A Machine Learning Perspective on Predictive Coding with PAQ

Presentation by Phil Trommer

November 7, 2019

Overview

Introduction to PAQ

Introduction to PAQ

What is PAQ8

- What is it?
- How does it work?
- What makes it so famous?

Introduction to PAQ

Matt Mahoney

- Born 1955
- Recieved Ph.D in computer science at Florida Tech in 2003
- Released PAQ1 on January 6, 2002

Introduction to PAQ

Matt Mahoney

- Born 1955
- Recieved Ph.D in computer science at Florida Tech in 2003
- Released PAQ1 on January 6, 2002

What is PAQ?

- A lossless, open-source compression algorithm
- Brings high perfomance at the cost of increased memory usage and time consumption
- Related to PPM, is envisioned as PPMs improvement

Introduction to PAQ

Principles of PAQ

- Modeling combined with adaptive arithmetic encoding
- Open to additions and improvements
- Improves perfomance of PPM by including several predictors (i.e. models of data)
- Combines the result of the predictors

Introduction to PAQ

Exemplary Predictors

Introduction to PAQ

Exemplary Predictors

The order-n context predictor

- Examines the last n bits and counts the 1's and 0's
- Estimates probability whether next bit is 1 or 0 like PPM

Introduction to PAQ

Exemplary Predictors

The order-n context predictor

- Examines the last n bits and counts the 1's and 0's
- Estimates probability whether next bit is 1 or 0 like PPM

The sparse context predictor

- Context consists of a specific amount of non-contiguous bytes before the current bit

PAQ & Predictors

- PAQ encoder looks at the beginning of input file for deciding which predictors are used
- Ways to combine predictions change through with the different versions

PAQ8L

PAQ8L

PAQ8 - What's new?

- Predictors don't produce a pair of bit counts anymore
  $\hookrightarrow$ those counts get weighted and normalized into the interval $[0, 1] \subset \mathbb{R}$
- Instead each predictor already outputs a probability
- *paq8l* is a stable version of paq8, released by Matt Mahoney

PAQ8L

PAQ8 - What's new?

- Predictors don't produce a pair of bit counts anymore
  $\hookrightarrow$ those counts get weighted and normalized into the interval $[0, 1] \subset \mathbb{R}$
- Instead each predictor already outputs a probability
- *paq8l* is a stable version of paq8, released by Matt Mahoney

PAQ8L - Machine Learning Perspective

- paq8l is the version of PAQ used by *Byron Knoll & Nando de Freitas*
- They try to show the possibilities of PAQ beyond data compression

Applications for PAQ8

References