Università degli Studi di Palermo

Dipartimento di Matematica e Informatica

Words and Automata Research Group

SEMINAR ANNOUNCEMENT

Title: Bit-optimal Lempel-Ziv compression

Rossano Venturini, University of Pisa

Thursday 14th November 2013, 3 p.m.
Room 7
Via Archirafi 34, 90123 Palermo

Abstract:

One of the most famous and investigated lossless data-compression schemes is the one introduced by Lempel and Ziv about 35 years ago. This compression scheme is known as ``dictionary-based compressor'' and consists of squeezing an input string by replacing some of its substrings with (shorter) codewords which are actually pointers to a dictionary of phrases built as the string is processed. This strategy is the core of the most famous compressors, like gzip, pkzip, lzma, 7zip, rar, lz4, snappy, and so on. Surprisingly enough, although many fundamental results are nowadays known about the speed and effectiveness of this compression process, we are not aware of any parsing scheme that achieves optimality under any constraint on the codewords other than being of equal length. Here optimality means to achieve the minimum number of bits in compressing each input string, without any assumption on its generating source.

In this talk we will address some issues pertaining to the bit-complexity of LZ77-based compressors, the most powerful variant of the LZ-compression scheme, and we will present algorithms which achieve bit-optimality in the compressed output size by taking efficient/optimal time and optimal space. We will also address the problem of optimally trading the compressed size/decompression time of LZ77 parsings by introducing the Bicriteria LZ77-Parsing problem. The goal is to determine an LZ77 parsing which minimizes the space occupancy in bits of the compressed file, provided that the decompression time is bounded by a user defined parameter T. Experiments on real texts will motivate the practical interest of the proposed solutions.

All interested people, in particular students, are invited to participate.

For further information: Gabriele Fici, Dip. di Matematica e Informatica, Via Archirafi 34, 90123 Palermo gabriele.fici@unipa.it