MEDICAL UNIVERSITY – PLEVEN
FACULTY OF PHARMACY

DIVISION OF PHYSICS AND BIOPHYSICS, HIGHER MATHEMATICS AND INFORMATION TECHNOLOGIES

LECTURE No4

ORDER AND PROBABILITY INFORMATION AND ENTROPY

Thermodynamic probability and entropy. Boltzmann equation of entropy. Statistical definition of entropy. Shannon relation of information content. Maxwell's demon

Prof. M. Alexandrova, DSc

Thermodynamic Probability and Entropy

Let us consider the molecular basis of entropy.
Since all natural processes lead to a net increase in entropy and also since there exists a natural tendency for an ordered set of things to get disordered through interplay of natural forces, we can identify increase of entropy (dS) as increase of disorder or randomness of a system on a molecular level.

Thus, we can say:

1. Entropy is a measure of the degree of randomness or disorder of a system.

2. The entropy of a system increases when it becomes more disordered.

3. A process can occur spontaneously only if the sum of the entropies of the system and its surroundings increases.

Let us start with the assumption that entropy is a measure of randomization of a given distribution.

We will consider a system of maximum entropy as a system in maximal disorder.

Furthermore, let us demand that the entropy be an *extensive* parameter. Therefore, like volume, or mass, but in contrast for example to temperature or density, the entropies S_1 and S_2 of two systems can be added, if these systems come together: $S_1+S_2 = S$.

How can we now define a parameter, which indicates the degree of randomization or, on the contrary, the degree of order? What does order of organization mean?

Let us consider the distribution of four distinguishable spheres in two compartments of a box.

Let each of these spheres, independently of the three others, fall just by chance into one or the other compartment of the box. All of the 11 possibilities of the distribution have the same degree of probability, because the probability of each sphere individually falling into compartment 1 or into compartment 2 is equal.

Summary: there is only one way to realize the distributions 0:4 and 4:0. In contrast, there are four ways to realize the distribution 3:1 and 1:3, and, finally, six ways for equal distribution: 2:2.

Let us now ignore the fact that the spheres are distinguishable. Let us simply ask: how large is the probability that just by stochastic distributions one of the relations 4:0, 3:1, 2:2, 1:3, or 0:4 occurs?

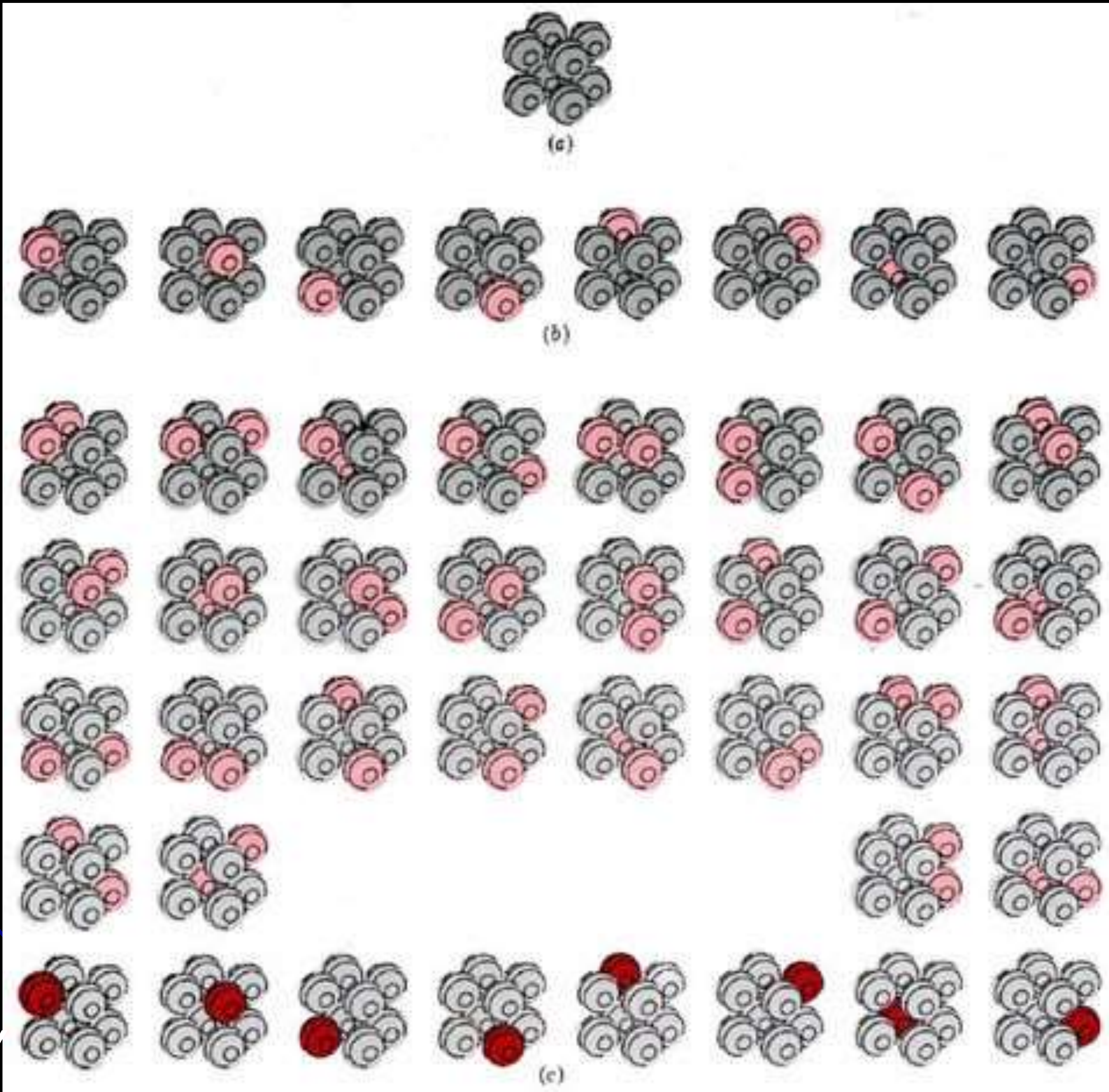
The probability of any kind of **distribution will be larger** if it can be realized **in a larger number of ways.** The distribution mode 2:2, for example, is six times more likely than the distribution 4:0, or 0:4.

The number of ways, which lead to the realization of a definite situation is a measure of the probability of the occurrence of it. We will designate this number of ways by the parameter *W* that we will call *thermodynamic probability.*

$1 < W < \infty$ $0 < P < 1$

Conclusion: if *W* really is a measure of the probability of getting a definite distribution, and if an increase of the degree of order is the most uncertain result of a stochastic distribution and finally, if the entropy (S) is a parameter, indicating the degree of disorder, then S should be a function of *W*.

The thermodynamic probability W of a crystal containing eight atoms at three different T. (a) At 0 K there is only one way in which the crystal can be arranged, so that W = 1. (b) If enough energy is added to start just one of the atoms vibrating (color), there are eight different equally likely arrangements possible, and W = 8. (c) If the energy is doubled, two different atoms can vibrate simultaneously (light color) or a single atom can have all the energy (dark color) ; W = 36.



If two situations with probabilities W_1 and W_2 are connected together, then the probability of this combined situation results from the product (W_1 x W_2).
$S=f(W)=S_1+S_2=f(W_1)+f(W_2)=f(W_1 \times W_2)$.

This demand is met by the logarithmic function:
$\ln A+\ln B=\ln(A \times B)$.
Hence entropy is proportional to logarithm of *W*:
$S=k \ln W$ This is the Boltzmann equation of entropy.

Boltzmann's constant k was defined as a universal constant later by Max Planck. It must have the same unit of measure as entropy.

$$k = \frac{R}{N_A}$$

k= 1.3806488(13)×10^{-23} J/K

Information and Entropy

Shannon [Claude Elwood (1916 -2001), an American mathematician, known as "The father of Information Theory“] introduced a parameter into information theory which was related to S and was named *information*. The information of a message depends on the effort required to guess it by a highly set system of questions.

- Not difficult to guess the result of the toss of a coin, since there are only two possibilities of equal probability.
- To guess a certain card in a full deck of playing cards is much more difficult - a large number of yes-no questions have to be answered.

The information content of a playing card > of a tossed coin. Should a deck consist of cards, which are all the same, and this is known, guessing will not make sense at all. The information content of each of these cards is zero.

Def. The probability by which possibilities are turned into reality seems to become a measure of information.

In information theory the mathematical term **probability** (P) is used, which is defined as follows:

$$P = \frac{\text{number of favorable cases}}{\text{greatest posible number of cases}}$$

On average, coins tossed a hundred times will land with heads up in 50 instances. Hence, the probability of heads facing is ½. On the other hand, the probability of throwing a "six" with some dice is only P = 1/6, whereas the probability of throwing one of the three even numbers would be higher: $P = 3/6 = 1/2$. Whereas the thermodynamic probability (W) is always larger than 1, the value of the mathematical probability lies between 0 and 1 ($0 < P < 1$). $P = 0$ means *impossibility*, while P = 1 expresses *absolute certainty*.

Information (I) is a function of mathematical probability (P): $I=f(P)$. The conditions for the function f are again satisfied by a logarithmic function, since the multiplication rule for the calculation of probabilities must be valid.

Therefore: $I=K \ln P$ - the *Shannon equation* of information theory. The unit of I is determined by the unit of the factor K. The bit ("binary digit") is most commonly used. It expresses the number of binary yes-no decisions, which are needed to determine a given message.

E.g. the one side of the coin can be guessed by one single decision, its information value is 1 bit. Five binary decisions will be sufficient to guess a card from a deck. Hence, the information value of one card is 5 bits.

The factor $K = -1/\ln 2 = -1.443$ must be used to calculate I in bits. If Boltzmann's constant (k) is used ($K = -k$), then information is obtained formally in entropy units.

The idea of a link between information and entropy was first suggested by Boltzmann [Ludwig Eduard (1844 –1906), an Austrian physicist]. Erwin Schrodinger (1944) made the frequently quoted statement: *"The living system feeds on negative entropy"*. This is the reason why sometimes the term "*negentropy*" is used.

What is the real importance of Shannon's information equation in biophysics?

Ans: It is possible to calculate the information content of a protein by the above approach. The requirements are, firstly, a statistical record of the frequency of the occurrence of the individual amino acids in proteins.

- This will provide the probability (P) for the presence of a given amino acid at a certain locus in the protein.
- Using Shanon equation the information content of each monomer can be calculated.
- The information content of a whole protein can be obtained by addition of the values of its monomers.

The information content of a nucleic acid can be obtained in the same way.

One mammalian mitochondrial DNA molecule consists of about 15000 nucleotides. Assuming that the four possible types of nucleoside bases (adenine, cytosine, guanine, thymine) have an equal probability of occurrence, then the information content of each single nucleotide will have a value of 2 bits. The information capacity of one DNA molecule, in this case, amounts to 30000 bits.

Example: the information content of an English text can be calculated from the frequency of the use of individual letters. In this way one can derive the information content of a word, a sentence, and even of a textbook.

It is obvious that this parameter does not reveal anything about the "information value" of the book as generally understood. The same information (*I*) would be given by any other book with the same number of meaningless strings of English words.

Does this situation invalidate Shannon's information concept? Everybody knows how important the calculation of information is today in the field of computer science.

Is it possible then to quantify biologically important information? Does it mean that the information concept is not applicable to biological systems at all? There is no reason for scepticism.

A distinction has to be made between a syntactic measure of information, and a semantic measure. The *syntactic* information content of a DNA molecule, as calculated above, provides some information on the maximum storage capacity of the genome.

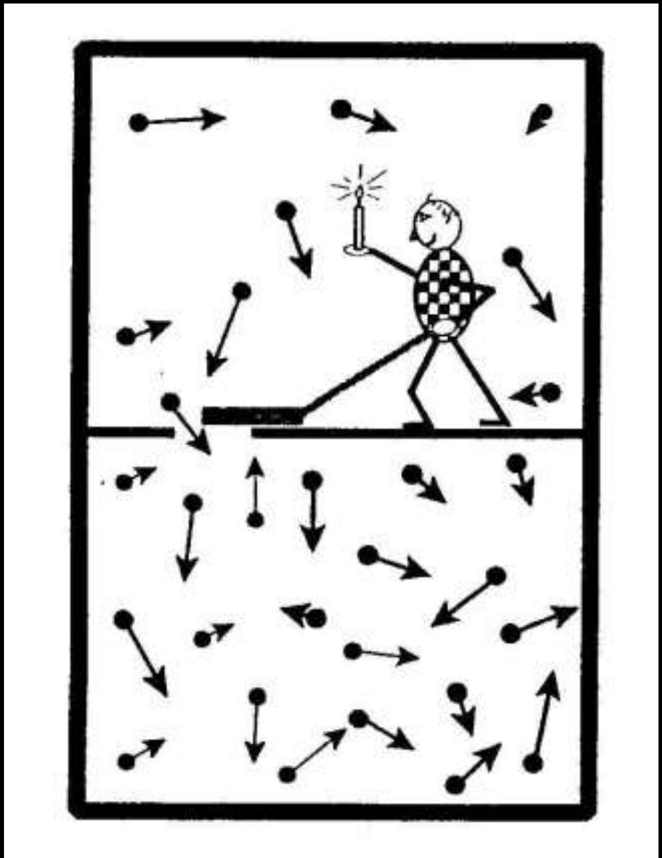
The amount of information actually stored is even lower, if the redundance in the storage of genetic information, which is required for the protection of the information, is taken into account. Estimates of information capacity in the genome vary between 3×10^2 bit and 10^{12} bit.

The *semantic* information, in contrast to the syntactic information should contain some kind of validation of the content. Despite many attempts, quantification of semantic information has not yet been achieved.

The interconnection of information with entropy in terms of thermodynamics may be illustrated best by a thought experiment conceived by James Clerk Maxwell in 1881 and still discussed today.

Maxwell proposed a room, which is connected with another by an opening. This opening can be closed by means of a slide. Both rooms are filled with a gas which is, in the beginning, in equilibrium, e.g. of equal pressure and temperature. An intelligent creature, called "Maxwell's demon", is able to handle the slide between the two rooms with ease. This "demon" can observe accurately the direction and the velocity of the molecules in his room.

Velocity and direction of these particles are statistically distributed. If a particle flies accidentally toward the opening, the demon opens the slide to let the particle pass. As a result of such sorting, the pressure in the lower room would rise.



The demon could also take another approach. For example, he could separate fast from slow particles.

In this case, a difference in the temperature between the two rooms would occur. In both cases, the entropy of the whole system would be reduced and energy might be generated by an appropriate turbine.

The result would be a "perpetuum mobile", a perpetual motion machine of the second order, as it would contradict the second principle of thermodynamics.

The problem may be resolved by the following consideration: the "demon" requires information to carry out the sorting. He collects this information by "watching" the molecules. In order to "see" he needs light. For this, the demon symbolically carries a lit candle. Yet, a body will only be able to emit light in a state of non-equilibrium relative to its environment.

This, however, contradicts the equilibrium condition at the beginning of the experiment. The same would apply to any other approach to acquisition of information. This resolves the apparent contradiction with the second law of thermodynamics.

Maxwell's demon demands particular interest in biophysics because of its analogy to various functions of living systems.

The living cell, too, reduces its entropy at the expense of its environment, using information processing. Yet, in this case it is not the energy of statistical fluctuations, which is used. The biological system selects molecules from its environment, which are rich in G, and correspondingly have a low content of entropy. It uses this energy and extrudes molecules with lower free energy and larger entropy.

The basic information for this process of selection is stored in the structure information of the proteins, which are responsible for the recognition of these molecules, and eventually for their metabolism.

These proteins get their information during the process of synthesis via the RNA, from the DNA of the genome.

This example raises the question: what is the threshold value of information, which is required to control the processes of living systems?

What is the threshold information that carries out not only the metabolic function of the primordial organism but additionally its replication? How did the first accumulation of information in a molecule occur ? A purely accidental combination of monomers to build their first functional macromolecule must be ruled out. The probability for this occurrence is too low by far. Today, so-called *prebiotic evolution* is assumed, i.e. chemical selection of developing polymers even before the stage of the biopolymer.