The Management of Indeterminate Measurement Data

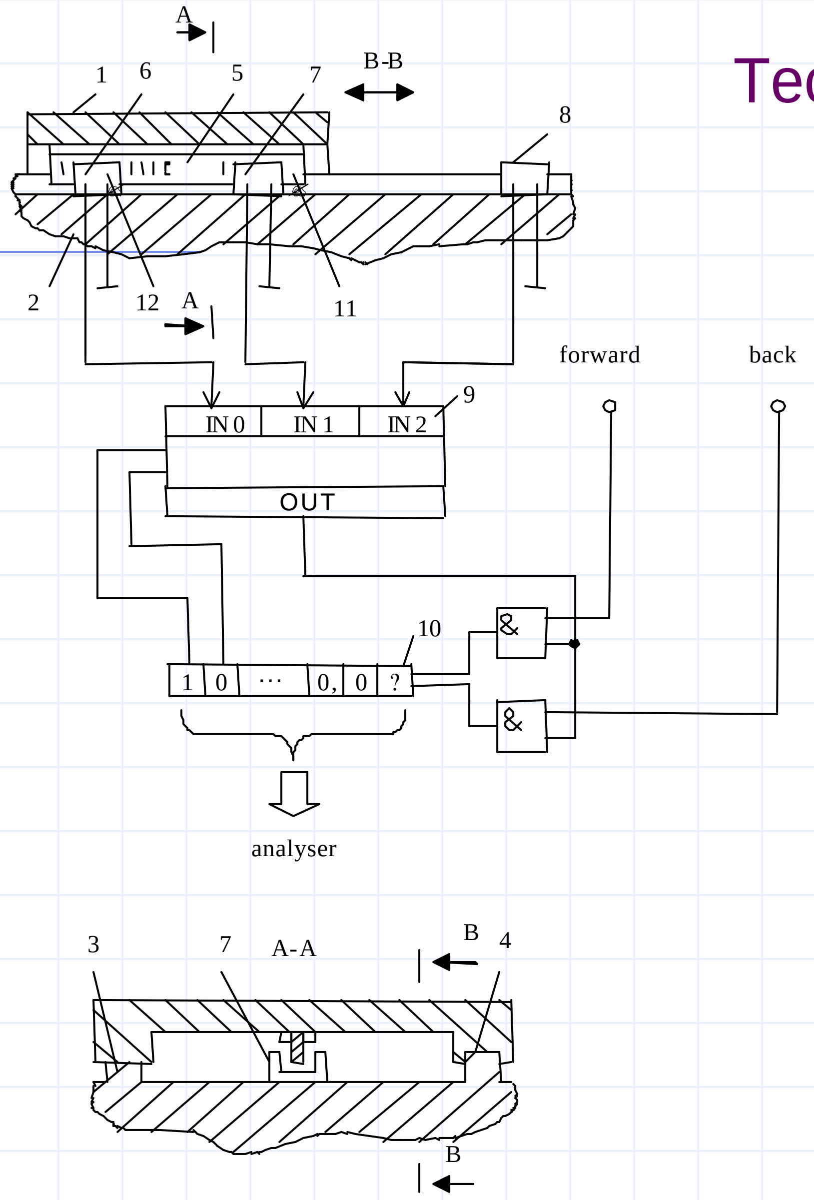
Mindaugas Rybokas

Vilnius Gediminas Technical University,
Vilnius, Lithuania

Introduction

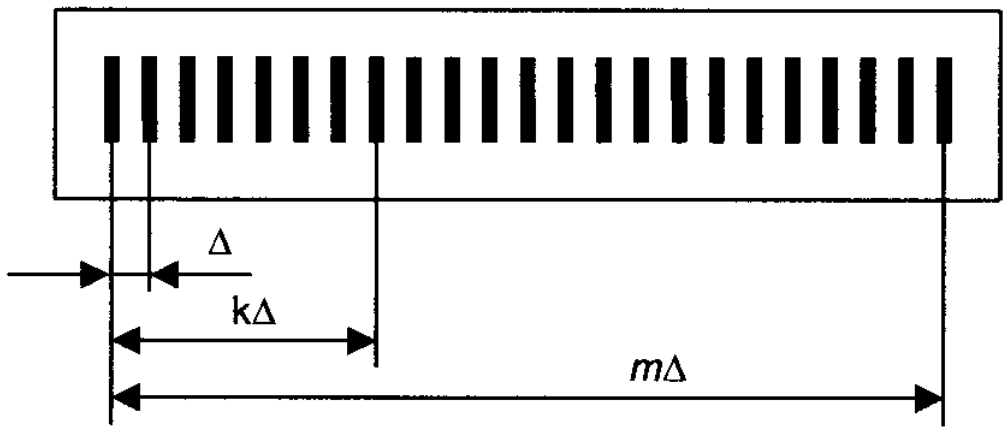
- Large data set evaluation is strongly connected with information theory which is considered by many a subset of communication theory.
- Nevertheless, very fast development of the optical, radio, GPS, electronic and mechatronic means of sensing, data collection, transmitting and evaluating permits to consider a wide range of application of those means to substitute mechanical measuring instruments and machines to non-contact measurements, such as optical, radio, ultrasound measurements.
- The quality control of machine and equipment parts is strongly connected with the use of automatic coordinate measuring machines (CMM) and large data sets evaluation.

Technical Implementation



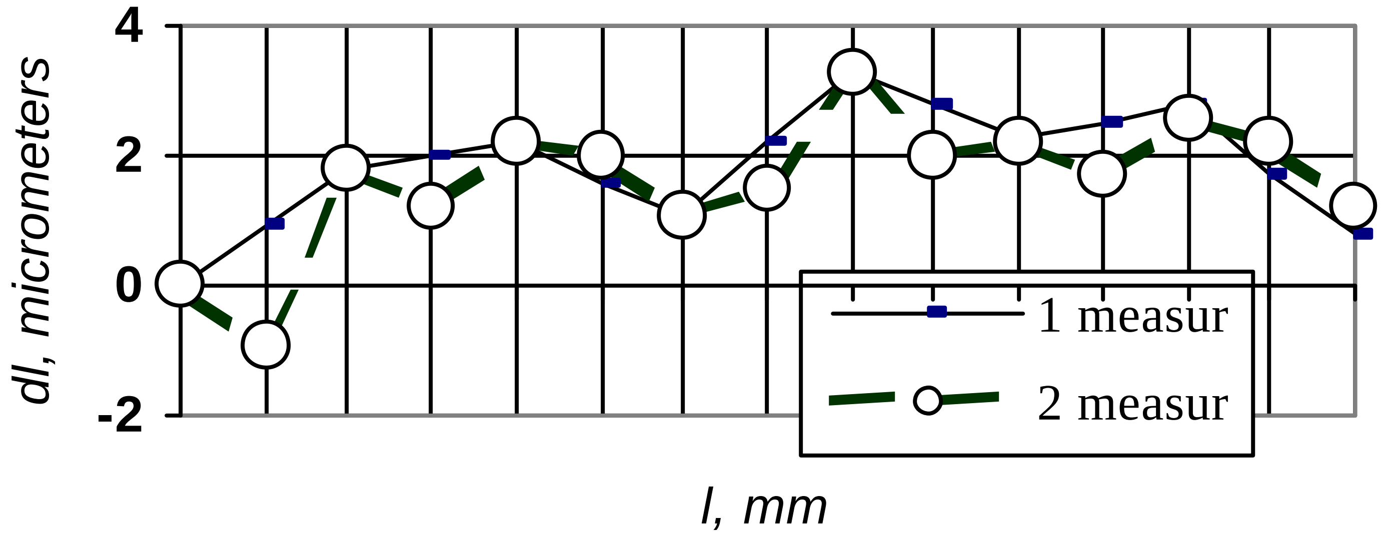
Arrangement of information – measuring system of the machine using a smaller part of the linear transducer for measuring the full stroke of the machine

Raster Scale



Raster scale having the length *l=m?*, the pitch of the scale *?* and the pitch of the calibration *k?*.

Influence of Pitch on Measurement Results



Results of measurement of the linear transducer during its calibration at the pitch of every 2 mm (the 1st measurement) and at every 1 mm (the 2nd measurement)

Information Uncertainty vs. Indeterminacy

The basic definitions used in this presentation:

information uncertainty and ***information indeterminacy***. As for uncertainty is widely used an expression for probability of data taking account of random data with some value of S as the estimate of s (in measurements s is taken usually as 1s = 0.68, 2s = 0.95 and 3s = 0.995).

By ***indeterminacy*** we mean that the data have not been checked, tested, calibrated (in measurement), validated to some extent by reference measure (standard measure).

It is conceived that the both terms compose (***information uncertainty*** and ***information indeterminacy)*** the other sometimes used expression "inexactness".

Entropy and Information

The entropy *H(X)* in bits of information message is:

$$H(X) = -\sum_{i=1}^{n} p_i(x)\log_a p_i(x),$$

where *p(x)* is a probability mass function of a random variable *X*, and *a* – the base of logarithms, n – number of outcomes. In case of a uniform distribution entropy will be: $H(X) = -\sum_{i=1}^{n} \frac{1}{n}\log_a \frac{1}{n} = \log_a n$,

where *n* – the number of measurements.

Entropy is the uncertainty of a single random variable. The reduction ***I*** in uncertainty due to the information assessed (in our case – information received after the calibration) is:

$I = H_0 - H_1$,

where H_0 is the entropy before receiving the information , H_1 – after receiving it.

Measurement Data Investigation

For each measurement a large data set is used to construct the final measurement result expression (of part dimensions, machine movement magnitude):

$$X \ ? \ \bar{x} \ ? \ ?,? \quad ,$$

where X – measurement result;

$\bar{x}$ - the systematic component of measurement result,

? - measurement uncertainty, expressed as $? \ ? \ \frac{tS}{\sqrt{n}}$, where t – Student coefficient, S – the estimate of mean standard deviation, P- confidence interval (0,68; 0,95; 0,99).

$$H_0 \;?\; ? \underset{?}{?} \log_a \frac{1}{m} \;?\; \log_a m \,.$$

The information received after the calibration, i.e., determination of accuracy of part of the scale will be $H_1 \;?\; \log_a b$,

where $b \;?\; \dfrac{m}{k}$ is the number of calibrated strokes in the scale. These strokes were measured $\boldsymbol{c}$ times each for the statistical evaluation. Then the reduction in information uncertainty (indeterminacy) due to information received will be:

$$I \;?\; H_0 \;?\; H_1 \;?\; \log_a m \;?\; \log_a b \,;$$

then $\log_a b \;?\; \log_a m \;?\; I$; and $b \;?\; a^{(\log_a m \,?\, I)} \;?\; m \;?\; a^{?I}$.

Since the total number of measurements is ***n* = *bc*** (each calibration measurement is performed ***c*** times), the expression of measurement result (3) at given the probability and the reduction in information indeterminacy becomes:

$$X \; ? \; \overline{X} \; ? \; \frac{t\,?S}{\sqrt{a^{?\,I} mc}} \;, P,\; I(H_0, H_1) \,.$$

It means that the measurement result is determined with the ***uncertainty*** assessed by probability level *P* and with the ***indeterminacy*** of the result assessed by the entropy $I(H_1, H_0)$ of evaluation of that part of all data in question.

Conclusions

- The new approach to the measurement data evaluation gives a full information of the measurement process performed and the quantity of data assessed during this process.
- While investigating the information measurement systems of machines there has been observed that large systemic errors accumulate at the length of the measuring distance. Using the displacement measurement system offered the absolute systemic error can be limited to a desired interval by applying the minimisation of information entropy of the scale.

???