Seminar Information and Coding Theory

Robert König robert.koenig@tum.de
Time & Location: TBD

Guidelines

- The presentation of each topic should ideally last less than 50 minutes. The talks should be given in English.
- Presentations should be given predominantly on the blackboard but you may use a projector for showing, e.g., pictures or graphs. You may also use slides e.g., for reviewing materials as long as the total time doing so does not exceed 15–20 minutes.
- The target audience are the other students attending the seminar. Try to make sure it is understandable for everyone.
- You are asked to "pair up". Each pair of students will be assigned two of the topics below, and each student is asked to give a non-negligible fraction of the associated presentation. If there is an odd number of participant, one student will be assigned a single (more advanced) topic to present.
- You are encouraged to submit a summary (at most 2 pages) of the assigned topic, preferably at least one week before your talk. This will be corrected and/or provided as a handout for the other students.
- You can arrange a meeting with me or one of my assistants prior to the presentation. This is intended to help identify key concepts to be presented, address specific technical questions, and to make sure your presentation is ready for public consumption. If possible, these meetings should organized topic-wise (in pairs).
- Attendance at all talks except in justified circumstances will be required to receive credits.

We recommend having a look at Prof. Dr. Manfred Lehn's website `http://www.alt.mathematik.uni-mainz.de/Members/lehn/le/seminarvortrag` (in German), which gives advice on how to prepare a good seminar talk.

Topics

This seminar will serve as an introduction to the theory of information and coding. Topics to be discussed include information-theoretic quantities (entropies) and corresponding inequalities, Shannon's channel coding theorem, data compression and basic concrete coding strategies. Some alternative uses of information theory e.g., to noise-tolerant computation will also be discussed.

The following concepts should be reviewed prior to the seminar:

> discrete and continuous random variables, probability density, conditional distributions, Bayes rule, independence, expectation value and variance, Markov's inequality/Chebyshev's inequality, weak law of large numbers and tail bounds (Chernoff) for i.i.d. random variables.

We will discuss the following topics. For each topic in the following list, we give a number of keywords as suggested concepts to be covered in your talk. While these should all be covered in some form, it is up to you to make a reasonable selection of subtopics which you discuss in more detail. Please try to include at least one full proof of a result, and otherwise try to convey the main ideas/concepts.

Some suggested references are given; if you use complementary materials, please try to make sure to follow similar notation as in "standard" textbooks such as [1], [3] and [5].

1. **Fundamentals**: Entropy, Markov chains/discrete memoryless channels (DMC), Fannes' inequality. Examples of noisy channels (binary symmetric and binary erasure channel). Relative entropy and mutual information. [1, Chapter 2].

2. **Mutual information and capacity-cost function:** Chain rule, data processing inequality. Capacity-cost function for a DMC and its additivity. Statement and (weak) converse to channel coding theorem.[5, Chapter 2], [1, Chapter 7.9]

3. **Achievability of capacity:** Maximum likelihood decoding, random codebook generation, achievability of capacity for the binary symmetric channel. Statement and proof sketch of general channel coding theorem. [1, Chapter 7.7] [5, Chapter 2].

4. **Data compression**: Asymptotic equipartition property (AEP), data compression. Lossless compression: Kraft's inequality, bounds on optimal codes, Huffman codes. [1, Chapter 3 and 5]

5. **Codes for communication**: rate & code distance, linear codes, generator- and parity check matrix, Hamming code and Hamming bound. Syndrome/Maximum likelihood decoding decoding, Weight enumerators and MacWilliams identities. [5, Chapter 7]

6. **Code concatenation**: Singleton-bound, Gilbert-Varshamov bound, code concatenation, Zyablov bound, polynomial time construction of Zyablov-bound achieving codes, decoding of concatenated codes. [3]

7. **Concrete codes and decoders**: BCH codes, Reed-Solomon codes and decoders, the Berlekamp-Welch algorithm. [4]

8. **Expander codes**: Tanner graphs, expander graphs, greedy algorithm for decoding. [9]

9. **Fault-tolerant computation:** upper bounds (von Neumann), signal propagation and depth lower bound on noisy formulas. [2, 7]

References

[1] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, Wiley-Interscience; 2nd edition (2006).

[2] W. Evans and L. J. Schulman, *Signal propagation, with application to a lower bound on the depth of noisy formulas*, Proceedings of FOCS 1993, available here.

[3] V. Guruswami, A. Rudra and M. Sudan, *Essential Coding Theory*, 2018. Available here.

[4] L. Guth, The Polynomial Method, online lecture notes, Chapter 2, available here.

[5] Robert J. McEliece, *The Theory of Information and Coding*, Addison-Wesley, 2nd edition (1977)

[6] David MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press; First Edition edition (2003). Available at `http://www.inference.phy.cam.ac.uk/mackay/itila/book.html`.

[7] Nicholas Pippenger, *Reliable Computation by Formulas in the Presence of Noise*, IEEE Trans. Inf. Th. vol 34, no. 2, March 1988, available here.

[8] Claude E. Shannon, *A Mathematical Theory of Communication*, Bell System Tech. J. 27, (1948). 379–423, 623–656. Available at `http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf`.

[9] Michael Sipser and Daniel Spielman, *Expander Codes*, IEEE Transactions on Information Theory, vol. 42, number 6 1996. available at here

Note: You may have to follow the Proxy setup instructions to access some of these materials through the university's network.