

科学技術リテラシーとしての 情報学リテラシー



早稲田大学理工学術院 基幹理工学部情報理工学科教授 笥 捷彦

難しいお題を頂戴してしまいました。いわゆる「情報リテラシー」というと、誰でも知っていてほしいITの読み書き（使い方）を意味するのが普通だ。そこに、わざわざ「科学技術リテラシーとしての」という修飾がついているのだから、いわゆる「理系」の人にとっての、しかも「理系」の作業をする場合のITの読み書きについて述べよ、ということであろう。また、「情報学リテラシー」と明記してあるところを見ると、より原理原則まで立ち戻ったものを書け、という注文であろう。可能な限り、その注文に即して書いてみる。

IT (information technology, 情報技術) は、コンピュータを使った情報を記録・複製・編集・加工・転送する技術、およびその技術を使った仕組みや装置などを指す。より古くから使われていた用語に、情報処理 (information processing) というのがある。最近では、特にインターネットの普及・発展もあってコミュニケーションの道具として活用されることを強調したICT (information communication technology, 情報通信技術) と呼ばれることもある。

さて、そのITは、その名称からすると情報を直接に扱う技術だと思ってしまうが、そうではない。正確にいうとITはデータを扱う技術である。ここでいうデータとは、何らかの情報を表現したもの、それもデジタルな表現を指す。平たくいえば、データとは、文字とか記号とかを並べたものである。

すると、情報学リテラシーの第一歩は、ITがやってくれることがデータの記録・複製・編集・加工・転送といった処理にすぎない、ということをお腹に銘じておくことである。当然ITにできることは、与えられたデータが特定の方式で情報を表しているとしたとき、そのデータを元に、それらが含意している情報を表すデータを作り出すことに過ぎない。情報とデータとの間の対応づけは、ITの外の話なのである。

データは情報そのものではない。データがどんな方式で情報を表しているのかについて知っていなければ、どんなデータも意味をもち得ない。古代文書を入手しても、そこに書かれた記号が何を表し、それらの列がどんな意味を運んでいるのかが分かっていなければ、紙くずと変わらない。膨大な数値列を入手したところで、それらの数値が何をどのように表したものかを知っていなければ、これまた、ごみと変わらない。

さて、科学技術で扱うデータのほとんどは数値である。ニュートン力学での大成功があって科学技術が急速に発展したこともあって、科学技術が対象とする現象や事象の多くは、実数を用いて扱われている。その実数そのものをデータとして、つまり数値として記録することは不可能である。そこで近似が行われて、有限長の数字列、それもほんの少数の桁数の数字列として記録される。そこに誤差が入る。

誤差の入っているデータを使って計算した結果

にどれほどの誤差が入ったものになるか、については、科学技術の勉強の中でいやというほど学ばされたはずである。手でやろうが、ITを使ってやろうが、その誤差について変わるところはない。しかし、人は、ITを使って出したのだから、とITによってもたらされた結果を、つい鵜呑みにしがちである。誤差の話は、ITの発展以前から確立しているものであるが、実は、科学技術における情報リテラシーの根本原則として心しておくべきことでもある。

加えて、実数をコンピュータで扱うときに用いられている浮動小数点表現にまつわる注意が必要になる。浮動小数点表現とは、物理定数などを表すのに使われているものと同じで、 3.14159×10^0 などと特定桁数の有効数字（この例では6桁、3.14159）と、大きさを示すべき乗数（この例では0）との組で実数の近似を表す。同じ桁数の有効数字をもつ二つの近似値を使って計算した結果を同じ桁数に丸めたものは、ほとんどの場合、近似前の元の実数を使って計算した結果を同じ桁数に丸めたものにほぼ等しい。ところが、例外的にそうならないときがある。値が近い2数について引き算を行う場合、浮動小数点に近似した値を使って計算した結果は、大幅に有効数字の桁数が減ってしまう。例えば、何らかの実測値が 3.15×10^0 であったとき、 3.14159×10^0 との差は $0.05841 \times 10^0 = 5.841 \times 10^{-2}$ であって5桁の有効数字しか残っていない。この現象を桁落ちという。手計算でやっていれば、計算のやり方を変えなければならないことにたちまち気づくはずが、コンピュータに任せて計算が進んでしまうと桁落ちが起きていたことにさえ気がつかないままになりかねない。これも科学技術における情報リテラシーのうちとして心得ておくべきことである。

ITは、膨大な量のデータを高速に処理することを可能にしてくれた。それは大いなる便宜をもたらしてくれているが、同時に、それは瞬時に多量のデータが流出してしまうことも可能にしていることに留意する必要がある。情報セキュリティの意識をもつことが大切になる。

多量のデータが流出したからといって、それだけでだちに情報が漏洩したことになるとは限らない。データは、どんな情報をどのように表しているのかが分かっているのはじめて「情報」として意味をもつものだからである。データが一定の明快な規則で情報を表しているなら、そして、一群のデータがこんな情報を表しているはず、ということが分かっているなら、その情報を表している方法を復元することも、ITの高速さをもってすれば大した時間をかけずにできてしまう可能性が高い。そこで、データを暗号化する、ということが行われる。情報セキュリティを考えるとときの重要なポイントである。

ITの高速多量のデータ処理は、コンピュータによって行われる。コンピュータは、与えられた指示に従って動く。その指示をプログラムという。そのプログラムもまた、情報であるから、方法を定めてデータとして表すことができる。処理対象となるデータも、処理内容の指示を与えるプログラムを表したデータも、データであることに変わりがない。それらを一緒に並べて記録しておいて、後はコンピュータという機械が動いて、そのプログラムどおりにデータに処理を施していく。

プログラムをデータとしてどのように表現するかを定めたものを、コンピュータの機械語という。プログラムを表す書き方を定めているのだから、日本語、英語と同様に言語だと呼んで差し支えないだろうということでの命名である。この機械語という言語は、すこぶる簡単な規則でできている。たかだか数十から数百程度の、命令と呼ばれる単位データ形式が用意されているだけであり、プログラムはこの命令を並べることで表現する。コンピュータは、その少数の命令を正しく解釈できさえすればそれでいいことになる。現時点でのコンピュータは、いずれも電子回路として実現されている。これをハードウェアと呼ぶ。ハードウェアに誤りがないかどうかは、用意された命令について全数検査して確認することができる。もちろん、その検査を人手で行っていたのでは話

にならないから、それを自動的に行うプログラムを作って別のコンピュータを使って実施するのである。したがって、コンピュータという電子回路そのものに間違いが残されている可能性はほとんど起こり得ないと考えていい。

一つのコンピュータの上で、プログラムを与えさえすればどんなデータ処理も行うことができる。これがITを成り立たせている根幹の仕組みである。コンピュータというハードウェアは、何をやらせたいのか、という目的によらず一定であっていい。プログラムを差し替えるだけで別の用途の装置として働いてくれる。さらに、やりたいことの内容を増強したり変更したりするのは、ハードウェアを変更する必要がなく、プログラムだけ取り替えればいい。このようにいかようにも調整が効く、という意味で、プログラムのことを、ハードウェアに対比してソフトウェアと呼ぶ。

コンピュータというと、デスクトップPCとかノートPCとかをすぐ思い浮かべるが、実はそれら以上にコンピュータは世の中に溢れているのである。身の回りにあるいろいろな装置にコンピュータが入っている。しかも、複数個入っていたりもする。電気釜であれ、冷蔵庫であれ、クーラーであれ、果ては自動車であれ、飛行機であれ、コンピュータが入っている。専用の電子回路を個別に用意するのではなく、同じコンピュータを使っても適切なプログラムを組み合わせれば、それぞれの特定目的に合わせた装置として動いてくれるのである。もちろん、通信装置としてしか意識していない、携帯電話・スマートフォンも、ゲーム機も、すべてコンピュータを使って作られている。どこのスマートフォンが便利か、どのゲーム機が面白いかなどはすべてソフトウェアで勝負がついているのである。

このように、IT装置がIT装置として機能する大本は、いかにしていいソフトウェアを作るか、すなわち、高機能なプログラムを作るか、ということに尽きることになる。そのプログラムを作る作業をプログラミングという。ITを使って得られた結果は、誰かが作ったプログラムの指示通り

にコンピュータが自動的にデータ処理を行った結果にすぎない。つまり、その結果が正しいかどうかは、そのプログラムがいつでも意図どおりの結果をもたらすように正しく作られていたかどうかによって依存する。ところが、プログラムを正しく作ることはたやすいことではない。どうしても間違いが混入してしまう。

では、プログラムを作る作業をコンピュータに自動的にやらせればいいではないか、と思うかもしれない。ところがプログラムを作る作業を自動的に行わせるためのプログラムというものは作り得ないことが分かっている。経験的にそうだが、いうのではなくて、理論的に証明されているのである。だから、どうしてもプログラミングは人智によるほかはないのである。

ITがプログラムで動かされているものであって、そのプログラムが人智によるほかではなく間違いが混入してしまっているかもしれないものだとすると、ITで得られた結果であっても、間違いが混入している可能性があるということになる。それを心にとめて、得られた結果の蓋然性を確認する習慣を身につけよう。教科書に書いてあることであっても、論文として公表されていることであっても間違っていることがある。それと同じで、使い古されているプログラムであってもそこに間違いが入っていることがあり得る、と心する必要がある。

ITは、高速多量のデータ処理を可能にしてくれた。しかも速度もデータ量も、3年でほぼ2倍になるという勢いで上がってきている。それでもなお、身近な天気予報一つをとってみても思うように精密な計算を適切な時間内に行うことができるにはまだ遠い。それに、いつまでもこの勢いで速度も量も伸びていくわけにはいかない。コンピュータというものの新しい方式の研究が急がれている。

同じデータ処理を行うにも、何通りものやり方が可能だし、何通りのものプログラムの作り方が可能である。データ処理対象の問題によっては、最高速な方式が分かっているものがある。中に

は、原理的にしらみつぶしにやるのと同等の方式しか見つからないものもある。N個の対象の並べ方の中で何らかの最適解を求める問題がしらみつぶしで調べるほかないとすると、世界一速いコンピュータをもってしても、N=30の答えを出すのは世の終わりを待つほかはないことになる。情報セキュリティに用いられている暗号の中には、こうした問題を逆手にとって暗号を解くやり方は既知でも具体的に解くには最速のコンピュータをもってしても世の終わりを待つほかはない、という方式を使っているものもある。

具体的に解く手順が分かっている問題については、人手によらなくても適切なプログラムを使えば解が得られる。たとえば、積分をしたり、微分方程式を解いたりすることも、解法が既知の場合に対してなら数式処理プログラムを使えばたやすくできるようになってきた。こうしたプログラムを使って結果が得られた場合でも、その結果が正しいものであると鵜呑みにしてはならない。そのプログラムが特殊な条件下でなければ成り立たない性質を使っているかもしれないし、そのプログ

ラムが間違っているかもしれない。結果の蓋然性を自ら確認することが大切である。筆算で何かの結論を出したときは検算する。それと同じことである。

ITは、科学技術における計算や思考を大いに助けてくれる。しかも、高速多量に処理してくれる。しかし、それはあくまで情報を映したデータに対して、プログラムによって起こされるものである。情報とデータの対応、それにプログラムの信頼性が得られてはじめて意味をなすことを肝に銘じておくことが肝要である。



初級C言語 —やさしいC—

早稲田大学教授 笥 捷彦 監修

日本電子専門学校 後藤良和・高田大二・佐久間修一

B5判/304頁/定価2,310円(税込)

プログラミング初心者向けに書かれたC言語のテキスト。

図や具体例を数多く示し、プログラムの流れはフローチャートを用いずに説明しています。

<目次>はじめてのプログラム/変数/画面への表示/簡単な計算/キーボードからの読み込み/処理の流れ/たくさんの値の扱い—配列—/文字列の扱い—文字配列—/ポインタ/プログラムの小分け1—関数—/データのまとめ—構造体—/ファイルの処理/用意されている関数—標準関数—/プログラムの小分け2—いろいろなデータ—/変数の有効範囲—記憶クラス—/プログラムの字面の書き換え—前処理—