

第4章 データの扱い

I. データとデータモデル

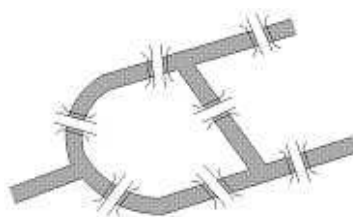
データという言葉は色々な意味合いで使われるが、ここでは「コンピュータの処理対象となる符号化された情報」の意味で使うものとする。データを体系的に扱うために上手く表したモデルを**データモデル**と呼ぶ。

II. 代表的なデータモデル

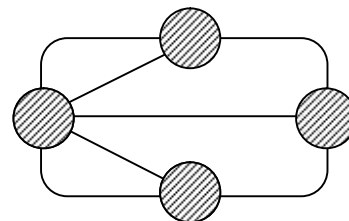
1. ネットワークモデル

ケーニヒスベルクの橋の問題

プロイセンのケーニヒスベルク (Königsberg, 現ロシア領カリニングラード) にプレーゲル川という川が流れており、そこには 7 つの橋が架かっていた。同じ橋を 2 度渡ることなく全ての橋を渡れるか？ 尚、どこから出発してもよい。



上の問題は、川と橋で考えるとわかりにくいですが、右のように4つの陸地(ノード)を結ぶ7つの道(エッジ)としてグラフ化すればわかりやすい(グラフについては第2章I節を参照)。このグラフはケーニヒスベルクの橋の問題の良いモデルとなっている。このグラフのように、繋がり方を表すモデルを**ネットワークモデル**という。因みに、この問題のように全てのエッジを重複なく辿る路を**オイラー路**という。オイラー路の判定は容易である。



ウェブの例

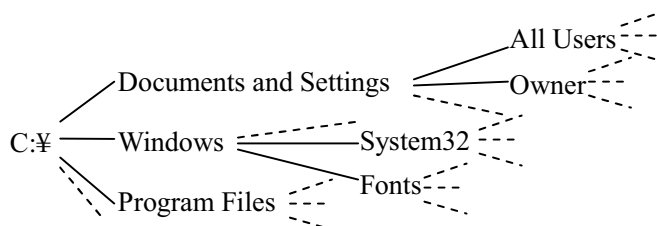
各ウェブページをノードと見做すと、リンクはその有向エッジである。このようにウェブをグラフの構造から分析すると、「重要なページからリンクされているページは重要」のような規則の方程式を作って解くことで、各ページの重要度を定めることができる。Google 等の検索エンジンでは、このようにして検索を効率化している。

2. 階層モデル

対象の理解の方法として“分類”はよく使われる方法の1つである。以下の例のように、分類した物事の枝別れの構造を**階層モデル**や**木構造**という。木構造は有向グラフの一種だが、向きが自明なので普通は矢印を描かない。

一. 階層型ファイルシステム

Windows や Macintosh などのコンピュータのファイルシステムは、右のように木構造になっている。このようになっていると、次のような利点がある。



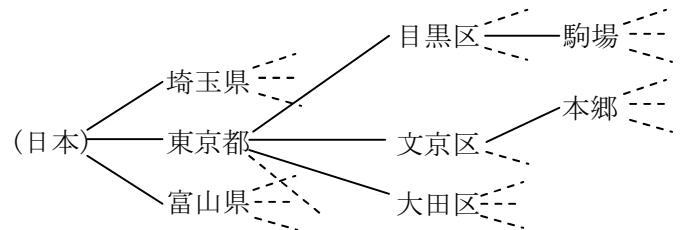
1. フォルダの中にサブフォルダを作って細かい仕分けができる。
2. フォルダを分けておけば、それぞれのフォルダのサブフォルダが混ざり合ってしまうことはない。
3. フォルダを辿ることで、全てのファイル・フォルダに対して重複なく一度ずつ同じ処理を行うことができる。例えば、全体のバックアップを取るなどである。
4. あるフォルダ内のファイル及びサブフォルダ全体に対して行う処理は、その最上位のフォルダに対する処理として代表して行うことができる。例えば、全体を移動させたいときは最上位のフォルダを移動させればよい。

二．住所の階層性

右図のように、日本国内の場所であれば
都道府県→市・郡→区・町村→町名(字)→
丁目→街区符号→住居番号などといったよ
うに一意的に所在地の表示が定まる。

例えば東京大学駒場 I キャンパスなら、
東京都→目黒区→駒場→三丁目→8 番→1 号
といった具合である。

このように、日本の住所表示は階層的な構造を成している。そのため、目黒区と文京区はそれぞれ独立にその下の住所を管理できる。



第 4 章の試験範囲は以上です。