

夏休み情報科学シリーズ「はじめての情報科学」第 2 弾

情報科学メジャーの学生はもちろんのこと、それ以外のメジャーの学生、まだメジャーを決めていない 1-2 年生の視聴も大歓迎です。教員の皆さんもふるってご視聴ください。言語は英語です。

日 時: July 26th Monday 9:30-10:40

場 所: <https://icu.zoom.us/j/98375967288>

Meeting ID: 983 7596 7288 PW: 446761

題 目: An Introduction to Data Compression

演 者: Dr. Marie-Luce Bourguet

Senior Lecturer (Associate Professor), School of Electronic Engineering and Computer Science, Queen Mary University of London

データ圧縮は情報科学の最も重要な分野の 1 つであり、様々な場所で利用されています。デジタル世界におけるデータの爆発的な増加に伴い、データを効率的に保存・送信するための技術の必要性が高まりました。たとえば、Web 経由で取得し、コンピューターまたはスマートフォンに保存されるすべての画像や動画は圧縮されています。

データ圧縮とはそもそもどのような技術でしょうか?これはシステムティックなデータの削除と、その後に行われる冗長性を活用したデータの再構成、あるいは重要性が低いデータの削除でもあります。他方で、データ圧縮はコンピューターにおいてのみ実現されるものではなく、人間の脳において、感覚から絶えず得られる膨大な量の情報の保存と利用を可能にするための処理として実施されるものです。本レクチャーでは、ヒューマンメモリ、自然言語処理、視覚認知から例を挙げ、データ圧縮について説明します。

コンピューターの処理においては、データはデジタル形式(ビット)で表現されるため、コンピューターにおけるデジタルデータの圧縮とは、ビットの一部を削除することです。それでは何故データ圧縮が必要なのか?一層増え続けるデジタルデータを日々作成、加工、および共有するに際し、コンピューターによるデータ圧縮技術は不可欠なものとなりつつあります。たとえば、最新のストレージおよびネットワークテクノロジーはすべて、データ圧縮に依存していま

す。データ圧縮はまた、電話会議やビデオオンデマンド等のアプリケーションにおいても不可欠な技術となっています。

次にデータ圧縮がどのように機能するか?についてですが、データ圧縮は「可逆」、「非可逆」いずれの場合においても、データによって表現される情報量を可能な限り保持しつつも、冗長なデータを削除します。本レクチャーでは、データの冗長性に関する手法とデータ圧縮手法について説明します。例えば、RLE は画像の空間的な冗長性を活用したロス・レス圧縮技術であり、Huffman は画像データの統計的冗長性を活用します。Jpeg と mp3 は不可逆であり、知覚の冗長性を利用してそれぞれ画像と音楽データを圧縮します。また mpeg は画像の不可逆的な圧縮技術であり、画像データの時間的冗長性を利用しています。

NS forum の Moodle があります。授業のコメントシートはこちらに投稿してください。また大学院生の方はコメントシートの提出で出席とします。

<https://moodle3.icu.ac.jp/course/view.php?id=3836>



Summer Information Science Series

“Information Science for beginners” Vol. 2

Not only students of the Information Science major, but also students of other majors and first and second year students who have not yet decided on a major are welcome to watch. Faculty members are also welcome to attend. The language is English.

Time: July 26th Monday 9:30–10:40

Place: <https://icu.zoom.us/j/98375967288>

Meeting ID: 983 7596 7288 PW: 446761

Title: An Introduction to Data Compression

Speaker: Dr. Marie-Luce Bourguet

Senior Lecturer (Associate Professor), School of Electronic Engineering and Computer Science, Queen Mary University of London

Data compression is used just about everywhere. It is one of the most important area of information science. Explosive growth of data in the digital world leads to the requirement of efficient techniques to store and transmit data. For example, all the images and videos you get on the web and store in your computer or phone are compressed.

So, What is Data Compression? It is the systematic removal and, later, reintroduction of redundancy from data, and possibly also the removal of insignificant data. Data compression has not been invented with the computer though. In the brain, compression is what makes possible the storage and use of the enormous quantity of information that is constantly coming through the senses. In the lecture, we will discuss examples from human memory, natural language, and visual perception. In computers, data is represented in digital format (bits), and so for computers, digital data compression is about removing some of the bits.

Why is digital data compression needed? Computers use compression to handle the ever-increasing amount of digital data created, processed, and shared everyday by

computer users. For example, all the latest storage and network technologies rely on data compression. Data compression is what enables every-day applications such as teleconferencing and video-on-demand.

How does compression work? Compression can either be lossless or lossy, but in both cases, it strives to remove redundant data while preserving as much as possible the information that is represented by the data. In the lecture, we will discuss several types of data redundancy and techniques used to compress data. For example, RLE is a lossless compression technique, which exploits spatial redundancy in images, whereas Huffman exploits statistical redundancy to compress images. Jpeg and mp3 are lossy and exploit perceptual redundancy to compress image and music data respectively. Finally, mpeg is a lossy compression technique for videos, which exploits temporal redundancy in video data.

There is a Moodle for the NS forum. Please post comment sheets for the class here. If you are a graduate student, your attendance will be counted by submitting the comment sheet.

<https://moodle3.icu.ac.jp/course/view.php?id=3836>
