University of Washington

Abstract

List Decoding and Property Testing of Error Correcting Codes

Atri Rudra

Chair of the Supervisory Committee:
Associate Professor Venkatesan Guruswami
Department of Computer Science and Engineering

Error correcting codes systematically introduce redundancy into data so that the original information can be recovered when parts of the redundant data are corrupted. Error correcting codes are used ubiquitously in communication and data storage.

The process of recovering the original information from corrupted data is called decoding. Given the limitations imposed by the amount of redundancy used by the error correcting code, an ideal decoder should efficiently recover from as many errors as information-theoretically possible. In this thesis, we consider two relaxations of the usual decoding procedure: *list decoding* and *property testing*.

A list decoding algorithm is allowed to output a small list of possibilities for the original information that could result in the given corrupted data. This relaxation allows for efficient correction of significantly more errors than what is possible through usual decoding procedure which is always constrained to output the transmitted information.

- We present the first explicit error correcting codes along with efficient list-decoding algorithms that can correct a number of errors that approaches the information-theoretic limit. This meets one of the central challenges in the theory of error correcting codes.

- We also present explicit codes defined over smaller symbols that can correct significantly more errors using efficient list-decoding algorithms than existing codes, while using the same amount of redundancy.

- We prove that an existing algorithm for a specific code family called Reed-Solomon codes is optimal for "list recovery," a generalization of list decoding.

Property testing of error correcting codes entails "spot checking" corrupted data to quickly determine if the data is very corrupted or has few errors. Such spot checkers are closely related to the beautiful theory of Probabilistically Checkable Proofs.

- We present spot checkers that only access a nearly optimal number of data symbols for an important family of codes called Reed-Muller codes. Our results are the first for certain classes of such codes.

- We define a generalization of the "usual" testers for error correcting codes by endowing them with the very natural property of "tolerance," which allows slightly corrupted data to pass the test.