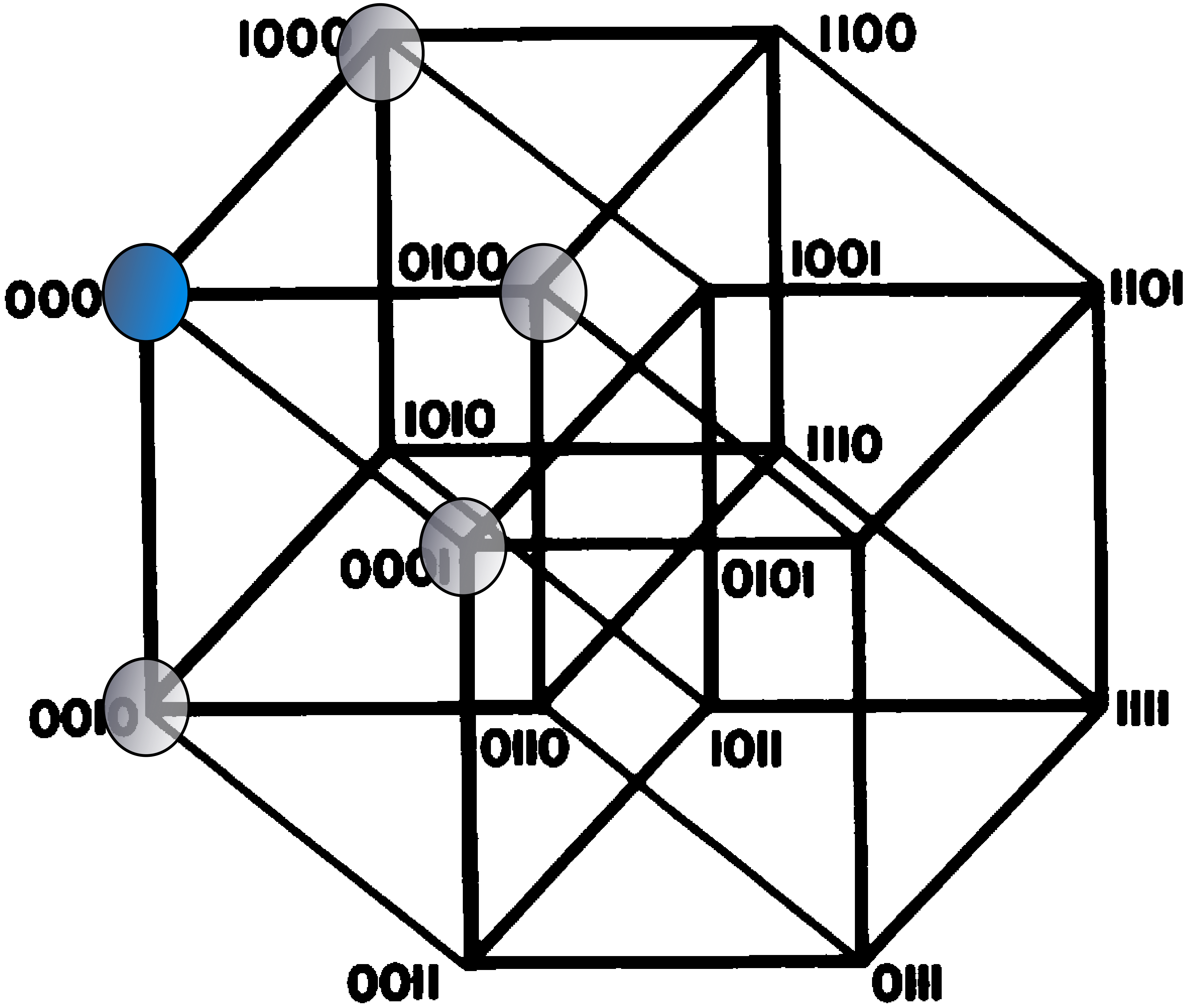

1
0
10
11
00
01
100
110
000
010
101
111
001
011
1000
1100
0100
1001
0000
1101
1010
1110
0001
0101
0010
1111
0110
1011
0011
0111

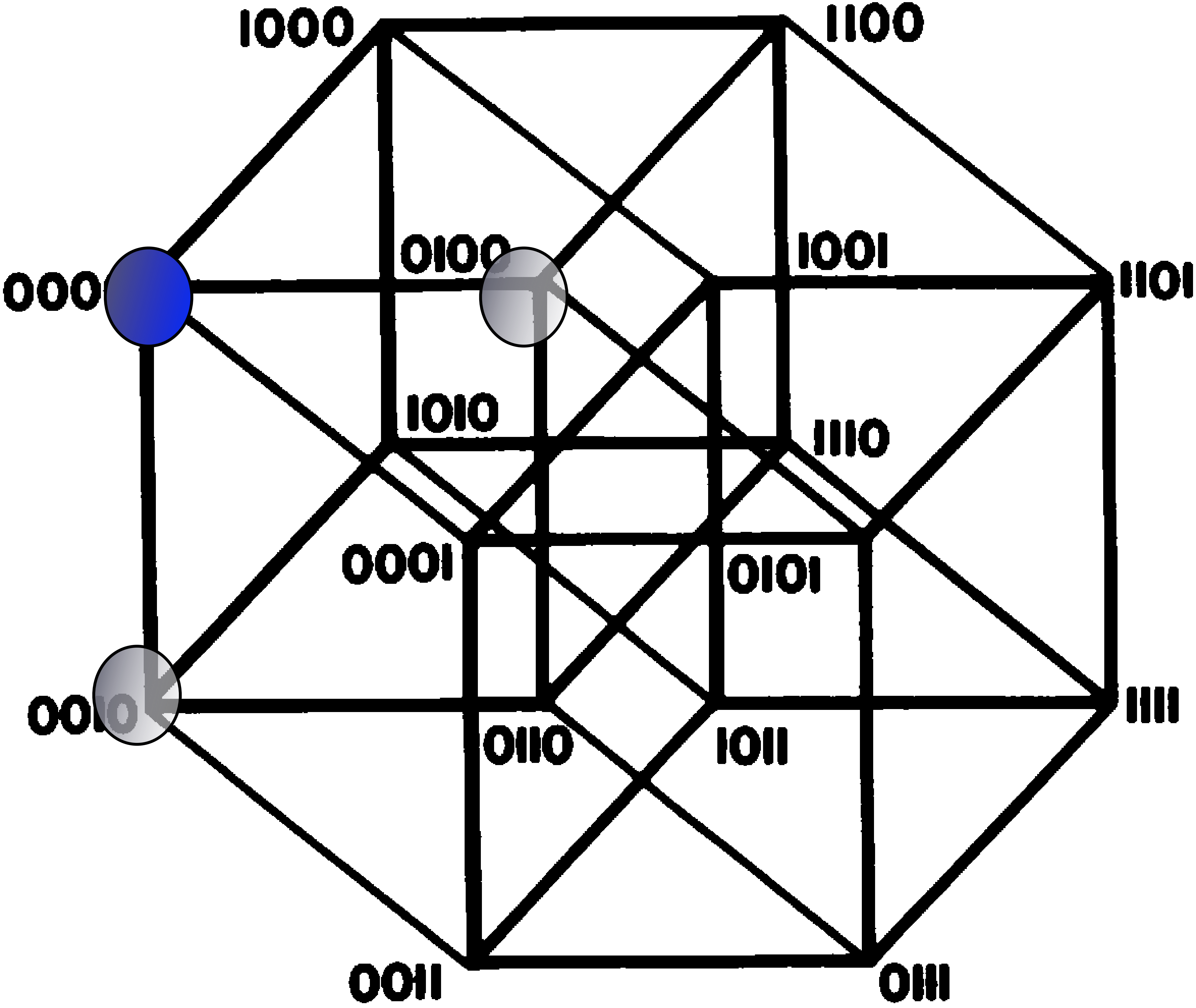
Replicator-Mutator Equation

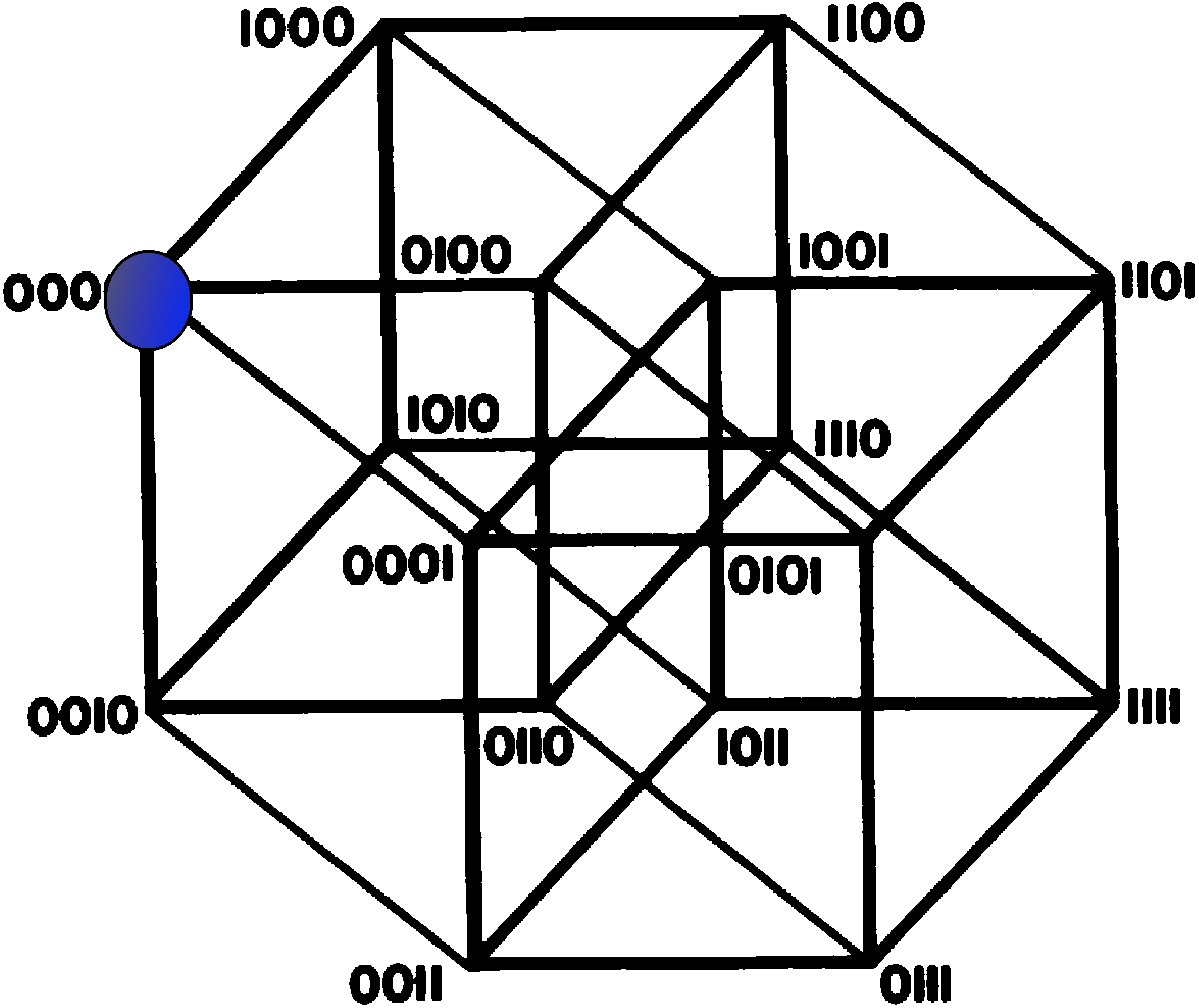
$$\dot{g}_i = \sum_{j}^{2^n} g_j r_j(\mathbf{g}) m_{ij} - g_i \bar{f})$$

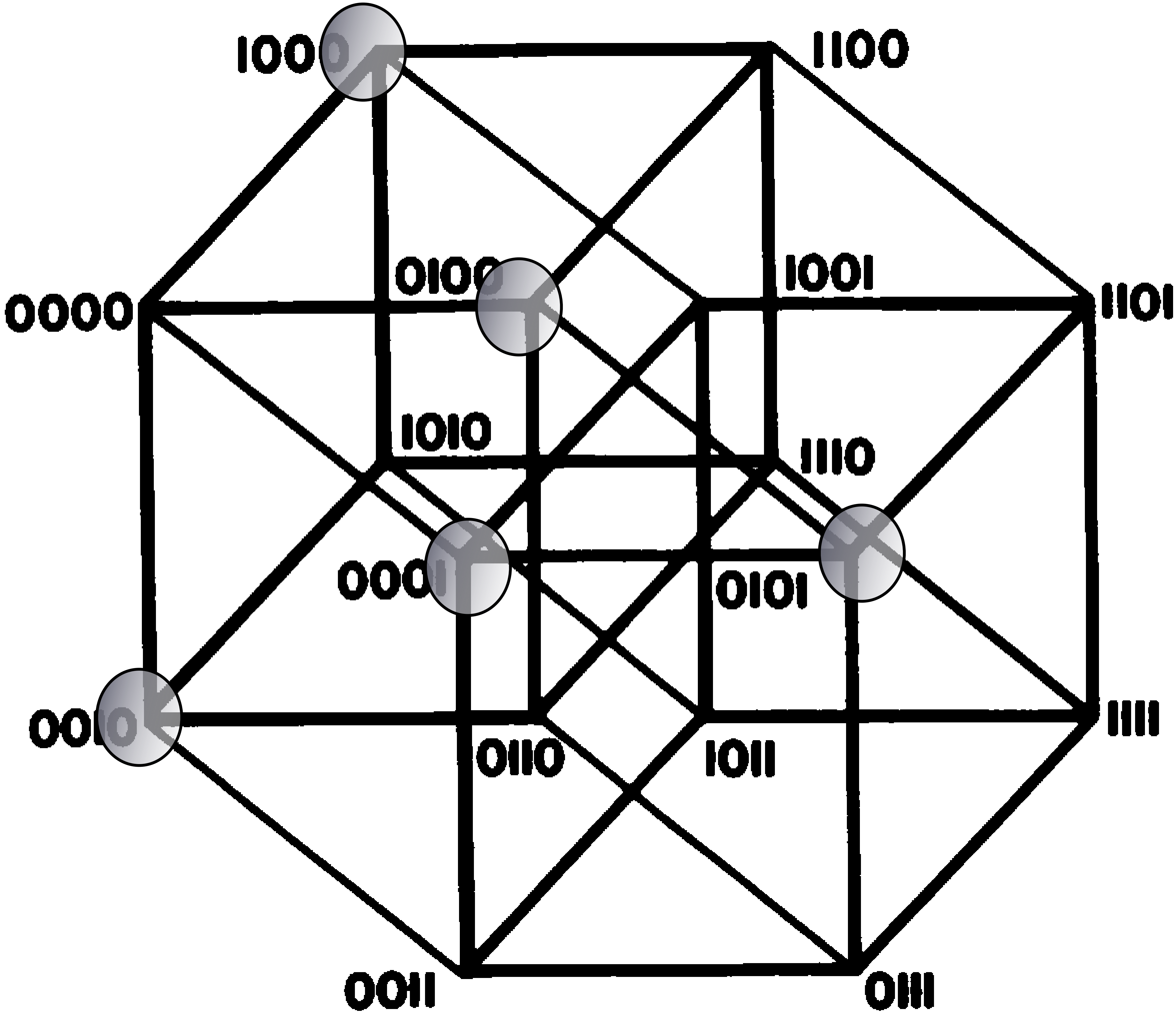
$$m_{ij} = \mu^{H(i,j)}(1-\mu)^{L-H(i,j)}$$

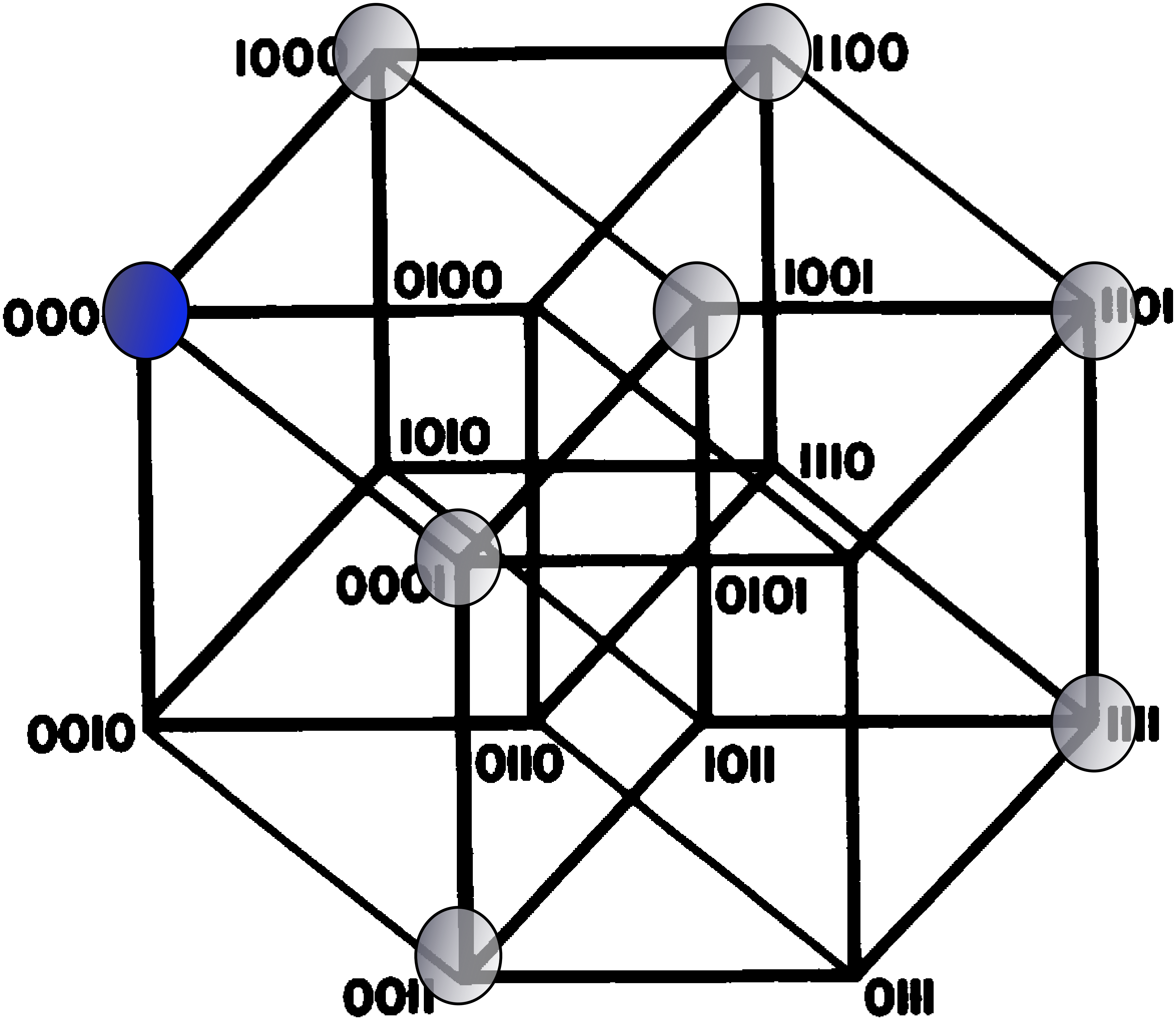

1000
1100
0100
1001
1101
000
1010
1110
0001
0101
0010
0110
1011
1111
0011
0111

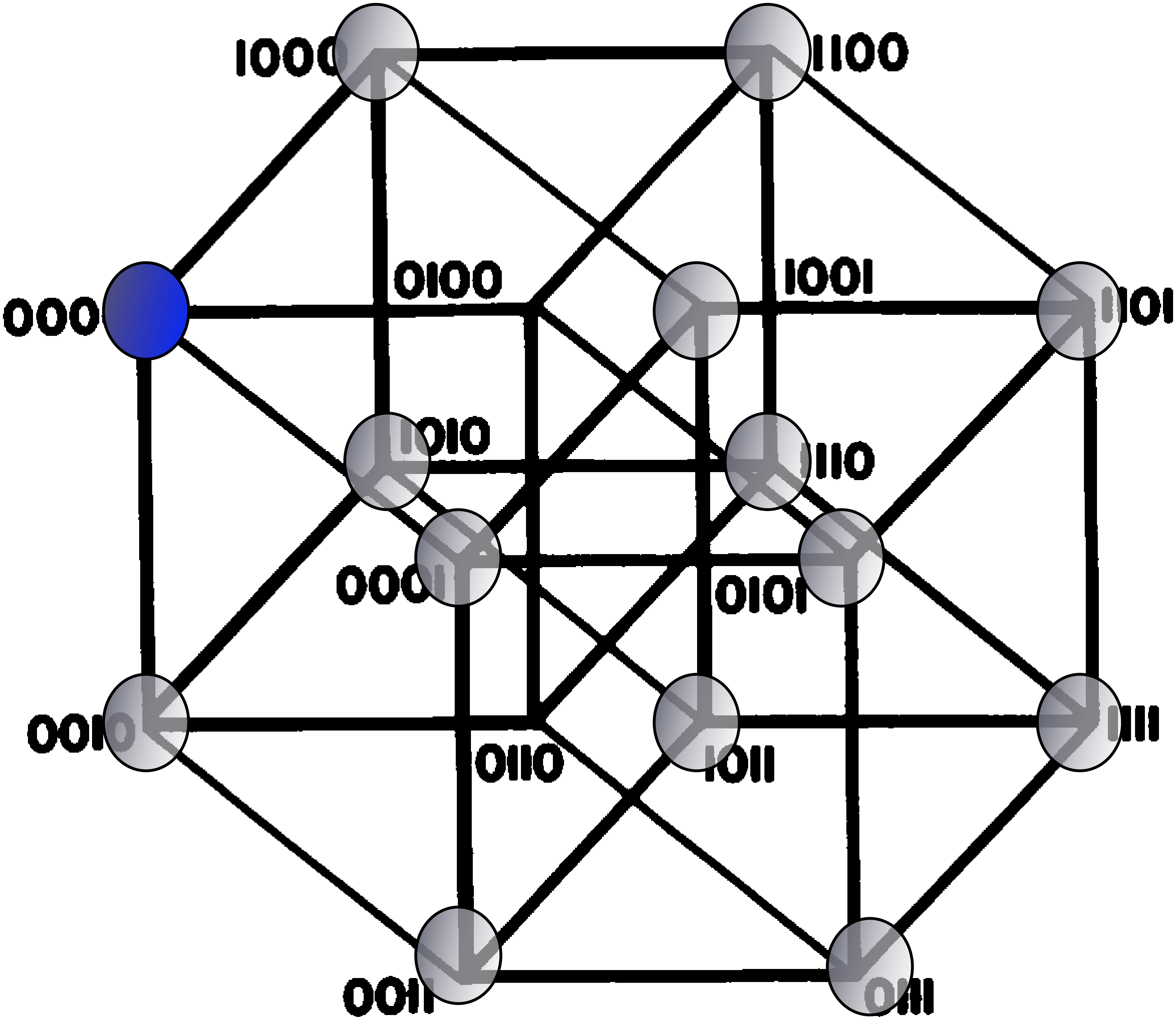

1000
1100
0100
1001
000
1101
1010
1110
0001
0101
0010
0110
1011
1111
0011
0111


1000
1100
0100
1001
000
1101
1010
1110
0001
0101
0010
1111
0110
1011
0011
0111

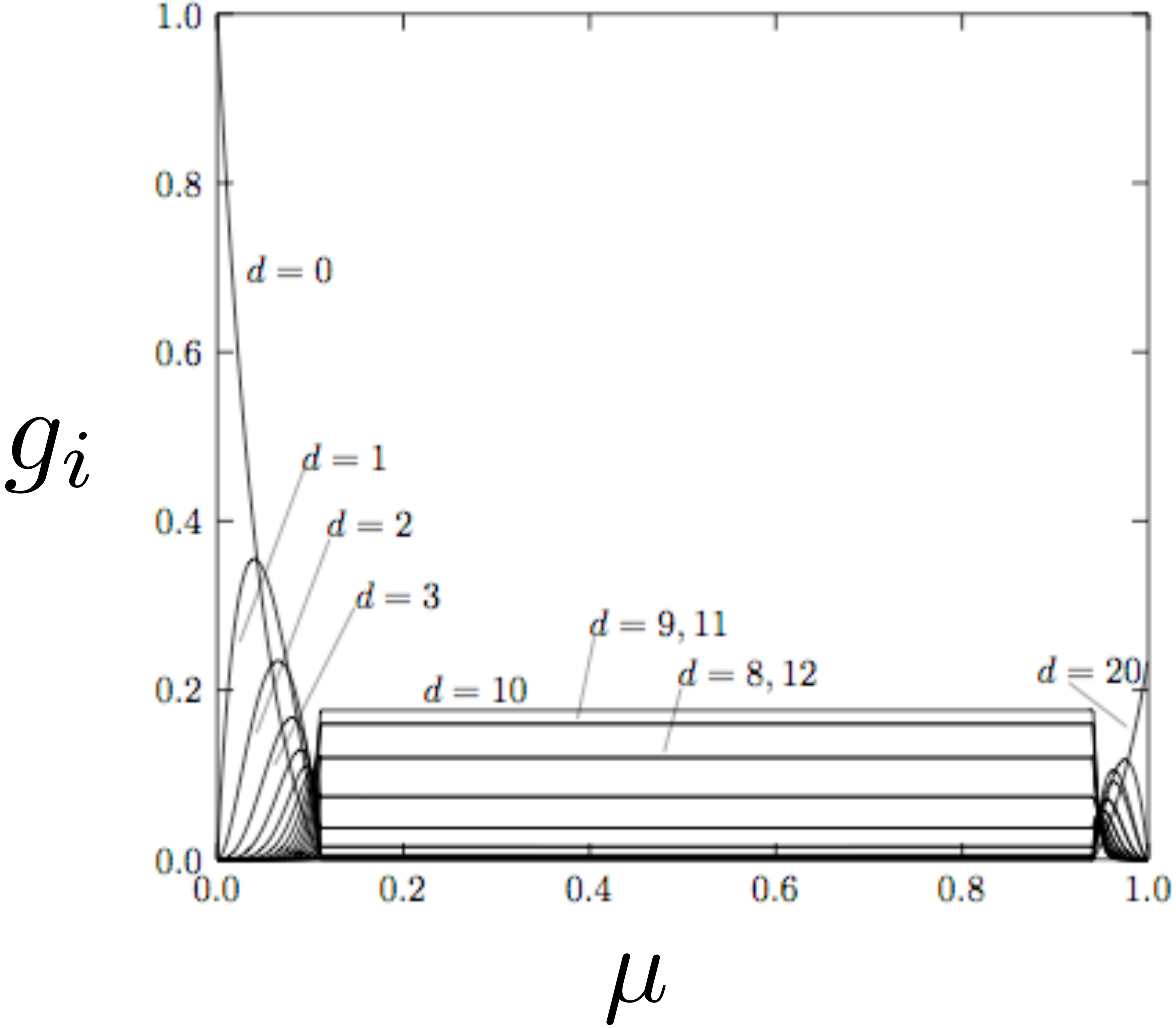

1000
1100
0100
1001
0000
1101
1010
1110
0001
0101
0010
1111
0110
1011
0011
0111


1000
1100
0100
1001
0000
1101
1010
1110
0001
0101
0010
0110
1011
1111
0011
0111


1000
1100
0100
1001
000
1101
1010
1110
0001
0101
0010
0110
1011
1111
0011
0111


1000
1100
0100
1001
000
1101
1010
1110
0001
0101
0010
1111
0110
1011
0011
0111

Error Threshold



Fitness Landscape

Delta function:

$$\mu < \frac{s}{L} = \frac{1}{L}$$

Multiplicative function:

$$\mu < s$$

Kimura's Neutrality Inequality

Expectation of Mutation

$$N\mu$$

Probability of Fixation

$$\frac{1}{N}$$

Condition for Neutrality

$$sN < 1$$

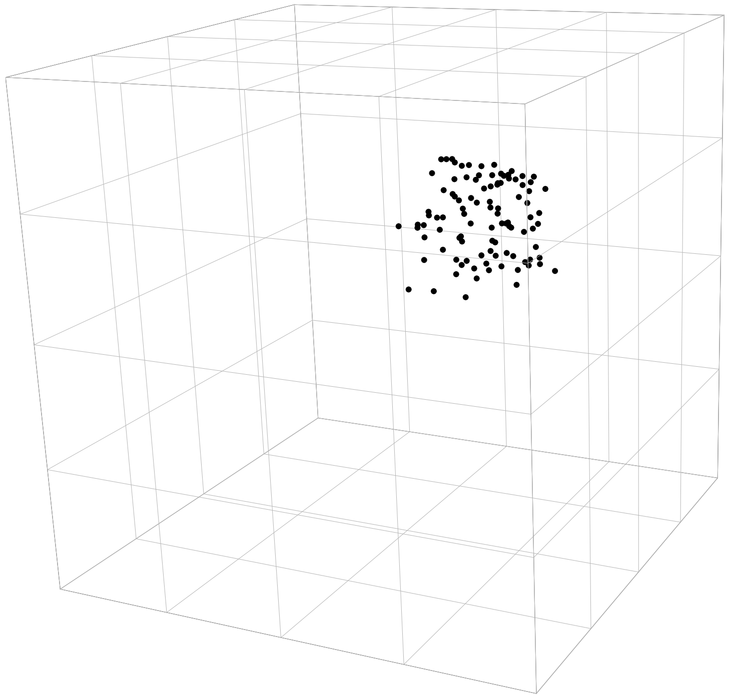
Evolutionary Information Storage

Information Conserved

organismal regularity

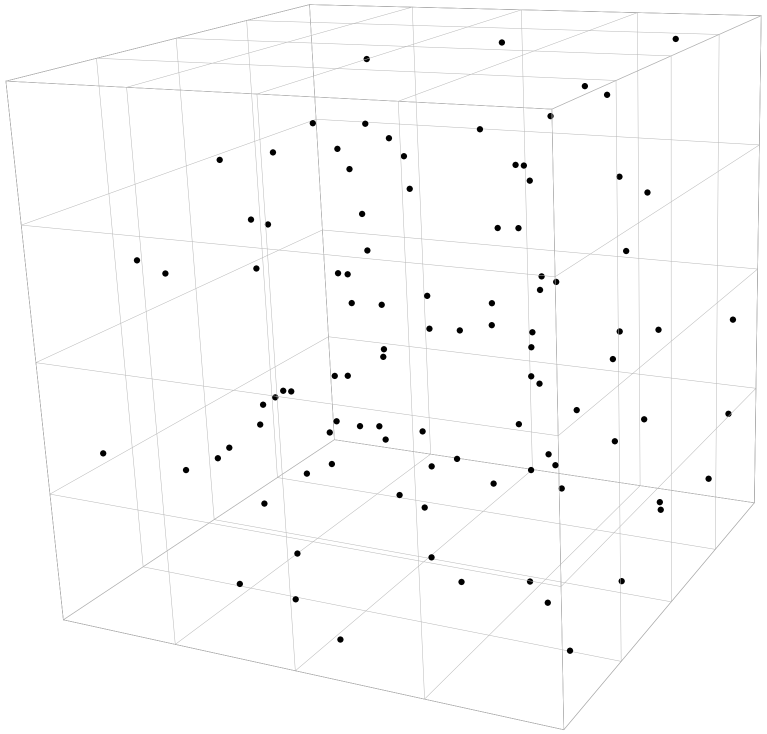
Environmental regularity

$$s \geq \frac{\mu L}{N}$$



Information Lost

$$s < \frac{\mu L}{N}$$



Evolution, Localization & Information

(with a 4 letter alphabet)

Information as Selective Uniformity

Aligned genomes

ACGTC...T
ATGTG...T
ATCTG...A

1 2 3 4 5...l

$$H_i = -\sum_j p_j^{(i)} \log_4 p_j^{(i)}$$

$$I_i = H_{max} - H_i$$

Information in Population

$$C = L - \sum_i H_i$$

Information-Selection as a Compression Ratio

Aligned genomes

ACGTC…T
ATGTG…T
ATCTG…A

1 2 3 4 5…l

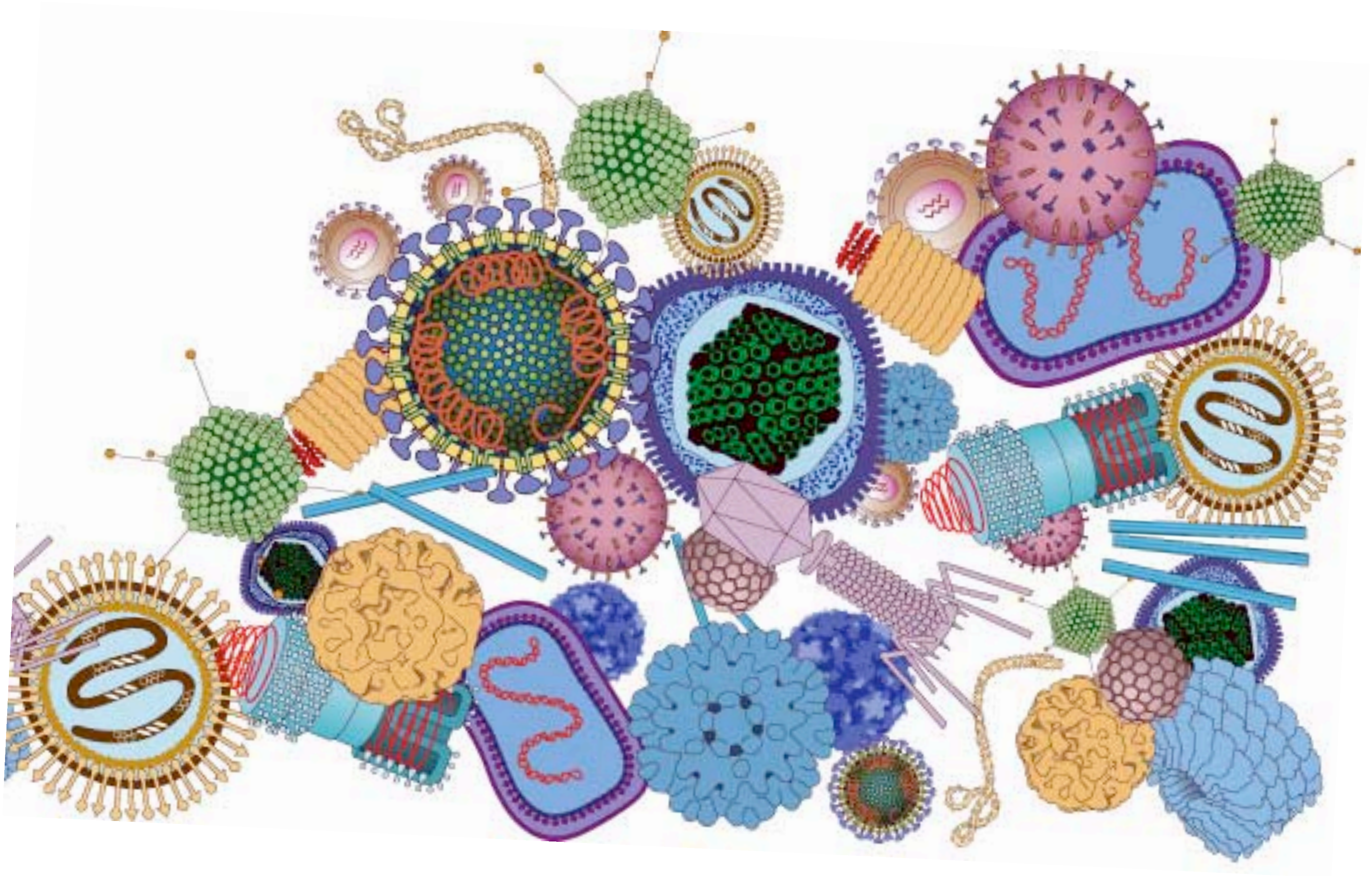
$$H_i = -\sum_j p_j^{(i)} \log_4 p_j^{(i)}$$

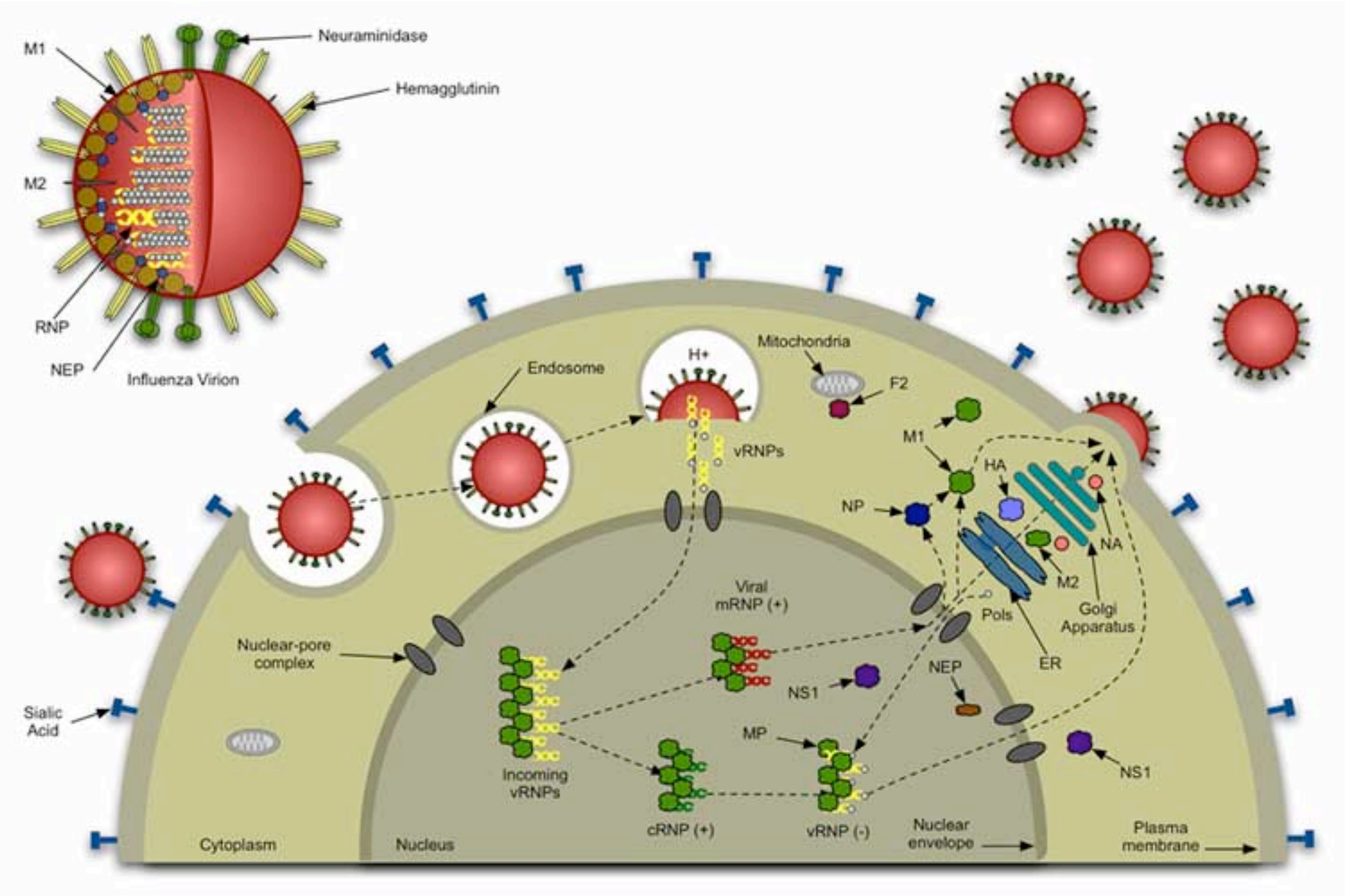
$$C = \frac{L H_{max}}{\sum_i^L H_i} = \frac{L}{\sum_i^L H_i}$$

$$sN > \mu L$$

The Tendency to Population Multiplicity & Individual Minimality

THE SPACE GOD VERSUS THE ANDROMEDA STRAIN





www.genomeknowledge.org

KEY

virus catalyzed

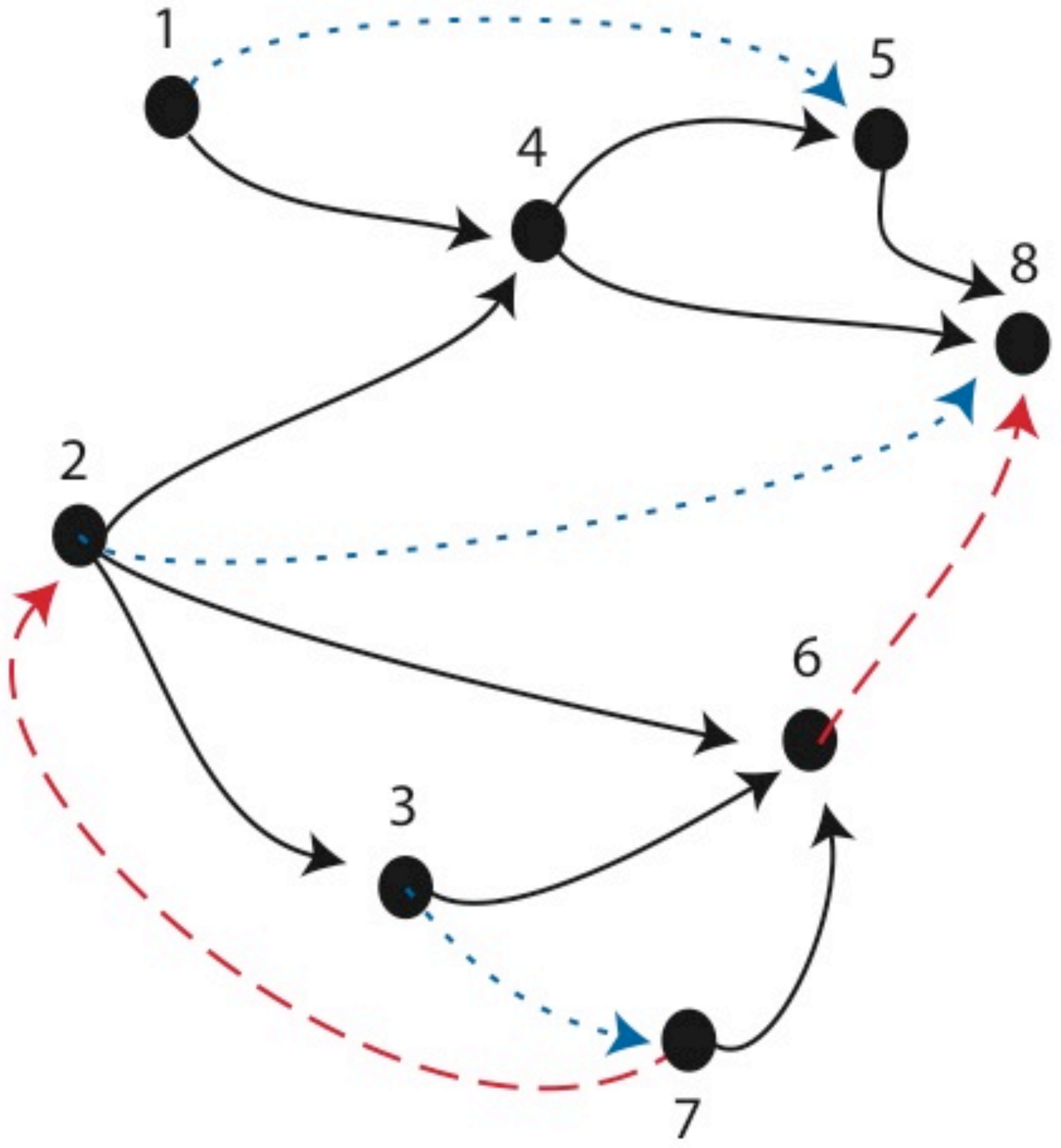
→

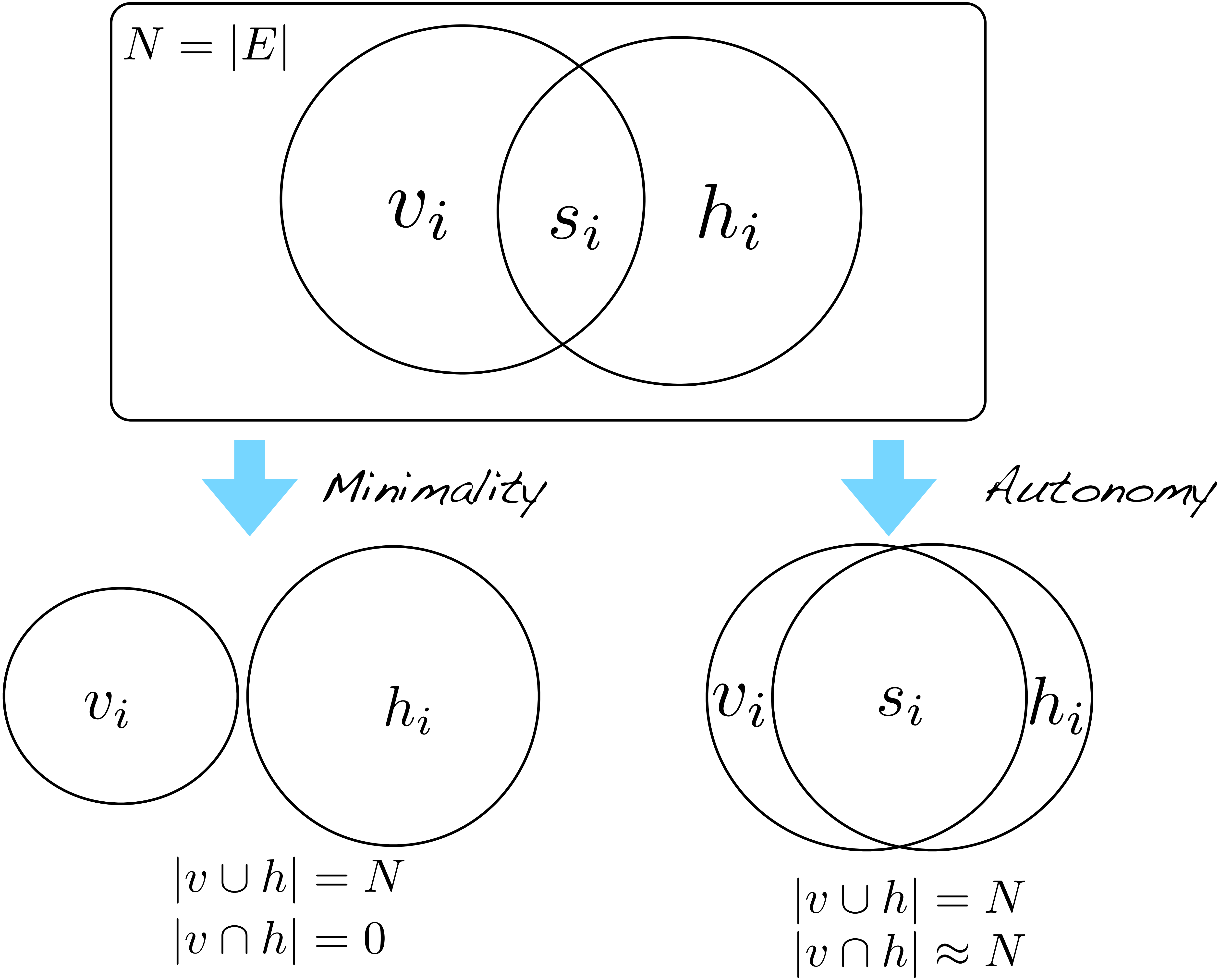
host A catalyzed

⇢

host B catalyzed

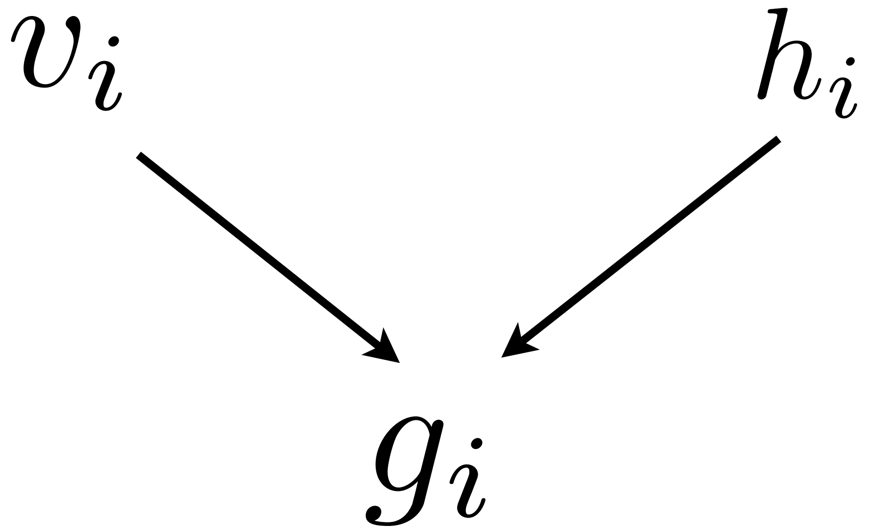
⇢




$N = |E|$
v_i
s_i
h_i
Minimality
Autonomy
v_i
h_i
v_i
s_i
h_i
$|v \cup h| = N$
$|v \cap h| = 0$
$|v \cup h| = N$
$|v \cap h| \approx N$

$$w = \prod_{i \in E} f(g_i) \sum_{j \in NE} h(g_j)$$

$$w_E = \prod_{i \in E} f(g_i)$$



$$\begin{tikzcd}
v_i \arrow[dr] & & h_i \arrow[dl] \\
& g_i &
\end{tikzcd}$$

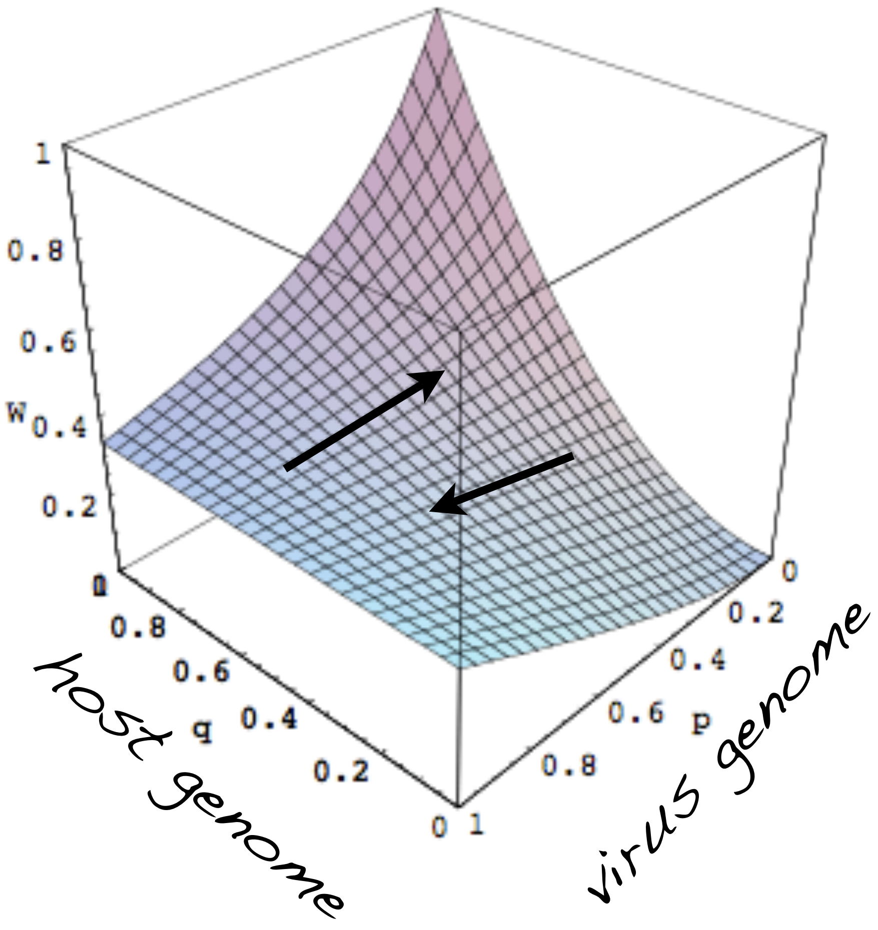
$$g_i = h_i \vee v_i$$

$$Prob(h_i = 1) = q \qquad Prob(v_i = 1) = p$$

$$L_v = \sum_i v_i, \qquad < L_v >= pN$$

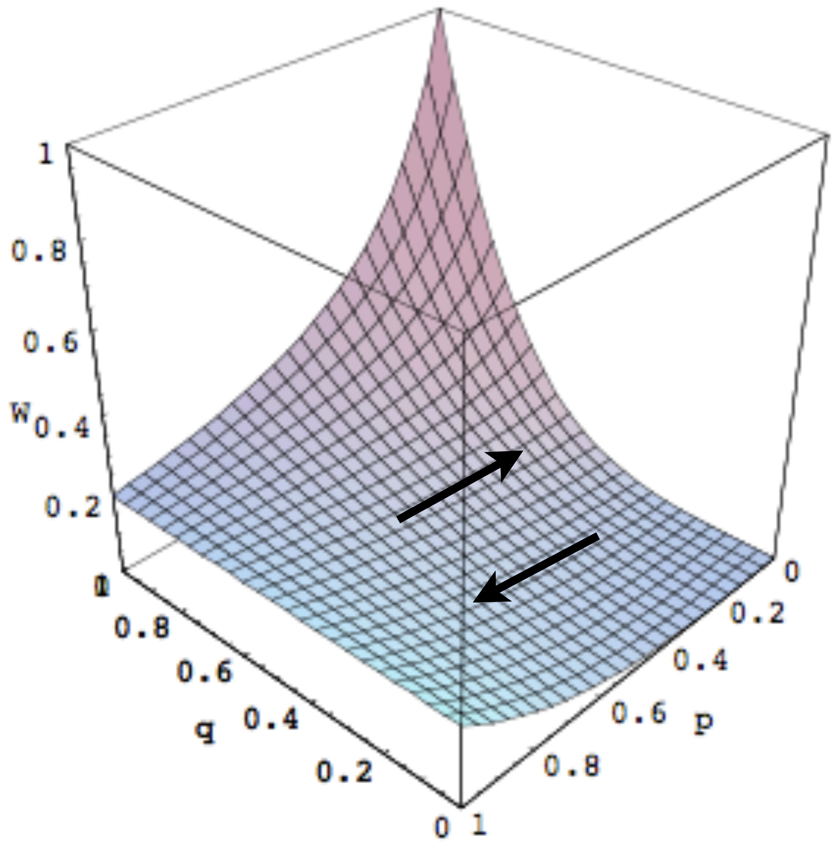
$$w = \frac{\prod_{i \in N} g_i}{(1 + L_v)} \qquad < w >= \frac{(q + p - qp)^N}{1 + pN}$$

N=4

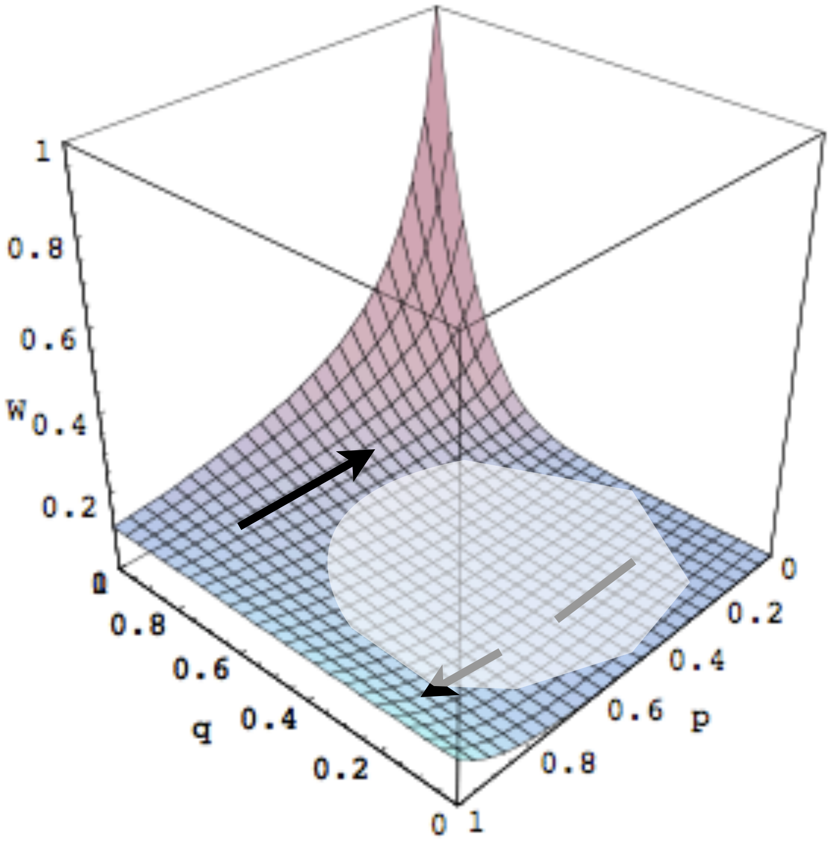


1
0.8
0.6
W
0.4
0.2
0
0.8
0.6
q
0.4
0.2
0
host genome
1
0.8
0.6
p
0.4
0.2
0
virus genome

N=8

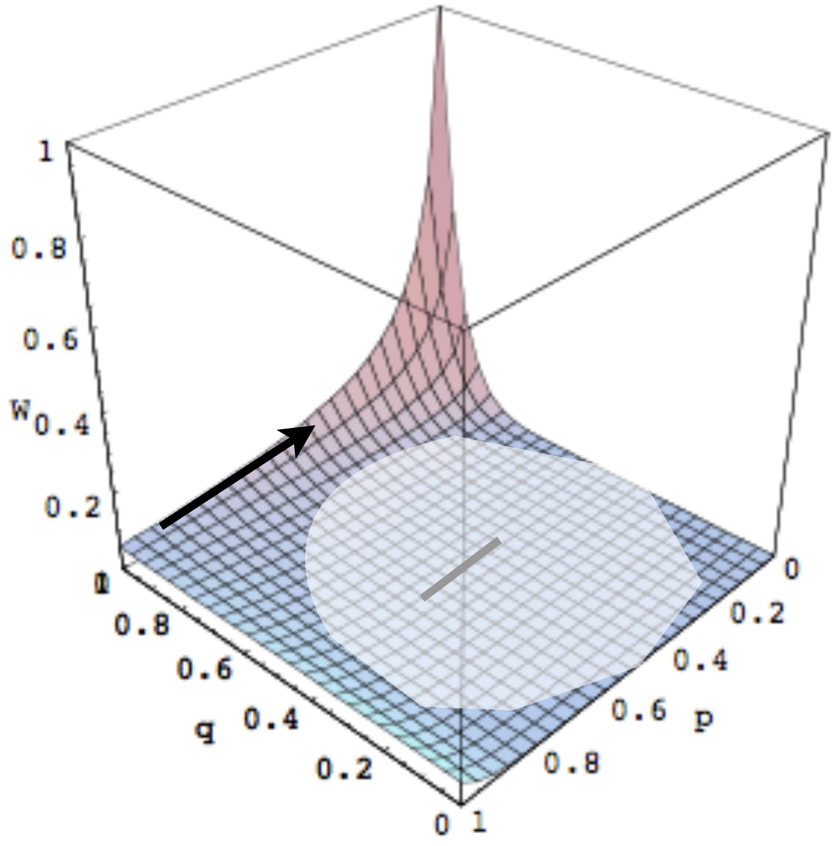


1
0.8
0.6
W
0.4
0.2
0
0.8
0.6
q
0.4
0.2
0
1
0.8
0.6
p
0.4
0.2
0



1
0.8
0.6
W
0.4
0.2
0
0.8
0.6
q
0.4
0.2
0
1
0.8
0.6
p
0.4
0.2
0

N=16



1
0.8
0.6
W
0.4
0.2
0
0.8
0.6
q
0.4
0.2
0
1
0.8
0.6
p
0.4
0.2
0

N=32

Evolution & Minimality

- Evolutionary **theory** is concerned largely with the frequency, variety & relationships among chemically improbable sequences
- Replicator/Mutator Dynamics tends to favor small sequences which can be preserved
- Neutral theory requires large populations for effective selection of target sequences -- more likely with small sequences
- Evolution can be thought of as an inferential, model fitting process (Bayesian) and selection as a mechanism for injecting information into sequences

- Genomes tend to "neutralize" and/or eliminate redundant sequences - genomes need not encode perfectly predictable resources
- Increasing autonomy (often size) reflects greater inferential uncertainty
- Genetic dissipation in co-evolutionary contexts requires a consideration of rates of gene inactivation in all interacting agents
- Genome growth beyond competitive persistence is driven by, e.g. robustness & control

Whence Complexity?