

連載

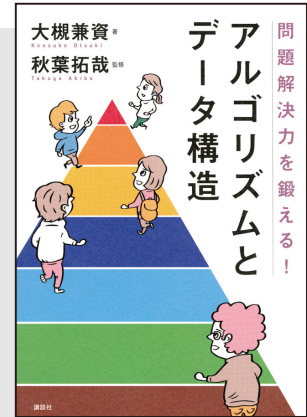
ビブリオ・トーク
- 書評 -

… 石井一夫 (久留米大学)

問題解決力を鍛える! アルゴリズムとデータ構造

大槻兼資 著, 秋葉拓哉 監修

講談社 (2020), 3,300 円 (税込), 368p., ISBN: 978-4-06-512844-2



アルゴリズムとは

アルゴリズムとは、『問題を解くための方法、手順』のことです」と本書の最初に述べられている。Wikipedia 的には、『計算可能』なことを計算する、形式的な (formal な) 手続きのこと、あるいはそれを形式的に表現したもの」と述べられている。本書では、アルゴリズムの動作を説明するだけでなく、「どうしたらよいアルゴリズムを設計できるか」をメインに捉えて述べられている。

データ分析とアルゴリズム

データ分析の観点から言えば、アルゴリズムとは、データ分析の際のデータのフロー形式を記述したものと言えると思う。たとえば、本書で紹介されているアルゴリズムのうち、線形探索法はデータ分析という総当たり (ブルートフォース) 分析法だし、その中の1つの手法として再帰アルゴリズムが含まれる。また、分割統治法はデータ分析という並列処理だし、動的計画法も並列処理を含んでいる。配列、連結リスト、ハッシュテーブルなど、データ構造の概念はデータ分析に必須だ。グラフ構造、木構造、ネットワークフローの考え方、アルゴリズムも機械学習で出てくる。ディープラーニングそのものが、ニューラルネットワークと呼ばれるグラフ構造である。このように考えると、データ分析を深く学び使いこなすためには、アルゴリズムやデータ構造を学ぶことが重要であり、それが問題解決力を身につけることになる。

離散数学とアルゴリズム

興味深いのは、本書の章立てが、前回のビブリオ・トークで紹介した『情報系のための離散数学 (共立出版)』¹⁾ と非常に似ていることである。『情報系のための離散数学』では、前半に命題、論理式、集合、数学的帰納法、証明技法について述べ、後半で、グラフ理論やネットワークの解説をしている。それに対し、本書は3部構成で、最初にアルゴリズムの概要と計算量の評価方法を説明した後に、第1部として線形探索法、分割統治法、動的計画法などのアルゴリズム設計技法が説明される。その後、第2部として配列、連結リスト、ハッシュテーブル、グラフ構造、木構造、ネットワークなどのデータ構造が語られ、最後に第3部としてグラフ構造に基づいたアルゴリズムが紹介される。冒頭にある監修者のことばにも「離散数学の深遠な理論が広がっており」というコメントがあるので、著者や監修者らは離散数学とアルゴリズムの類似性を意識していると思われる。見ようによっては、離散数学をコンピュータ上で実装するための技法が、アルゴリズムという捉え方もできる。

本書の構成

本書の章立てであるが、18章の構成である。

1章「アルゴリズムとは」では、アルゴリズムとは何かについて述べ、アルゴリズムの例として「線形探索法」と「二分探索法」を、そして、「深さ優先探索」と「幅優先探索」を最初に分かりやすい事例をまじえて紹介しているのが印象的である。アルゴリズムの記

述方法としてC++のコードが使用されており、これをもとに実際にアルゴリズムを実装できるようになっている。これは、実行できない疑似コードよりは実感しやすい。

2章「計算量とオーダー記法」では、アルゴリズムを評価する指標として「計算量」と「計算量のオーダー記法」について紹介している。

3章から7章にかけては、アルゴリズムの具体的な設計技法が説明されており、3章で「全探索：線形探索法など」、4章で「再帰と分割統治法」、5章で「動的計画法」、6章で「二分探索法」、7章で「貪欲法」が紹介されている。

8章から11章にかけては、アルゴリズムに使われるデータ構造が説明されており、8章で「配列、連結リスト、ハッシュテーブル」、9章で「スタックとキュー」、10章で「グラフと木」、11章で「Union-Find」が紹介されている。10章では、重み付きグラフ、順序木と二分木、ヒープ、二分探索木などが特に詳しく説明されている。11章の「Union-Find」とは、グループ分けを行うための特殊な木構造のことである。

12章は、特にソートアルゴリズムについて取り上げており、挿入ソート、マージソート、クイックソート、ヒープソート、バケットソートなどを説明している。

13章から16章にかけては、グラフ構造を用いたアルゴリズムが紹介されており、13章で「グラフ探索」、14章で「最短路問題」、15章「最小全域木問題」、16章で「ネットワークフロー」が説明されている。

13章では、深さ優先探索と幅優先探索を中心にグラフ探索例を取り上げている。

14章では、最短路問題の方法として、動的計画法、ベルマン・フォード法、ダイクストラ法、フロイド・ワーシャル法などを取り上げている。

15章では、最小全域木問題の方法として、貪欲法に基づくクラスカル法を取り上げている。最小全域木問題とは、たとえば、いくつかの通信拠点をすべて通信用ケーブルでつなぎ、すべての建物間で通信できる

ようにしたいときに、最小コストでこれを実現する方法を問うような問題である。

16章では、ネットワークフローの問題を取り上げている。ネットワークフローとは、輸送ネットワークにおけるトラフィックを考える問題で、最大フロー問題（単一の始点から単一の終点へのフローネットワークで最大となるフローを求める問題）と最小カット問題（カットとは、グラフをあるノードを含む部分と、別のノードを含む部分に、エッジ部分で分割することである。その分割されたエッジのうち、最小の容量のもの（エッジ）を求める問題）などを取り上げている。

17章は「PとNP」と題し、問題を解くアルゴリズムを見つけるのが難しい問題をNP困難問題として取り上げ、18章は「難問対策」と題し、そのような解くのが難しい問題に対し、現実的な解がどのようにして得られるかについて、特殊ケースで解ける場合、貪欲法で解ける場合、局所探索と焼きなまし法、分枝限定法、近似アルゴリズムなどを紹介している。

本書をどのような人たちに推薦したいか

アルゴリズムの入門書として、プログラミングを学び、コーディング力を鍛えたい人に入門書として推薦したい。特に、プログラミングを用いたデータ分析を実践する際に、よいアルゴリズムを自分で考える力を身につけたい人にはお薦めである。

参考文献

1) 猪股俊光, 南野謙一: 情報系のための離散数学, 共立出版 (2020).

(2020年11月19日受付)

石井一夫 (正会員) ishii_kazuo@med.kurume-u.ac.jp

久留米大学バイオ統計センター准教授。専門分野: ビッグデータ分析, 計算機統計学, データマイニング, 数理モデリング, 機械学習, 人工知能。2015年度本会優秀教育賞受賞。2019年度本会シニア会員。2020年(株)エヌ・ティール・エス学術顧問。日本技術士会フェロー, APEC エンジニア, IPEA 国際エンジニア。

