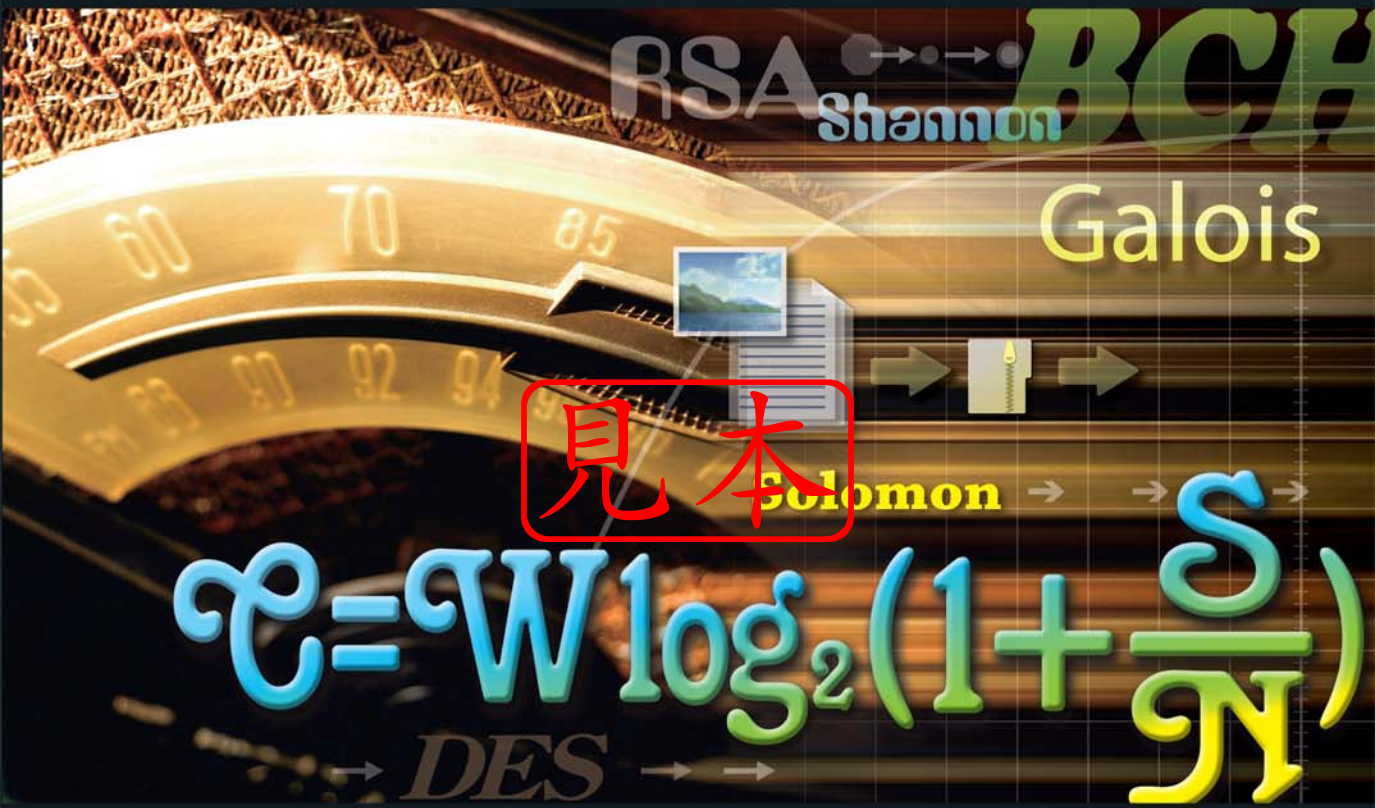
ディジタル信号処理シリーズ

Scilabで学ぶ
情報基礎，誤り訂正符号，暗号

【改訂新版】

やり直しのための工業数学 情報通信編

●三谷 政昭 著



CQ出版社

まえがき

理数離れが言われている今日このごろ，『物理数学』，『電気数学』，『情報数学』，『信号数学』，…といった“数学”という言葉を見ると，身体が硬直して二度と見たくないと思われる人が案外多い（ほとんど例外なく，そうかもしれない）のではなかろうか．しかし，数学を理解して使いこなせるとなれば，周りの人からも一目も二目も置かれ，製品開発，回路設計，ソフトウェア作成などの仕事面でのプラスは計り知れないものがあることも事実である．

『数学は仕事に役立たなければ，何の意味もない』

2001年1月には，実学としての情報通信や信号解析の数学がもつ面白味，醍醐味を味わっていただくことを念頭に，CQ出版社より「やり直しのための工業数学（情報通信と信号解析）」と題する書籍を発刊した．好評を博したかどうかは知る由もないが，さらなるパワーアップを目指して，本書「改訂新版 やり直しのための工業数学（情報通信編）」を出版することにした次第である．

近年，地上/BS/CSデジタル放送をはじめとして，インターネット，ADSL，ディジタル通信，携帯電話，…などのように，ありとあらゆる情報がすべてディジタル・データとして統合され，多種多様なサービスが提供されている．いずれのサービスも，ディジタル情報通信時代を支える“誤り訂正技術”，セキュリティ技術としての“暗号”などの底力なしに実現しえなかったであろうことに疑問の余地はない．

ところで，“誤り訂正技術”の「データ誤りを発見して修正するメカニズム」，“暗号”の「計算アルゴリズム」などの土台は，『情報理論』，『符号理論』，『暗号理論』の基礎数学であることも明白な事実である．さらには，これらの基礎数学が新しい発見・発明のヒントを隠し味としてもっていること，簡略計算テクニックの裏付けになっていることも特筆すべきであろう．

本書は，情報・符号・暗号に関係する数学を基礎からもう一度やり直したい人，仕事でばりばり使いたい人，…，そういった皆さんをターゲットに，数学が魔法のツール（道具）として仕事上の大きなパワーの源として，基礎数学を実務に直結させて有効に活用できる術をマスターするための知恵，知識を提供するものである．

本書は3部から構成され，主に数式表現と情報理論的な取り扱い，誤り訂正や暗号化するメカニズムとの関わり合いにスポットを当てて説明する．その際，“数学”的な難しい説明はほどほどに，数式のもつ物理的なイメージを中心に，Scilabによる具体的計算例を示しつつ，“わかりやすい”解説を心がけている．

まず，第1部は『情報基礎』と題し，情報理論的な側面（確率，統計など）や情報数学（符号や暗号などの基礎理論）を中心に説明する．次に，第1部の内容を受けた，第2部『誤り訂正符号』では符号理論に必要な数学，いろいろな誤り訂正符号の符号化/復号化などについて具体的に解説する．第3部『暗号』では，符号理論から暗号理論への橋渡しを行い，代表的な暗号としてRSA暗号とDES暗号をとりあげる．

とにもかくにも，数学的な素養として，情報理論，誤り制御技術，暗号理論に習熟しておくことがもっとも重要であることを体感してもらいながら，本書を読み進めていくプロセスにおいて，情報・符号・暗号の基礎をしっかりと習得してもらいたい．

おわりに，本書の出版にあたり，何かとご面倒をお掛けしたCQ出版社の山形孝雄氏/山縣妙子氏ならびに元社員の大野典宏氏に感謝の意を表します．

2011年2月　著者しるす

第1章 情報数学が使われる分野

近年，「マルチメディア情報ネットワーク」，「プライバシー情報」，「情報セキュリティ」，……などなど，情報という2文字を含んだ言葉が世の中をとびかっている．よく耳にする情報は，ディジタル通信ネットワークの世界にどっぷりと浸って生きていく私たち，とくに技術者にとって必要不可欠のものであり，その本質をしっかりと見据えておかなければならない．

また，画像や音声などのマルチメディア情報を処理するための製品開発，ソフト作成などの仕事面では，情報理論的な思考方法，情報理論的なセンスが要求されていることも事実である．

情報理論的な側面(確率，統計など)をもう一度基礎からやり直したい人，情報数学(符号，暗号，データ圧縮などの基礎理論)をしっかりと理解しておきたい人……，そういったみなさんにとって，第1部『情報基礎』は大いなる知識，知恵を提供するものである．

ところで，私たちの身の周りには，メモリ・オーディオ，スマートフォン，インターネット，電子メール，WWW(World Wide Web)，地上/BS/CSデジタル放送などを利用する場合において，文書，画像，音声などのさまざまなマルチメディア情報が氾濫している．こうしたマルチメディア情報は，ディジタル信号(0と1)として一元化されることが大前提であり，これまでのテキスト処理とは異なり，画像を中心とした大容量情報が主体となっている．

こうしたマルチメディア情報ネットワーク時代に必要とされる数学的な基礎をしっかりと理解しておくことは，ディジタル情報処理全般にわたる総合的な理解を深める際に大いに役に立つことに疑問の余地はない．

本章では，おもにディジタル情報，論理的な処理，情報数学の概要などについて説明する．その際，“数学”的な難しい説明はほどほどに，情報数学のもつ“とっつきにくさ”を解消してもらうことを最大の目標にして，わかりやすく解説する．とくに，本章の内容は，みなさんが情報理論的なセンスを身に付けてもらううえでのウォーミングアップになるものなので，しっかりと読み進めていってもらいたい．

1.1　情報数学で何ができるのか？

情報数学といえば，確率，統計という実用的ではない数学だというイメージがつきまとうようである．実生活の場面では，せいぜい天気予報で「今日は雨が降る確率が高いから，傘をもっていこう」とか，競輪・競馬などのギャンブルや宝くじなどでの当たり/はずれで「○△という馬が勝つ確率が高い」，「宝くじは当たらないね」，という言い方がされるぐらいであろう(**図1.1**)．

ところが，情報数学に土台をなす情報理論となると，さまざまな実生活の場面で登場してくるのである．たとえば，携帯電話やスマートフォンなどを使って友達と話をしたり，CDやメモリ･オーディオで音楽を聴いたり，DVD/ブルーレイで映画を見たり……など，ほとんどの日常場面で，知らず知らずのうちに情報理論の恩恵にあずかっているといっても過言ではない．無線，有線が入り乱れたディジタル通信ネットワーク・システムにしても情報理論のかたまりであるし，CDやメモリ・オーディオではディジタル信号の誤り訂正を入れても少々の傷ぐらいでは，ガリガリと耳障りな雑音が入らずクリアな音楽を再生できるようになっている(**図1.2**)．

こうした現在の電脳社会を実現する基盤が情報数学に根差しているわけで，基本的な考え方や数学的表現に精通していることは，21世紀に生きる技術者にとっては絶対に必要な知識であると言い切れよう．

まずは，情報数学の適用例をたとえ話にして，そのイメージの解説を行う．なお，情報数学はこれまではアナログ信号をおもな対象としていたが，現在は画像や音声を一元化したマルチメディア情報でディジタル信号であり，「ディジタル情報数学」，あるいは「ディジタル情報理論」と言い換えるのが妥当な感じがする．

● データ圧縮(図1.3)

わかりやすい例をあげると，「トラギ」(3文字)といえば，賢明なるCQファンの読者なら「トランジスタギジュツ」(10文字で，トランジスタ技術のこと)だと理解されるであろう．ここに，データ圧縮の基本的なコンセプトが眠っているというわけだ．正確にいうと10文字かかるところが，たったの3

図1.1　情報数学の利用される分野

図1.2　ディジタル化による利点

	データ圧縮	略語
World Wide Web	→	WWW
Integrated Circuit	→	IC
パワーステアリング	→	パワステ

図1.3　データ圧縮の基本的な考え方

文字でわかるので，データ量が3/10で済むことが理解される．

また，「勉強する」という動詞の活用形は，「勉強しよう（未然形），勉強します（連用形），勉強する（終止形），勉強すること（連体形），勉強すれば（仮定形），勉強しろ（命令形）」となるが，「運動する」や「処理する」なども同様で，変化したところだけ，つまり活用語尾「しよう，します，する，すること，すれば，しろ」を記憶すると，他の類似した動詞にも適用できるのである．こうした変化した部分のみを情報として浮きだたせる処理も，データ圧縮の一手法であるといえる．

● 符号化/復号化

何人かが順に，限られた時間内で話を伝えていって最後の人が理解した内容と原文との違い，すなわち最初の人の話をどれだけ正確に伝えられたかを競う伝言ゲームは，符号化/復号化の概念に近いものが含まれている（**図1.4**）．

まず，かなり長い話を限られた時間にまとめるという作業（符号化）を行って次の人に伝達し，話を聞いた人は頭の中で原文の内容を再構築するという作業（復号化）を行い，また次の人に伝えていくという処理（情報伝送）を繰り返す．こうした繰り返しの作業では情報が失われるという情報誤りがつきもので，原文とは似ても似つかぬ内容が伝達されることが往々にして起き，この伝言内容変化のプロセスを楽しむのである．

ところが，パソコン通信，インターネットなどのデータ伝送では，こうした情報誤りは絶対に避けなければならないわけで，誤りが発生しにくいしくみ（誤り検出，誤り訂正）を組み込む必要性が出てくる．

たとえば「トラギ」という言葉を誤りなく正確に相手に伝えるためには，「トラギトラギトラギトラギ」と同じことを繰り返して言うとか，「とまとのト」，「らいおんのラ」，「ぎんこうのギ」と言うとか，みなさんもごく自然に誤りが起きにくいような工夫をしているのである（**図1.5**）．

図1.4　符号化/復号化―悪い例

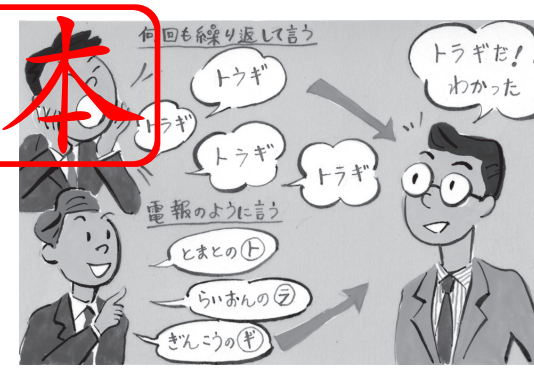


図1.5　情報伝送誤りを防止するには

図1.6　暗号の基本―符丁

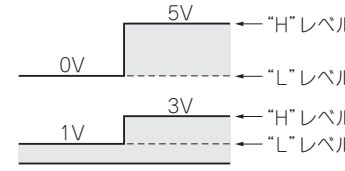


図1.7　2値動作の電圧レベル

● 暗号

私事で恐縮だが，家では椅子に腰掛けて，「オイ」と言うとお茶が，「オーイ」と言えば新聞紙が運ばれてくるという夢のような生活をしている(実際は大きく違っていて，みなさんのご想像どおりではあるが)．「お茶をもってきてほしい」とか「新聞紙をもってきてほしい」と言わなくても，仲の良い意思疎通の完全な夫婦であれば，「オイ」とか「オーイ」という二人の間だけで通じる会話が成立するのである．これは，まさしく暗号で，心が安らぐ感じ，安心できる符丁(符号)という気がする(**図1.6**)．

こうした他の人にはわからない言葉で，当該者同士がお互いに話ができるというしくみが「暗号システム」とよばれるものであり，情報が外に漏れないようにすることの重要性が認識され，驚くべきことに暗号を商売にする会社までも出現している．

1.2　論理回路を知ろう

情報数学の基本的な考え方を用いると，データ圧縮，符号化/復号化，暗号などのしくみ(アルゴリズム)を作れるわけだが，実際に電子回路やソフトウェアで実現するとなると，0と1の2進数の世界での演算処理を行う．2進数の演算に関しては，ブール代数，論理代数，真理値表，論理式が重要な意味をもち，ディジタル回路(論理回路，スイッチング回路)を用いてハードウェアで実現される．

まずは，数学的な話をする前に，0と1の演算処理を実現するための論理回路を考えてみよう．一般に論理回路は数学的な取り扱いに偏る傾向があるが，電子回路的な動作とともに理解しておくことにより，論理代数の数学的な理論との接点も明らかになるので，簡単に説明しておく．

ディジタル電子回路は，回路の二つの異なる動作状態(たとえば電流が流れているか/いないか，あるいは電圧値が高いか低いかなど)だけで表される．つまり，電流や電圧そのものの値にはよらないことが多い(**図1.7**)．

第2章 情報数学の基礎

第1章では，データ圧縮，符号化/復号化，暗号など，情報数学がどのような分野で使われるか，その概要を解説した．

ところで，長さ，重さ，時間などは，いずれも国際的に定められた標準量と比較して表されることは物理学の基本的な約束であるが，目に見えない情報ということになると，長さや重さのように標準になるものと比較することはできそうにない．情報数学の難しさはこの点にあるともいえよう．

本章では，情報数学の基礎概念として「情報を量る」に焦点を当てて，情報の定量的な表し方(情報量，確率など)について例題とともにわかりやすく解説する．

2.1 情報量とは

まずは，二つの対になった文章の例をいくつかならべて，情報量の大小を実感することから始めてみることにしよう．

[例1] (a) 8月の天気予報 「明日は晴」
(b) 8月の天気予報 「明日は雪」

[例2] (a)「旅客機が墜落炎上した」
(b)「オートバイが道路横の溝に落ちた」

[例3] (a) 不動産屋の宣伝文 「徒歩10分で駅に近くて便利」
(b) 居住する人の声 「徒歩10分で駅に近くて便利」

[例4] (a) 秀才のK君がL大に合格した．
(b) 天才のM君が留年した．

[例1]では，(b)のほうが情報量が多い．なぜなら8月は晴れるのが一般的で，雪が降るなどとは想像できない．

[例2]では，(a)のほうはニュース速報で報道すべき大事件であり，情報量が多い．

[例3]では，(a)は広告臭がしてどうもうさんくさそうで，(b)のほうが情報量が多い．

[例4]では，もちろん(b)のほうが情報量が多い．

[例1]～[例4]における情報量の大小を比較してみると，[例1]－(b)，[例2]－(a)，[例3]－(b)，[例

4]－(b)のほうがいずれも珍しいことで，情報量が多いことがわかる．このことを別な言い方にすれば，珍しいこととはほとんど起こらない(起こる確率が小さい)事柄であり，

起こる確率が小さい事柄ほど，情報量が多い

という統計学的な表現で記述される．つまり，情報量は起こる確率に対して単調減少する関数でなければならない．

また，情報量には加法性という性質も重要である．例をあげてみよう．Tさんの自宅はPアパート(4階建)の2階の右から5番目の部屋である(**図2.1**)．このアパートの各階には8室の部屋があり，1階の部屋には右から左へ順に11，12，…，18，2階は21，…，3階は31，…と番号が振られている．

たとえば，Tさんの自宅が「25号室である」という表し方(通報Z)は「2階にある」という通報Xと「右から5番目にある」という通報Yが合わさったものであると考えられる．このようなときは，二つの通報のもつ情報量を加えてほかの一つの情報量になるという関係，すなわち，

$$I_Z = I_X + I_Y \quad \cdots\cdots(2.1)$$

と表されることが妥当であろう．ここで，I_X，I_Y，I_Zはそれぞれの通報X，Y，Zの情報量を表す．

ところで，情報量の単位がビット[bit]であることはみなさんご存知のとおりであるが，二者択一(表と裏，0と1)のように二つの中から一つが起こる事柄を1ビットという情報量として定義している．つまり，確率pで起きる事柄の情報量I[bit]は，底を2とする対数関数として，

$$I = -\log_2(p)\ [\text{bit}] \quad \cdots\cdots(2.2)$$

と表すのである．たとえば，1枚の硬貨を投げたとき表か裏のいずれが出るのかという情報量は，表か裏の出る確率はそれぞれp=1/2であり，式(2.2)を適用すると，

$$I = -\log_2\left(\frac{1}{2}\right) = \log_2(2) = 1\ [\text{bit}]$$

と計算される．

同様にして，「Tさんの自宅は2階にある」という情報量I_Xは，4階建の中から一つの階を指定する確率p_Xが1/4であるから，

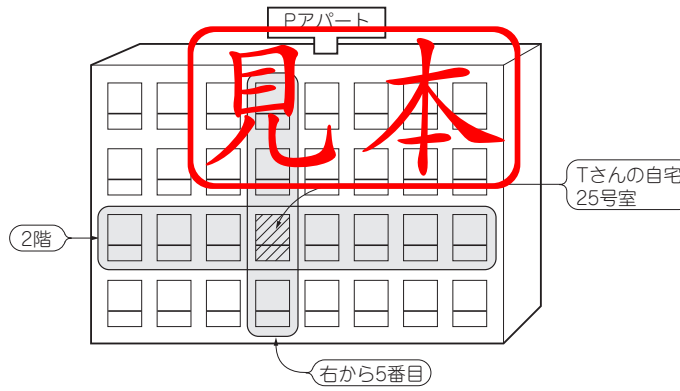


図2.1　情報量の加法性

第7章 誤り訂正符号の基礎

第1部では，情報理論的な側面（確率，統計など）や情報数学（符号などの基礎理論）を中心に説明してきた．主なキーワードとしては，以下のようなものが挙げられる．

情報量，確率，シャノン線図，平均情報量，情報エントロピー，通信モデル，通信容量，
符号化/復号化，ハフマン符号，ハミング距離，誤り検出/訂正，データ圧縮

ところで，IT（Information Technology：情報技術）革命の進展を支える技術の一つに，「誤り訂正が可能」というディジタル化にともなうシステム構成上のキー・テクノロジーがある．すなわち，ディジタル化した情報に何らかの誤りが生じたとしても，これを元の正しい情報に自動的に復元できるしくみで，放送や通信における品質向上に不可欠な技術といえる．さらには，記憶メディアにおける高密度記録化につながることから，CD，ディジタルVTR，DVDなどの実用化の基礎技術ともなっている．

こうした状況下において，誤り訂正符号の考え方をきちんと理解しておくことは，ディジタル放送，携帯電話などの移動体通信をはじめとする多様な分野で大いなる力を発揮することが期待される．そこで，第2部では「誤り訂正符号」にフォーカスして，符号理論に必要な数学，いろいろな誤り訂正符号の符号化/復号化などについて具体的に解説していくことにする．

なお，符号理論に必要な数学を第12章に総括してあるので，第7章の前に数学知識の再確認をしてから読み進めていくことをお勧めしたい．もちろん，順に読み進めていってもらってから，最終的に第12章を読んで「誤り訂正符号」の考え方を整理してもらってもよい．

本章では，具体的な誤り訂正符号の第一歩として「ハミング符号」を取り上げ，ハミング符号がパリティ符号を拡張したものであること，そして符号化と復号化のしくみについてわかりやすい形で説明する（付録Cを参照）．さらに，より汎用性の高い線形符号の導入を行う．

7.1 誤り検出/訂正のしくみ

ある一定のハミング距離を保った符号語の中から，いくつかの符号語を用いることは，別の視点から見てみると，任意に選んだ符号系列に適当な符号を付加して冗長度を与えることであるとも考えられる．つまり，情報符号と誤りの検出/訂正の検査符号を分ける考え方に，誤り検出/訂正の基本的な

コンセプトが隠されている．

たとえば，家庭の主婦が洗濯した後，物干しに干そうとして，「あれっ，靴下が足りない」とすっとんきょうな声を発したとする．これは，靴下の数をかぞえていたのではなく，靴下には「偶数である」という情報があることに基づいている（**図7.1**）．

この「偶数である」という情報は靴下に冗長度として付加されているもので，実はこれだけでも意外と多くの情報を誤りなく伝えることができる．これに反し，たとえば同時にハンカチがなくなっていたとしても，なかなか気づきにくいことも事実である．

靴下と同じように，情報符号に検査符号として「偶数である」という情報を常にもたせるようなしくみが誤り検出/訂正の考え方の基本であり，これが**パリティ検査**とよばれるものである．

一つの例を挙げてみよう．いま，**図7.2**の左の3ビットの情報符号にもう1ビットの検査符号を付加して0000からの**ハミング距離**（**重み**ということもある）を偶数になるようにする．それが右の誤り検出のための符号である．

また，**図7.2**の誤り検出符号を1列ずつ横に並べて8列にし，さらに1行ごとに偶数パリティを付けてみることにしよう（**図7.3**）．ここで，**図7.4**に示す○で囲んだビットが誤ったとすると，○を含んだ行と列（□の中）に含まれる‘1’の個数が奇数に変化することから，誤りが発生したビットを特定して誤りを訂正できることになる．

つまり，パリティ検査のビットを組み合せることで誤りの訂正が可能になる．ここに，ビットを特定するしくみとなる誤り訂正符号の基礎をなす考え方がある．

図7.1　偶数パリティで洗濯物をチェック

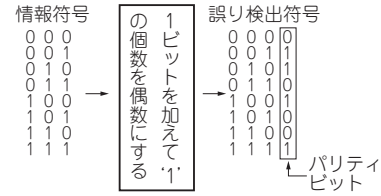


図7.2　パリティ検査とは

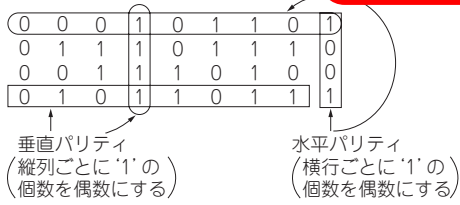


図7.3　垂直水平パリティ符号の構成

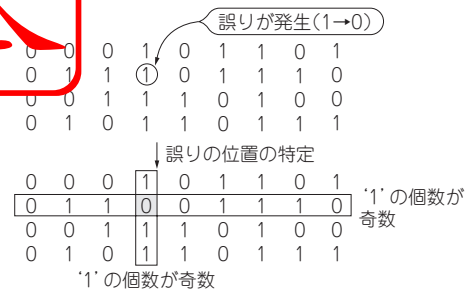


図7.4　垂直水平パリティ符号による誤り訂正例

第13章 暗号とは何か？

第2部では，誤り訂正符号を中心とする符号理論について解説してきたが，第3部では，新しいテーマとして「暗号」を取り上げることにする．

暗号という言葉は，10年ほど前まではあまり耳慣れないものであったようだが，ネットワーク化された情報化社会と呼ばれる現在においては，個人としての秘密やプライバシーの保護のために暗号は極めて重要な地位を占めている．安全で信頼性の高い情報化社会の実現のために必要不可欠な基盤技術の一つとして位置づけられている．

本章では，暗号の導入として「暗号とは何か」という話から始め，符号理論から暗号理論への橋渡しと同時に，暗号で用いられる用語を中心に解説する．一般に暗号は情報セキュリティの中核技術で‘守り’の技術として見られがちではあるが，リーマン・ショック後の不景気な時代においては，社会・経済を活性化するための‘攻め’の技術であることも知っておいてもらいたい．

13.1 暗号の役割

暗号は，かつて軍事や外交の機密文書の世界で使われていたためか，どうしても暗い過去を引きずっている感がぬぐいきれない．しかし，現在はどうかと言えば，暗号はもはや陰の存在ではなく，陽の当たる場所に出てきて堂々と表の世界を闊歩（かっぽ）しているのである．

これまで，これほど暗号が脚光を浴びたことはないのだが，今では高度情報化社会の根幹をなすキー・テクノロジとして位置づけられているようである．過去の暗いイメージを払拭し，安号（あんごうと読ませる造語で，安心を与えるための符号形態を表す）としての役割を果たすための技術が‘暗号’であると言うこともできようか．

最近，携帯電話，電子メールなどのデータ通信の急速な進展につれ，大切なデータを保護する，あるいは秘匿する必要性が日に日に高まっている．これまでは，紙幣であれば「透かし」，手書き署名や実印などが暗号の役割を担っていた．こうした実体のある暗号が，電子化された情報ネットワーク社会においては仮想的なものになってしまい，実体はなく‘目に見えない’ものになっているのである（**図13.1**）．

暗号の歴史は古く，バビロニア時代にさかのぼるが，ローマ時代の「シーザー暗号」，第2次世界大

図13.1　暗号の役割

戦を中心に用いられた日本の「紫(パープル)暗号」などのように，古代から近代にかけては軍事・外交などの用途が中心であった．つまり，ある限定された組織内で重要な情報を秘密裏に伝達するという枠組みの中で使用された．

ところが，20世紀後半に至り，コンピュータと通信の発展が暗号の世界に一大変革をもたらし，その適用領域の急速な広がりを見たのである．産業界では企業の秘密を守り，自治体では住民のプライバシーを保護するというように，暗号は一般社会で日常的に用いられるようになってきた．

また，手書き署名や実印(ハンコ)の印影もコンピュータ通信の中では，0と1の記号の並びにすぎなくなってしまい，簡単にコピーされてしまうことにもなる．こうした本人の確認や相手の確認，あるいは文書に対する署名の機能(「認証」という)も暗号技術で実現される．

さらに，1990年代の後半あたりから，インターネットをはじめとするパソコン通信の普及とともに，その上で電子取引や電子決済が行われるようになると，暗号のもつ秘匿性に加えて，金額や文書の改ざん防止，相手確認などの認証がきわめて重要となり，暗号の果たす役割が再認識され始めた．

このように，インターネット時代の成熟にともない，ネットワーク上を多種多様な情報が飛び交うわけで，個人や情報を正しく認証し，安心できる信頼の基盤を構築するためのキー・テクノロジが暗号なのである．わかりやすく言うと，暗号は情報のセキュリティ(安全性)を確保するために不可欠な技術であり，情報の改ざん，破壊，盗聴などの好ましくない事態を防止するためのものなのである．

13.2　暗号系のモデル

暗号理論は情報理論の中の一分野であり，その基本的な理論は，かの有名なシャノンが構築し，情報理論の立場から見た暗号系のモデルを示している(**図13.2**)．

図13.2は抽象的なので，ここではもっとも簡単な‘鬼退治’暗号(洒落た名前かも?)と名づけた例によって具体的に説明する．次に示す文字列は，ある勇者の名前を‘鬼退治’暗号とよばれる方法によって暗号化したものである．

NPNPUBSP(暗号文)

簡単とは言っても，容易に解読できるかといえばそうでもない．ただ勘のはたらく人なら「鬼退治」からの連想で，

ISBN978-4-7898-5301-9
C3055 ￥4000E
CQ出版社

定価 4,400円（本体4,000円）⑩

9784789853019
1923055040007

ディジタル信号処理シリーズ

見本