IT Basics
2

Faculty of Cybernetics, Statistics and Economic Informatics

BUES

Prof. Răzvan ZOTA, Ph.D.

zota@ase.ro

https://zota.ase.ro/itb

Information Theory - Introduction

Information Theory – the science dealing with the studies of the laws referring to obtaining, transmitting, processing and storing information.

Due to the random characteristic of the phenomena that represent the study of *Information Theory*, this can be considered a branch of the probabilistic theory from mathematics.

In order to transmit information it is necessary to **code it** in a manner, meaning its transmission using a language made of symbols and signals (the signals can be acoustic or light oscillations, electrical impulses or others).

A basic problem in Information Theory is the one of finding the most **efficient methods of coding**, finding the probabilities of transmitting information with a minimum number of symbols.

Information Theory - Introduction

Another typical problem in Information Theory is the one of determining the optimal number and capacity for the transmitting channels such that no distortions or losses would happen on the route from source to destination.

The entropy - a measure for the **degree of uncertainty** of a state of the physical systems.

Any information contains all the data about a certain physical system.

The information have no sense if the state of the physical system is completely determined.

Information Theory - Introduction

The interesting cases are the ones when a physical system may be randomly in a state from a set of possible states – this is about the situations when the system has a certain degree of uncertainty.

Moreover, the information regarding a physical system is as consistent and complete as the uncertainty of the system is greater till the information is revealed.

Here we have the question: **what should we understand by a high or low degree of uncertainty and how should we measure the uncertainty degree of a physical system?**

Information Theory - Introduction

Simple example:

1) Physical system: **the coin** (the result of flipping would put the coin in one of the two states: ***head*** or ***tail***)

2) Physical system: **the dice** (the result of rolling the dice may have the result of leaving the dice in one of the six possible states)

? In which case do we have a higher degree of uncertainty?

Information Theory - Introduction

Let us consider the following example:

1) Physical system: **the coin** (the result of flipping would put the coin in one of the two states: head or tail)

2) Physical system: **a computer** (in functioning state with the probability of 99% and in malfunction state with the probability of 1%)

? Only the number of the possible states is influencing the uncertainty degree of a physical system?

Information Theory - Introduction

A possible definition for information:

Information = a message which brings a new statement in a problem with some degree of uncertainty.

The uncertainty **is lowering** as the information appears (**grows**).

Being the experiment X, with probability distribution:

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}$$

The system of the events is considered **complete**.

Shannon’s formula

https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf

(original paper from 1948 of Claude Shannon)

Claude E. Shannon considered the following formula for the degree of uncertainty:

$$H(p_1, p_2, \ldots, p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

H is called ***entropy***

According to C. Shannon, the unity measure of the information is the bit.

One **bit** (**B**inary dig**IT**) is defined as the **quantity of information gained by choosing one value from two equally probable values.**

Coding example

What does it mean that a coded signal has 1.75 bits/symbol?

= we may convert the original signal in a row of 1 and 0 such that the **average** is 1.75 binary digits for each symbol from the original signal.

Assume that we have 4 symbols: A, B, C, D with the probabilities:

P_A=1/2; P_B=1/4; P_C=1/8; P_D=1/8

$-\log_2 P_A$ = 1 bit, $-\log_2 P_B$ = 2 bits, $-\log_2 P_C$ = 3 bits, $-\log_2 P_D$ = 3 bits

As the Shannon's formula, the uncertainty is:

$$H = \frac{1}{2} \bullet 1 + \frac{1}{4} \bullet 2 + \frac{1}{8} \bullet 3 + \frac{1}{8} \bullet 3 = 1{,}75\ bits$$

Coding example (cont.)

If we use the binary representation for the symbols A,B,C,D:

A = 1; B = 01; C = 000; D = 001, then ABADCAAB will be coded as:

10110010001101

(14 binary digits used for coding the 8 symbols => the average is 14/8 = 1.75)

Obs. What is happening if we use the following coding:

A = 00; B = 01; C = 10; D = 11 ?

Entropy's properties

$$P1.\ H(p_1, p_2, ..., p_n) \geq 0$$

$$P2.\ H(p_1, p_2, ..., p_n) = 0 \Leftrightarrow \ni k \in \{1, 2 ... n\}\ a.i.\ p_k = 1$$

$$P3.\ H(p_1, p_2, ..., p_n) \leq H(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n})$$

Entropy's properties (cont.)

$$P4.\ H(p_1, p_2, \dots, p_n, 0) = H(p_1, p_2, \dots, p_n)$$

$$P5.\ H(X_1 \times X_2 \times \dots \times X_n) = H(X_1) + H(X_2) + \dots + H(X_n)$$

$$P6.\ H(X \times Y) = H(X) + H(Y/X), \forall X, Y$$

$$where: \quad H(Y/X) = \sum_{k=1}^{n} p(x_k) \cdot H(Y/x_k)$$

$$H(Y/x_k) = -\sum_{i=1}^{m} p(y_i/x_k) \log_2 p(y_i/x_k)$$

Entropy's properties (cont.)

$$P7.\ H(Y/X) \le H(Y), \forall X, Y$$

$$P8.\ H(X \times Y) \le H(X) + H(Y)$$

$$P9.\ H(X/Y) = H(Y/X) + H(X) - H(Y)$$

Arithmetic basis of the computers

http://www.ma.utexas.edu/users/mks/326K04/what.html

Number system = a collection of numbers together with operations, properties of the operations and a system representing these numbers. Collection of representation rules by using symbols (*digits*).

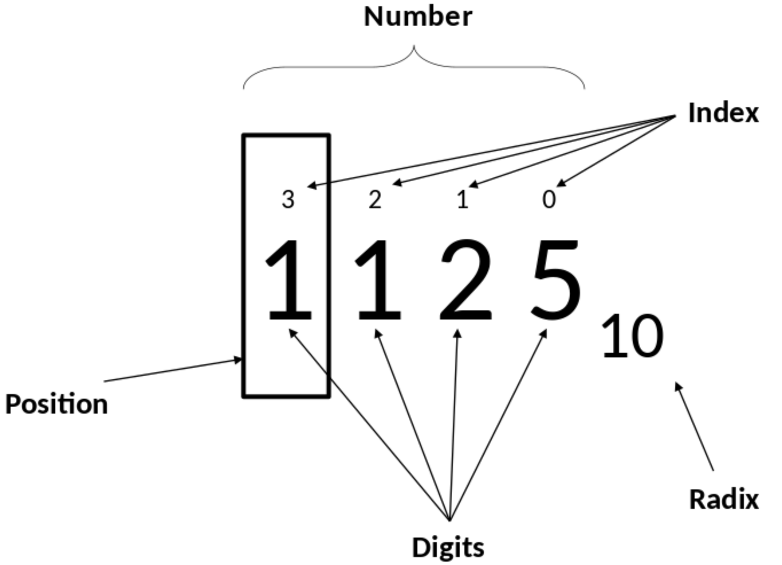
The number of allowed symbols is the base (radix) of the number system

Number systems

- positional
- non-positional (the Roman system)

Arithmetic basis of the computers

Positional number systems



Arithmetic basis of the computers

The roman system

I	X	C	M	V	L	D
1	10	100	1000	5	50	500

In the case when **a smaller value is positioned after a symbol with a bigger value**, the values are **added**.

III	→	1 + 1 + 1	=	3
VIII	→	5 + 1 + 1 + 1	=	8
XVIII	→	10 + 5 + 1 + 1 + 1	=	18
LXXII	→	50 + 10 + 10 + 1 + 1	=	72
CI	→	100 + 1	=	101
MMVII	→	1000 + 1000 + 5 + 1 + 1	=	2007
MDC	→	1000 + 500 + 100	=	1600

Arithmetic basis of the computers

In the case when a symbol with a smaller value is positioned before a symbol with a bigger value, the smaller value is subtracted from the other value.

IV	→	5 − 1	=	4
XIX	→	10 + (10 − 1)	=	19

Representing a number in a base

- Integer representation

$$N = \overline{a_n a_{n-1} \cdots a_0}$$

$$N = a_n \bullet b^n + a_{n-1} \bullet b^{n-1} + \cdots + a_0 \bullet b^0$$

This is the representation of integer N in base b.
The digits of number N have the following property:

$$0 \le a_i \le b-1,\ \forall i \in \overline{0, n}$$

Real number representation

- Real number R representation:

$$R = \overline{a_n a_{n-1} \cdots a_0 a_{-1} \cdots a_{-m}}$$

$$R = a_n \bullet b^n + a_{n-1} \bullet b^{n-1} + \cdots + a_0 \bullet b^0 + a_{-1} \bullet b^{-1} + \cdots + a_{-m} \bullet b^{-m}$$

This is the representation of a real number R in base b. The digits of R have the following property:

$$0 \le a_i \le b-1, \forall i \in \overline{-m, n}$$

there is no k such as $a_k = a_{k-1} = a_{k-2} = \cdots = b-1$

Base conversion

- Base conversion (integer part, fractional part)
- Quick conversion between numbers represented in bases that have the following relation:

$$b_1 = b_2^p, \quad p \in \aleph$$