Program of the course "Information theory" in autumn 2020 (lectures 1-5 and 11-13).

1. The game "guess a number" : non-adaptive and adaptive strategies; upper and lower bounds for the required number of operations in the worst case.
2. Sorting algorithms: un upper bound (the Merge Sort algorithm) and the information-theoretic lower bound.
3. "Fake coin problem" : search for an object with a wrong mass with balance scales: upper and lower bounds for the number of operations.
4. Combinatorial definition of the information quantity by R. Hartley; the basic properties of Hartley's information.
5. The game "guess a number" with a probability distribution on the set of answers: un upper bound for the required number of operations on average.
6. Shannon's entropy of a random variable. The bounds $0 \leq H(\alpha) \leq \log n$ for a random variable α with n values.
7. The inequality $H(\alpha, \beta) \leq H(\alpha) + H(\beta)$ for a pair of jointly distributed random variables (α, β).
8. Shannon's entropy provides a lower bound for the average number of operations in the game "guess a number."
9. Prefix codes and uniquely decodable codes. Kraft's inequality for the binary uniquely decodable codes. For every uniquely decodable code there is an equivalent prefix code.
10. Shannon's entropy provides a lower bound for the average length of a uniquely decodable code for a given distribution of probabilities on the set of messages.
11. Huffman's coding, its optinality
12. A simplified version of Stirling's formula: there exist constants $c_1, c_2 > 0$ such that for all integer numbers $N > 0$

$$c_1\sqrt{N}\left(\frac{N}{e}\right)^N \leq N! \leq c_2\sqrt{N}\left(\frac{N}{e}\right)^N.$$

13. Block coding for typical sequences: the n-letter words $x \in \{a_1, \ldots, a_k\}^n$ with frequencies of letter $p_1, \ldots, p_k$ can be represented by binary strings of length
$$\left(\sum_{i=1}^{k} p_i \log \frac{1}{p_i}\right) n + o(n).$$

14. *Expectation* and *variance* of real-valued random variables. The basic properties and the Chebyshev inequality.

15. Block coding for random sequences: let α_i, $i = 1, 2, \ldots$ be a sequence of independent identically distributed random variables; then for every $\varepsilon > 0$, all values of $(\alpha_1, \ldots, \alpha_n)$ (except for a set of values with a total probability smaller than ε) can can be represented by binary strings of length
$$\left(\sum_{i=1}^{k} p_i \log \frac{1}{p_i}\right) n + o(n),$$
where $(p_1, \ldots, p_k)$ is a distribution of probabilities on the values of each α_i.

16. Conditional Shannon's entropy $H(\alpha|\beta)$; the definition and basic properties.

17. Mutual information $I(\alpha : \beta)$ in the sense of Shannon's information theory: the definition and basic properties. Non-negativity and symmetry.

18. Conditional version of the mutual information $I(\alpha : \beta|\gamma)$ in the sense of Shannon's information theory. Equivalent representations and non-negativity.

19. The fundamental relations between different entropic quantities for pairs and triples of jointly distributed random variables. Venn-like diagrams for the standard information quantities.

20. Examples of non-basic information inequalities. A proof of
$$2H(\alpha, \beta, \gamma) \leq H(\alpha, \beta) + H(\alpha, \gamma) + H(\beta, \gamma).$$

21. One-time pad encryption scheme. Security of the Vernam cipher. Optimality: in every secure scheme the size of the secret key is not smaller than Shannon's entropy of the message.

22. Secret sharing: the secret sharing scheme of Shamir; a proof of its security.

23. The existence of a decompressor that is optimal (up to an additive constant) for the simple and for the conditional algorithmic complexity. The definition of Kolmogorov complexity.

24. Basic properties of Kolmogorov complexity. The existence of incompressible binary strings for each length n.

25. Kolmogorov complexity and Shannon's entropy: there exist constants d_1, d_2 such that for all binary strings x of length n with pn zeros and $(1-p)n$ ones we have

$$C(x) \leq \left(p \log \frac{1}{p} + (1-p) \frac{1}{1-p} \right) n + d_1 \log n + d_2.$$

26. Mutual information $I(x:y)$ in the sense of Kolmogorov complexity. The Kolmogorov–Levin theorem (without a proof):

    (a) $C(xy) = C(x) + C(y|x) + O(\log |x| + |y|)$;

    (b) $I(x:y) = I(y:x) + O(\log |x| + |y|)$.