[OpenGL 2D 2018 第01回]

コンピューターの仕組み

# コンピューターの機能

## はじめに

　コンピューターは数字や文字、画像や音声などの様々な情報を処理することができます。わたしたちは普段、電子メールやワープロソフトなどで文章を書いたり、画像を添付したり、場合によっては音声などを扱うこともあります。そのような時、わたしたちは文字や画像がコンピューター内でどのような仕組みによって表されているのかということはあまり考えていません。文字や画像をコンピューター上で自由に扱うためには、単にキーボードの配列を覚えたり、ワープロソフトの使い方に習熟すればそれで事足りると考えている人がほとんどでしょう。

　しかし、コンピューターを用いてゲーム開発のように高度な作業を行おうとした場合、必ずしもコンピューターの「使い方」を知っているというだけでは十分ではありません。コンピューターを自由に使いこなせるためには、単にコンピューターの使い方を覚えるだけではなく、その背後にある「コンピューターの仕組み」を理解することが必要となります。これは入学して間もない皆さんにとっては実感の湧かないことかもしれませんが、今後様々な授業やプロジェクトなどを通して、コンピューターの仕組みを理解することの重要性を実感していくことになるでしょう。

## 基本的な機能

コンピューターは、次に示す5つの機能を持っています。

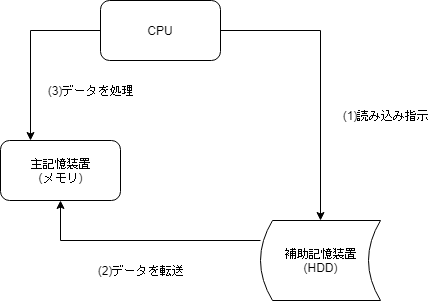
（１）制御  
（２）入力  
（３）記憶  
（４）演算  
（５）出力

　コンピューターの心臓部を構成するのは、中央処理装置(Central Processing Unit(せんとらる・ぷろせっしんぐ・ゆにっと)=CPU)と主記憶(一次記憶)装置です。これらの装置は、半導体素子を高密度に集めた「集積回路」と呼ばれる部品を組み合わせて作られています。コンピューターは命令を解釈し、それにしたがって演算を実行することで処理を行っています。中央処理装置(CPU)は制御機能と演算機能をもっています。制御機能とは、命令の流れやコンピューターを構成する装置の動作を管理する機能のことです。演算機能は命令を解釈して演算を実行する機能のことです。　コンピューターは、実行する命令やデータを中央処理装置に近いところに蓄えておき、これらを命令に応じて読み出して処理を行います。このために、コンピューターにはそれらの情報を記憶しておく機能が必要です。主記憶装置(メモリ)は一時的にプログラムやデータを記憶し、中央処理装置にそれらを直接提供します。

　パーソナルコンピューター(PC)などではプログラムやデータを保管するために補助記憶(二次記憶)装置と呼ばれるものを使用します。補助記憶装置は一般的に主記憶装置よりも大容量で、多くは周辺機器として接続れます。そのため、外部記憶装置とも呼ばれます。しかし、主記憶装置と比較すると読み書きの速度は遅い場合がほとんどです。外部記憶装置にはハードディスクドライブ(HDD)、ソリッドステートドライブ(SSD)、SDカード、USBメモリなどがあります。

　また、パーソナルコンピューターなどでは、人間とコンピューターとのあいだでコミュニケーションを行うために入力装置と出力装置を設けます。入力装置は人間の意図を伝えるために使われ、キーボードやマウス、マイクなどの装置があります。出力装置はコンピューターのデータ処理の結果を表示するもので、ディスプレイやプリンタ、スピーカーなどの装置が使われます。

　では、それぞれの装置について動作の仕組みを見てみましょう。まず、キーボードやマウスなどの入力装置を使ってソフトウェアを実行する指示を行うと、ソフトウェアがハードディスクなどの補助記憶装置から主記憶装置に読み込まれます。文書処理ソフトウェアや表計算ソフトウェアは、文書データや表計算データを作成、編集し、これらのデータも主記憶装置に置かれます。



　キーボードからデータを作成、編集するための命令を打ち込むと、ソフトウェアがこの命令を中央処理装置に伝え、必要なデータの一部が取り出されて、中央処理装置がデータに処理を加えます。データに処理を加えた途中経過や結果は、ディスプレイ上に表示され、作成した文書や表、計算結果はプリンタなどに出力することもできます。

　コンピューターの電源が切れると主記憶装置で記憶した内容は消えてしまいます。そのため、作成、編集が済んだデータを保存する場合は、主記憶装置から補助記憶装置に移し、そこに保存します。

## コンピューターの機能を実現する装置たち

　コンピューターの制御・演算・記憶・入力・出力の基本的な機能はどのようなハードウェアで実現されているのでしょうか。

**・中央処理装置（CPU)**

　中央処理装置は、演算機能と制御機能を担い、主記憶装置上のプログラムから一つ一つ命令を読み出して実行する装置です。その性能は、一つの命令で処理できるデータのビット数や動作速度で表されます。

　一つの命令で処理できるデータのビット数に応じて３２ビットCPUや６４ビットCPUなどがあり、データ処理の速度や規模に対応しています。また、中央処理装置の各部分の基本動作のタイミングを合わせるために、一定の時間間隔でクロックと呼ばれるパルス信号を発生させており、このクロック信号の間隔が短ければ短いほど中央処理装置の各部分の動作は速くなります。クロックが１秒間に何回あるかを示す数値としてクロック周波数が使われます。一般にその周波数が高ければ中央処理装置の動作速度も速いということができます。

**・記憶装置**

　記憶装置には、主記憶装置（メモリ）と補助記憶装置があります。主記憶装置は高密度の半導体記憶素子で構成されることが多く、その容量はコンピューターの性能を示すものの一つです。補助記憶装置には、ハードディスクのほかに、フロッピーディスクドライブやMO(光磁気ディスク）ドライブなどが使用されます。

　DVDドライブやCD-ROMドライブは、映画やゲームなどの大容量のソフトウェアの読み出しを行う補助記憶装置ですが、パーソナルコンピュータからDVDやCD-ROMを作成することができる装置でもあります。

　記憶装置の性能はその記憶容量や読み書きの速度で測られますが、容量や速度の組み合わせは様々で、用途・目的によって使い分けられています。

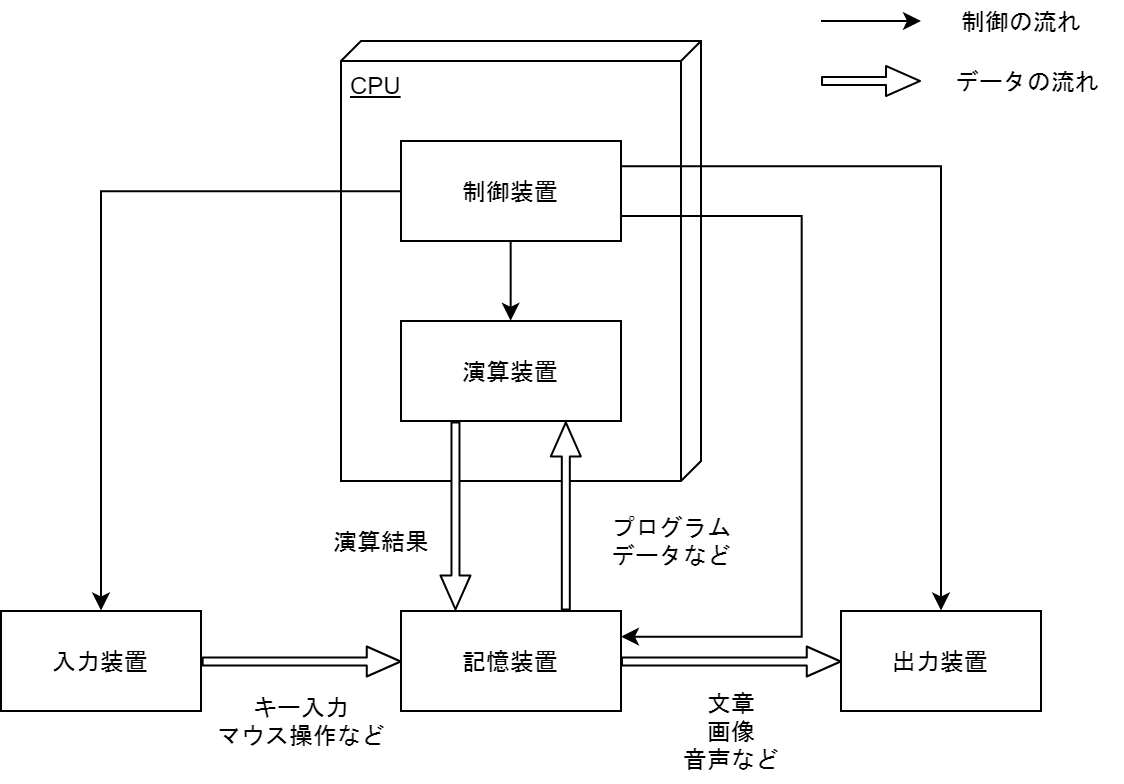
記憶装置は、入力装置または出力装置としても使われます。

**・入力装置**

入力装置にはキーボードやマウス、マイクのほかにも、カメラや指紋認証装置、トラックパッドなどがあります。温度センサーや明るさセンサーなども入力装置のひとつです。また、記憶装置も入力装置になり得ます。

**・出力装置**

出力装置にはディスプレイやプリンタ、スピーカーなどがあります。また、記憶装置も出力装置になり得ます。



## データとプログラム

　データやプログラムは、ハードディスク装置などの補助記憶装置にファイルとして保存されていたり、キーボードなどの入力装置から直接入力されたりします。これらのデータやプログラムは、そのまま中央処理装置が処理するのではなく、主記憶装置に送られ、そこに蓄えられます。

　中央処理装置は主記憶装置に蓄えられたプログラムやデータを読み込んで処理を行い、結果を再度主記憶装置に書き出します。そのため、データをハードディスク装置などにファイルとして保存するためには、保存操作を別に行う必要があります。主記憶装置は、一般にバイトという単位でデータを扱っています。１バイトは一般的に８ビットです。データを読み書きするために、このバイト単位のデータを収める小さな区画が論理的に設けられています。主記憶装置上のそれぞれの区画には、場所を示す数が割り当てられており、それぞれの区画の場所を示すそのような数を、アドレス（番地）と呼びます。

　中央処理装置や出力装置は、指定したアドレスから命令やデータを読み込んで処理を加えたり出力を行います。また、中央処理装置の処理結果や入力装置からの入力は指定されたアドレスに書き込まれます。

## コンピューターはプログラムをどのように実行するか

　プログラムはたくさんの命令が順番に並んだ集まりです。それを一つずつ「取り出し」、「解読し」、「実行する」という動作が繰り返し行われます。コンピューターは、基本的にこの命令を一つずつ実行していきます。そのため、命令を実行する順序は非常に重要です。

　命令は一般に処理の種類を表す命令部と処理するデータのアドレスを示すアドレス部の二つの要素から構成されます。制御装置には「次に実行する命令のアドレス」を記憶している場所があって、これにしたがって命令を取り出す仕組みになっています。命令は次のような手順で実行されていきます。

１　制御装置はプログラムによって設定された実行開始アドレスを「次に実行する命令のアドレス」として指定し、以下の２～６を繰り返します。

２　「次に実行する命令のアドレス」にある命令を制御装置に読み込む。

３　制御装置に設定された「次に実行する命令のアドレス」を一つ進める。

４　制御装置は命令部を取り出して解読する。

５　制御装置はアドレス部からデータのアドレスを決める。

６　解読された命令とデータのアドレスにしたがって処理を行う。命令が実行順序を変更する命令のときは「次に実行する命令のアドレス」を変更する。

　このように、コンピューターは主記憶装置に蓄えられた命令を順序だてて一つずつ取り出しては実行し、その流れは停止命令が実行されるまで続きます（実際には、プログラムは「それを管理している上位のプログラムに制御をゆだねる」という形で終了します）。複数の命令があっても、それぞれの命令を同時に行うのではなく、一つずつ実行することに注意しましょう。このようなデータ処理を逐次処理といいます。

# コンピューターはどのように情報を扱うか

## 2進法

　まず２進法と呼ばれる数字記法について説明をします。これは、コンピューターの仕組みを考える上での最も基本的な事項ですので、少し難しいかもしれませんが、頑張って理解しましょう。普段我々が日常生活で使っている数の表し方は１０進位取り記法（以下１０進法）と呼ばれる体系です。１０進法を用いると、一桁で１０個の別々のものを区別することができます。例えば、０は赤、１は青、２は緑・・・といった具合です。１１個以上のものを区別するためには、桁を増やすことによって対応します。

　さて、人間と違いコンピューターは２進法という表記体系を用いています。１０進法に慣れ親しんだ我々からすると非常に変わった体系に思われますが、後に述べるようにコンピューターにとっては非常に都合のよい体系なのです。１０進法においては０から９までの１０種類の記号がありましたが、２進法には０と１の二種類しか存在しません。つまり、２進法では１桁で２種類のものしか区別することができないのです。例えば、０が男、１が女、若しくは０が偽、１が真といった具合です。３つ以上のものを区別しようとすると、桁を増やす必要があります（二進法は２の累乗で桁上がりを起こします）。ここで、２進数と１０進数の対応を表にすると次のようになります。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10進数 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| ２進数 | 0 | 1 | 10 | 11 | 100 | 101 | 110 | 111 |

図１　２進数と１０進数の対応

　このように、例えば８個のもの（例えば、０から７までの数字）を区別する際に、１０進法では一桁しか必要ないのに対し、２進数では三桁必要となります。また、２５６個のものを区別するのに、１０進法では３桁あれば十分なのに対して、２進法では８桁も必要となります。随分２進法は効率が悪いと思われるかもしれませんが、電気回路であるコンピューターにとってはこちらの方が都合が良いのです。電気信号は一般にONかOFFの二通りの状態しか持たないので、ONを１、OFFを０とすると、ONとOFFの組み合わせで様々な情報を表すことができるのです。二進法を用いてコンピューターがどのように情報を表しているかについては後ほど詳しく説明をします。

## 情報量

　コンピューター内で表される数字や文字、画像などの情報は、必ず一定の量を持っています。「この画像ファイルは３メガバイトもある」とか「一秒間に５６キロバイトの情報を送れる」、もしくは「７ビットコード表」などということを聞いたことがあるかもしれません。この時、「３メガバイト」や「７ビット」というのは「ビット」や「バイト」という単位によって情報の量を表しています。一体、情報量とはどのようなことを意味しているのでしょうか。ここでは、情報量と情報量を表す単位であるビットやバイトについて学んでいきます。

## 情報量の単位としてのビット、バイト

　カードやコインには表と裏があるので、例えば表が黒、裏が白のカード一枚を使って、「晴れ」と「雨」の二つの天候を表すことができます。また、カードが２枚あれば「晴れ」、「雨」、「曇」、「雪」といったように４つの天候を、３枚あれば８つの天候を表すことができます。

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | 晴れ |
|  |  |  | 曇り |
|  |  |  | 雨 |
|  |  |  | 嵐 |
|  |  |  | 霧 |
|  |  |  | 雪 |
|  |  |  | みぞれ |
|  |  |  | 雹 |

図２

　さて、一枚のカードは二つの天候しか表すことができないのに対して、二枚のカードは四つの天候を表すことができますが、これは一枚のカードよりも二枚のカードの方が表すことのできる情報の量が大きいということを意味しています。三枚のカードは８つの天候を区別し、表すことができるので、二枚のカードよりも更に情報量は大きいということができます。このように、情報量とは「何種類の情報を区別し、表すことができるか」ということを意味する概念で、当然、一枚のカードより三枚のカードの方が表すことのできる情報の数が多く、情報量は大きいわけです。そして、情報量の単位として用いられるのが「ビット（語源は２進法の一桁を意味する、Binary Digitです）」といわれるもので、一枚のカードやコインのように、二つのものを区別できるものは、「情報量１ビットをもつ」といいます。表が黒、裏が白の３枚のカードを並べると８つのものを区別し、表現することができますが（図２）、この時３枚のカードによって表現されたものは情報量としては３ビットを持つといえます。同様に、図３のように１６枚のカードを並べた画像表現は１６ビットの情報量をもつといえます（すなわち、２の１６乗通りの異なるものを表現できるということです）。ためしに、表が黒、裏が白のカード６４枚がもつ情報量をビットで表してみてください。

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

図３

　さて、コンピューターの内部では、全ての情報はコインやカードの表裏のかわりに１か０かの二進数によって表現されます。先ほどのカードの表裏を二進数の１、０に置き換え、８つの天候を０から７までの数に置き換えると、図１のように対応づけることができます。

この場合、８個の数を区別することのできる二進数表現（すなわち、三桁の二進数）は何ビットの情報量をもつといえるでしょうか。二つのものを区別できるもの（例えば、一枚のコイン、一枚のカード）は情報量として１ビットをもつのでした。二進数において「一枚のコイン」の「枚」にあたるのは、桁数であるといえます。一桁の二進数ならば、１か０という二つの状態の区別ができるので、１ビットの情報量をもちます。同様に、２桁ならば１１、１０、０１、００という４つの情報を区別できるので、コインやカードで言えば２枚、すなわち２ビットの情報量をもつといえます。そうすると、０から７までの８つの数を区別するためには二進数で３桁必要なので、３ビットの情報量をもつことになります。一般に、（二進法は２の累乗で桁が増えるので）二進法ではN個のものを区別するためにLog N桁、すなわちLog Nビット必要であるといえます。大変わかりづらいところなので、各自図４を見てビットと桁数、区別できる情報の数の関係について確認してみて下さい。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 10進数 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2進数 | 0 | 1 | 10 | 11 | 100 | 101 | 110 | 111 |
| ビット | 1 | 1 | 2 | 2 | 3 | 3 | 3 | 3 |

図４ ビットと桁数の対応表

　また、２５６個の情報を区別するためにはLog２２５６＝８ビット必要ですが、一般に８ビットを１バイトと呼ぶことになっています（７ビットを１バイトとすることもあります）。さらに、１０００バイトを１キロバイト、１０００キロバイトを１メガバイトと呼びます。よって、「このファイルは３メガバイトもある」という表現は、「このファイルは３メガバイト＝３０００キロバイト＝３００００００バイト＝２４００００００(2400万)ビットの情報量を持つ」ということを意味しています。カードの例で言えば、２４００万個の表が黒、裏が白のカードを並べて表現されたものであるといえます。そして、２４００万ビットの情報量を持つ、とは２の２４００万乗個のものを表現できるということを意味しています。

## 16進数

　２進法の０と１を使ってコンピューターは様々な情報を表しています。このように、情報を表すために使用されている０と１の配列のことをビット列といいます。ビット列で情報を表す場合、情報量が増えると桁数もそれにつれて増えていきます。桁数が増えすぎると読みにくいため、ビット列を４桁ごとに区切って表現することがあります。このとき、４桁の２進数を１６進数で表します（図５）。

|  |  |  |
| --- | --- | --- |
| 2進数 | 10進数 | 16進数 |
| 0 | 0 | 0 |
| 1 | 1 | 1 |
| 10 | 2 | 2 |
| 11 | 3 | 3 |
| 100 | 4 | 4 |
| 101 | 5 | 5 |
| 110 | 6 | 6 |
| 111 | 7 | 7 |
| 1000 | 8 | 8 |
| 1001 | 9 | 9 |
| 1010 | 10 | A |
| 1011 | 11 | B |
| 1100 | 12 | C |
| 1101 | 13 | D |
| 1110 | 14 | E |
| 1111 | 15 | F |

図５ ２，１０，１６進数の対応表

　１６進数は１６とその累乗で桁上がりを起こしますが、０から９までの１０個の数だけでは不足なので、これに加えてA～Fの６個の英大文字を使って数を表しています。そのため、１０進法の１０は１６進法のA、１０進法の１５は１６進法のFとなり、１０進法の１６が１６進法での１０となります。１６進数の位取りの考え方は次のようになっています。

## 2進法と16進法の変換

２進法と１６進法の相互変換は次のように簡単にできます。たとえば、２進法の１００１１１００１１０を１６進数にする場合、以下の手順で変換します。

１． 与えられた２進数を、下の位から順に４つずつのグループに分ける。

１００　１１１０　０１１０

２． 各グループを、対応表を基に１６進数で置き換える。

１００→４　１１１０→E　０１１０→6

３． 答えは4E6

　１６進数から２進数に変換するには、これと逆の操作をすればよいわけで、１６進法と２進法との変換は１０進法と２進法の変換に比べてずっと簡単です。

## （演習）

次の２進数を１０進数に変換しなさい  
（１）１１０  
（２）１００１００

次の１０進数を２進数に変換しなさい  
（１）１５  
（２）２８７

次の２進数を１６進数に変換しなさい

（１）１０１０１１１１

（２）１００１００１００１

次の１６進数を２進数に変換しなさい

（１）７Ｆ

（２）３Ｅ８

# コンピューターはどのように文字を扱うか

さて、以上で文字や数値などの情報を２進法で表現するための準備が整いました。いよいよコンピュータの中で実際に使われている表現方法を学んでいきましょう。ここでは文字情報をコンピュータ上で表現する原理について学びます。

## コード化

　コンピュータで情報を表す第一歩は、様々な記号の中からコンピュータで表すべき記号を選択することです。例えば、漢字ならば使用頻度の高いものを選択する必要があります。次に、選択した記号をどのようなビット列に対応づけてコード化（符号化：encoding）するかを決めなければなりません。例えば、００が晴れ、０１が曇り、１０が雨、１１が雪という対応づけをすることも、立派なコード化です。コード化された記号全体をコード系（character encoding system）と呼び、通常は表の形でまとめられます。この表のことを、コード表（character encoding table）といいます。

　文字のコード化は、文字記号の少ないアルファベット圏では比較的容易ですが、漢字圏のように多数の文字があるところでは様々な問題があり、困難です。どの文字を採用するかを考えることも大変ですし、採用した文字とビット列の対応づけも面倒な作業になります。加えて、後で文字を追加するときには、先に作られたコード系といかに整合性を保つかといった多くの作業をこなさなければなりません。

## コード化の規約

　さて、実際にコード化はどのように行われているのでしょうか。かってなコード化が行われれば、ソフト間のデータのやりとりや通信に混乱が起こることは容易に想像がつきます。電話回線や通信回線を使用して文字データを送受信する場合にはお互いに文字の構成フォーマットを決めておかないと送信した文字情報を受信した文字情報とが一致せずに意志を伝えることができません。ですから、コード化の作業は、関係者や関係諸団体が集まって議論をしながら行われています。

　コンピュータは英語圏のアメリカで誕生したので、まず英語のアルファベットのコード系がつくられました。今日の社会では英語は事実上の国際語として機能しているため、このコード系はすべてのコンピュータ社会で使えるようになっています。英語圏の代表的なコード系としてはASCII(アスキー)があります。

しかし、コンピュータが広く日常的に使われるようになると、英語圏以外の国は自国の言語に応じたコード系を作るようになりました。それらのコード系はその言語圏でのみ利用することを前提として作られたので、インターネットのように国際的なコンピュータネットワークが普及しはじめると様々な問題が表面化してきました。特に、文字数の少ないアルファベット圏と漢字のような文字数の多い言語圏との間では、記号に割り当てるビット数の違いなど、大きな問題が生じたのです。

　現在、日本で使われている代表的なコード系としては、iso-2022-jp、シフトJIS、EUC-JP、Unicode(UTF-8, UTF-16) などがあります。

## Unicode(ユニコード)

　コンピュータが世界的に普及するに伴い、様々な言語の各種アルファベットと漢字などの表意文字を統一的に扱おうという提案がなされており、ISO(あいえすおー、国際標準化機構)という団体が、これを取りまとめて国際規格として採択しました。この規格はユニコード（Unicode）と呼ばれ、一部の例外を除き全ての文字を取り扱うことを目指しています。当初はUTF-16という1文字に16ビットを割り当てるユニコード規格が制定され、これによって、最大65536種類の文字を表現することができるようになりました。しかし、文字は各国、各民族に固有の文化が関係してくるため、16ビットでは全ての言語で使われる文字を統一的にもれなくコード化するのは難しいことでした。また、英語圏で広く普及していたASCIIコード系との互換性がなかったこともあり、なかなかUTF-16は普及しませんでした。その結果、現在ではASCIIコード系との互換性を保ちながらも、最大31ビットの情報量(文字にすると約21億種類以上)を持つ、UTF-8というコード系が普及しつつあります。

# まとめ

ここまで、コンピューターがどのような仕組みで動き、どのように情報を扱うかを見てきました。  
コンピューターは電気信号のONとOFFを２進数に見立てて、さまざまな計算を行っています。  
また、数値や天気、画像、文字などの情報はコード化することで、それらをコンピューター上で扱うことができるようになります。  
情報は補助記憶装置から主記憶装置へと読み出され、さらに中央処理装置(CPU)へと送られてさまざまな演算が行われた後、主記憶装置を経由して補助記憶装置に書き戻されたり、ディスプレイやスピーカーなどの出力装置に送られます。  
そして、プログラムはこの流れを制御する命令が順番に並んだ集まりです。プログラムを変えることによって、コンピューターにさまざまな動作をさせることが可能となります。また、プログラム自体も、コンピューターで処理される情報の一種です。