符号と暗号の技術について

技術解説

嵩　　忠　雄*

1．通信系に関する問題

図1は，典型的な通信系についての概念図である。左端の伝送あるいは記憶媒体は実際の伝送路とか記憶装置を表わす。それらの機能は，理想的には（雑音の影響，時間おくれ，波形歪等を別とすれば），入力信号に対し，それを空間的あるいは時間的に移動した信号を出力することである。空間的，あるいは時間的移動を除けば，“恒等写像”の実現を見指す。これは，一般の計算系と本質的に異なる所である。

雑音の性格は，通信路により異なり，例えば，衛星回線における熱雑音，半導体記憶における宇宙線，放射線の影響，磁気記憶装置におけるキズ，欠陥の影響等さまざまである。

通信系の問題は，通常，図1の右から，Ⅰ：情報源符号化，Ⅱ：通信路符号化，Ⅲ：信号波形（変復調）設計と分割し，単純化して考察される。

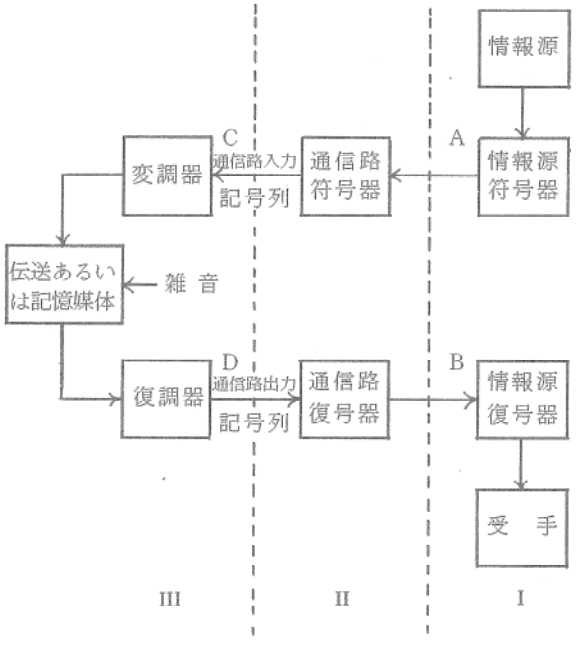

図1　典型的な通信系のブロック図

1.1　情報源符号化（冗長削減法）

情報源符号化問題では，図1の線ＡＢから左を，問題Ⅱ，Ⅲが理想的に解決したとして，理想的な恒等写像におきかえ，Ⅰの部分の解決に専心する(図2参照)。目標は，情報源の発生する通報がもつ“冗長”の削減法を考案し，通信路利用効率の改善による利得から冗長削減に必要な符号化，復号化のための代価を差引いた総利得をできるだけ大きくすることである。

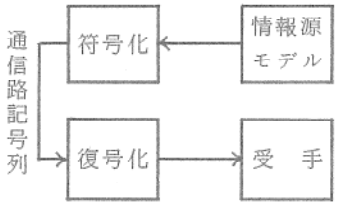

図2　情報源符号化問題

従来，冗長削減のために，情報源の発生する通報の統計的特性のみが利用されてきた。最近通報のもつ内部構造（例えば，文章の場合の文法構造）とか通報が表すもの（意味）にまで立入ることにより一層の冗長削減を目指す研究（知的符号化などと呼ばれている）が提唱されている[1]。これは究極の問題であり，個別の情報源についての地道な研究が待たれる。

1.2　通信路符号化問題

図1の線ＡＢより右を理想的離散的情報源（冗長のない情報源）と受手でおきかえ(図3参照)，図1の線ＣＤより左側は，通信路モデルでおきかえる。図1の変調器の入力デジタル信号（簡単のため2値とする）が0（又は1）であるとき，伝送媒体における一定時間幅の波形 w_0

*嵩　忠雄(Tadao KASAMI)，大阪大学基礎工学部情報工学科，教授，工学博士，情報工学

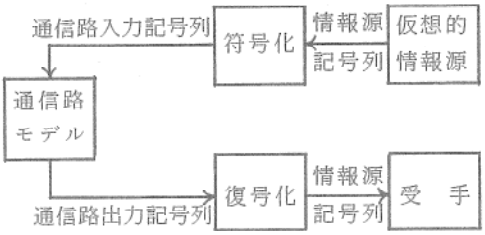

図3　通信路符号化問題

（又はw_1）に変換されて送信される。受信波形を入力として，雑音，歪等の特性を考慮してw_0かw_1のいずれが送信されたかを判定して，それぞれデジタル信号0か1を出力するのが復調器である。

一般に，0と1の系列に対して，波形w_0とw_1の系列が送られる。それを前後の受信波形とは無関係に，送信ビットごとに対応する受信波形に基づいて，送信ビットを0あるいは1と判定する復調法を硬判定であるという。硬判定の復調法は簡単ではあるが，本来利用可能な情報の一部を無視することからくる若干の損失は免れない。しかし，通信路符号化の問題では，通常，硬判定復調法を前提として，CDより左の部分を，送信ビットに対して復号器の出力ビットが異なる（誤りが起こったという）確率がいくらであるかと云った単純な通信路モデルでおきかえ，変復調問題と分離する。このように単純化することによって，誤り制御に関する問題の本質が浮かび上がる。

簡単な例として，計算機の半導体記憶装置において，1語の書きこみ，読み取りにおいて誤りが起こる確率が以前に起こった誤りとは独立で10^{-8}であるとする。1秒間に約10^6回，書き込み，読み出しを行なうとして，10分間に一度も間違いが起こらない確率は，

$$(1-10^{-8})^{6\times10^8} \approx 2.5\times10^{-3}$$

これでは，実用にならない。そこで，要求される信頼性を実現するためには，1）記憶媒体自体の信頼性を向上させる，2）誤り制御符号化方式を採用する，あるいはそれらの組み合わせの中でコスト的に有利なものが選ばれる。現在，2）の技術が成熟しつつあり，大部分の記憶装置において，2）の技術が利用されている。通信回線についてもほぼ同様である。

1.3　冗長の導入

情報源符号化の問題は，冗長の削減であったが，逆に誤り制御符号化の問題は，次に述べるように，如何に合目的に冗長を導入するかの問題である。例えば，2^{26}個の異なる“通報”を表わすために，最小26ビットが必要である。26ビットのみを用いると，1ビット誤っても異なった通報として受け取られる。仮に，もう1ビット余分に用い，最初の26ビットを2^{26}個の通報を表すのに使い，残りの1ビットを，最初の26ビット中1の個数が偶数なら0，奇数なら1と決めたとする。一般に，nビットを用いて，M個の通報を表わすとき，$\log_2 M/n$を伝送効率という。伝送路または記憶装置の利用効率を表わす尺度として用いられる。この例では，伝送効率が26／27であり，可能な0と1の組2^{27}個のうちの半分2^{26}個を通報を表わすのに用い，残りの半分は用いない。こうすることによって，伝送あるいは記憶媒質中で，仮に27ビットのうち1ビットの誤りが起こったとき，受手は誤りが起こったことを知ることができ（誤りの検出という），再送なり，再書き込みを要求することができる。

上の例では，誤りが起こったことを知っても，27ビットのうちどのビットが誤っているかは分からない。2ビット以上の誤りが起こる確率は十分小さいとし，1ビットの誤りがどのビットに起こっても正しく復元（誤り訂正という）したいとする。余分に5ビットを用いることにする。伝送効率は26／31となる。全体で2^{31}個の0，1の組のうち，通報を表わす2^{26}個（符号語と呼ぶ）をどのように選ぶべきかを考える。

$\bar{u}$，$\bar{v}$を異なる任意の符号語とする。$\bar{u}$に1ビットの誤りが起こった0，1の組（31個ある）に$\overline{u}$自身も含めた32個の0，1の組からなる集合を$N(\bar{u})$と書く。同様に$N(\bar{v})$を定義する。$N(\bar{u})$と$N(\bar{v})$が共通な0，1の組$\overline{w}$を含めば，$\bar{u}$に1ビットの誤りが起こって$\overline{w}$になったのか$\bar{v}$に1ビットの誤りが起こって$\overline{w}$になったのか区別がつかない。逆に，どの$\bar{u}$，$\bar{v}$についても，共通な0，1の組を含まないとすれば，仮に1

ビットの誤りが起こっても，どの符号語に1ビットの誤りが起こったかが一意にきまる。ハミングは，上の条件を満たす符号語を選ぶ簡単な方法を考案した。

なお，5ビット余分に使うと云ったが，少なくとも5ビットは必要である。各符号語$\bar{u}$について，$N(\bar{u})$は互いに重ならないから，少なくとも$2^{26}\times 32 = 2^{31}$個の0，1の組が必要であることが分かる。上述のような方式（誤り訂正・検出符号）は，現在半導体記憶装置をはじめ広く実用化されている。

1.4 符号理論

信頼性の向上のために，冗長を導入し伝送または記憶媒体の利用効率を犠牲にするというトレードオフの関係を，符号化，復号化のためのコストを考慮せずに，解明したのはシャノンである。

nビットずつまとめて（長さnのブロックという）送るとして，1ブロック中の少なくとも1ビットが誤って受手に渡される確率（ブロック誤り確率）で信頼性を測るとする。Rを通信路の統計的性質のみに依存してきまる定数（通信路容量）より小さい任意の実数とし，Pを任意に小さい正の実数とする。シャノンは与えられたR，Pに対して，nを十分大きくとれば，伝送効率がR以上，ブロック誤り確率がP以下であるようなブロック符号化が存在することを示した。Rを通信路容量より大きく選べば，上の性質を満たす符号化は存在しないことも知られている。

シャノンが示したのは，巧妙な存在証明であり，具体的にどのように符号化すればよいかは示していない。また，nを大きくとれば，符号化，復号化のコストが急激にふえるから，実用化するためには，これらのコストも考慮した具体的な符号構成法を考案しなければならない。この目標に向かって，シャノン以降精力的な研究（符号理論と呼ばれる）が行なわれてきた。

信頼性に対する要求を，機構化のコストの小さい単純な符号化によって実現すれば，一般に，通信路容量よりかなり低い伝送効率に甘んじなければならない。一方，伝送効率を上げようとすれば，機構化のコストの上昇を覚悟しなければならない。すなわち，通信路コストと計算コスト（符号化及び復号化操作のためのハードウェアのコスト）はトレードオフの関係にある。ハードウエア技術の急激な進歩により，複雑な符号化，復号化回路が一つのチップ中におさまるようになり，より高度な符号化方式が実用化されるようになった。

1.5 符号化変調方式

符号化問題と変復調問題とが分離され，それぞれの問題として研究され成果があげられてきたが，その境界において両者の統合を目指した研究がはじめられたのは自然な成行きである。とくに，誤り制御符号構成の方法を信号波形設計に持ち込んだ符号化変調方式の研究が活発に行なわれており，硬判定による損失の回避法として一部実用化されている[2]。今後の進展が期待される。

1.6 計算系への応用

信頼性向上のための符号化法を，伝送や記憶系のみでなく，計算機の演算装置，制御装置などに適用できないかと誰しも考えるであろう。事実，加算回路用の符号が研究されている。他方，論理和とか論理積のような非線形の基本関数については，例えば同じ演算回路を奇数回重複して作り，それらの出力の多数決をとるといった単純であるが非経済的な方法以外有効な方法がないことが示されている。一般の論理回路について，信頼性と回路の“冗長度”との間に，シャノンの定理が示すような関係は成立しない。高度の信頼性が要求される場合，1）ハードウェアの多重化，2）テスト回路の内蔵，3）入出力データに若干の冗長をもたせて誤りを検出し再試行する方式などが用いられる。

2．暗　　号

暗号の歴史は古いが，計算機という忠実で有能な助手が出現してから，新しい技術の展開が見られる。暗号の技術の基本は，知っておれば，ある計算を容易に実行できるが，知らなければ，その計算の実行が困難であるような“情報”（鍵

と呼ばれる）を，当事者以外には秘密に発生することである。

2.1 対称鍵暗号（共通鍵暗号）

最初に，昔からある暗号化方式（対称鍵方式又は共通鍵方式と呼ばれる）を考える。AからBへ秘密の通報文Mを送る場合，AとBは一つの鍵kを共有していなければならない。AはM（暗号化されていない通報文を平文と呼ぶ）を鍵kによって，暗号文e(k，M）に変換してBへ送る。この変換を暗号化と呼ぶが，暗号化の操作は，kとMを知れば，実行が容易でなければならない。また，同じkに対して異なる平文には異なる暗号文が対応する。暗号文からもとの平文Mを求めることは，鍵kを知っておれば，容易に実行できるが（復号化という），鍵kを知らずに，傍聴，盗聴，その他の方法で入手した暗号文から平文を求める（暗号解読と呼ばれる）のは，実用上困難でなければならない。

米国商務省標準局が1977年に採用した規格DESでは，鍵の長さ56ビット（パリティ検査用に余分に8ビットを用いる）によって64ビットの平文を64ビットの暗号文に変換する暗号化関数が具体的に指定され，LSIチップも市販されている。

対称鍵暗号の解読問題の理論を創ったのもシャノンである。暗号化は一種の冗長付加操作と考えられる。鍵のビット数を平文のビット数で割って得られる比をr_kと書く。一般に平文は冗長をもっているが，その冗長度をr_pと書く（nビットの“意味のある”平文の数は約$2^{(1-r_p)n}$）。nr_kビットの鍵を無作為に選ぶとする。すべての可能な鍵の選び方を考えたとき，一つの平文に対して，高々$2^{r_k n}$個の暗号文が対応するから，意味のある平文に対応する暗号文の総数は，$2^{(1-r_p)n+r_k n}$以下である。もし，$r_k < r_p$なら，nが大きくなるとき，異なる平文M_1，M_2に対して，鍵k_1，k_2があって，$e(k_1, M_1) = e(k_2, M_2)$となる可能性は，平均して十分小さくなる。すなわち，一つの暗号文に対して，平均して，一つ以下の意味のある平文が対応する。従って，暗号化関数eを知っており，与えられたビット系列が意味のある平文であるか否かの判定能力があれば，原理的には（所要計算時間，記憶量を無視すれば）解読可能であることが分かる。安全性のためには，同一鍵をあまりくり返し使わないとか，平文からできるだけ冗長を削減し暗号化する等の注意が必要である。

2.2 非対称鍵暗号（公開鍵暗号）

対称鍵暗号の欠点は，共有鍵の配布の問題である。秘密に配布するにはコストがかかる。また自己以外誰も信用できないとすればどうすればよいか？受信者のみが秘密鍵をもつ方式（非対称鍵方式又は公開鍵方式と呼ばれる）が，デフィー・ヘルマンによって，1976年に提案された。その後，いくつかの非対称鍵暗号が提案されたが，その中には安全性（暗号解読の困難さ）の条件が満たされないことが示されたり，安全性について疑念がもたれているものもある（このこと自体，問題の微妙さを示している）。現在まで残っている代表は，リベスト・シャミヤ・アーデルマンが1978年に提案したRSA暗号とその変形である。RSA暗号を例として説明する。

(1) 鍵の生成：各受信者は秘密裡に異なる素数pとqを選ぶ（例えば，10進100桁位の数を無作為に選び，素数判定アルゴリズム（効率的方法が知られている）で判定し，素数が求まるまで繰り返す。安全上の考慮から，さらにp－1，q－1が大きな素因数を含む必要があるので，これらの条件を満たすp，qが見つかるまで繰り返す。素数の分布定理から繰り返し回数の平均値が押さえられる）。$n = pq$とおく。秘密鍵（復号化鍵）k_dを$(p-1)$と$(q-1)$の最小公倍数Lと互いに素になるように選び，また公開鍵（暗号化鍵）k_eを，$k_e k_d \equiv 1 \pmod{L}$を満たすように求める。（x，y，zを整数とし，x－zがyで割り切れることを，$x \equiv z \pmod{y}$と書く。k_dとnから，k_eを求める効率的方法が知られている。）受信者は，nとk_eを名前（ID）と共に公開ファイルに登録する。

平文と暗号文をn－1以下の非負整数で表わす。mを平文，cを暗号文として，

(2) 暗号化：$c \equiv m^{k_e} \pmod{n}$，

(3) 復号化：$m \equiv c^{k_d} \pmod n$。

暗号化，復号化は，比較的容易に実行できる。平文と暗号文は一対一対応であり，n，k_eは公開されているから，秘密鍵k_dを知らずに，暗号文から平文を求めること（暗号解読）は，逆関数の計算であり，原理的に可能である。暗号化の鍵が秘密である対称鍵暗号と比べて，安全性を保証することがより難しい。適当な条件の下に，ＲＳＡ暗号解読は，公開されたnから秘密の素数pとqを求める素因数分解問題と同程度の難しさであることが示されている。なが年にわたる専門家の努力にも拘わらず与えられた正整数に対して，その素因数を求める能率的な方法は知られていない。

現在提案されている非対称鍵暗号の鍵の生成，暗号化，復号化の操作は，対称鍵暗号における操作と比べ，同程度の安全性を確保しようとするとかなり複雑になる。例えば，ＲＳＡ暗号では数百ビットの鍵を選ぶ必要があるといわれている。従って，高速処理が必要な多量のデータの暗号化には，対称鍵暗号を用い，その共通鍵の配布に非対称鍵暗号を利用する方式が実用的である。なお，公開ファイルに対して，他人を偽って登録したり変更したりするのを防止する必要がある。

2.3 デジタル署名

ＲＳＡ暗号などは，デジタル署名としても用いることができる。前提として，意味のある平文は十分の冗長をもっていると仮定する（必要なら，冗長をもたす。宛先，発信人の名前，日付，通し番号等が書いてあるとする）。

発信人Ｂ（秘密の復号鍵を$k_{d,B}$と書く）は平文ｍのＢによるデジタル署名文 $s \equiv m^{k_{d,B}} \pmod n$ を作り，自己の名前と共にＡに送る。ｓを受け取ったＡは，公開ファイルからＢの公開鍵$k_{e,B}$を知り，$s^{k_{e,B}} \equiv (m^{k_{d,B}})^{k_{e,B}} \equiv m^{k_{d,B} k_{e,B}} \equiv m \pmod n$ を得る。ｍがＢからの通報として意味のあるものであれば，Ａはｓをｍのデジタル署名として受け取る。暗号が安全であるとすると，Ｂの秘密鍵$k_{d,B}$を知っているのはＢのみであり，意味のある平文ｍについて $s \equiv m^{k_{d,B}} \pmod n$ を作り得るのは，Ｂ以外にないと考えられる。

2.4 ＩＤ問題

ＩＤ問題は，サービス機関が，他人になりすました要求でないことを確かめる問題であり，デジタル署名問題の特殊な場合である。現行の暗証番号方式は，新聞記事になったように比較的容易に破られる。

最近提案された方式[4]の筋書きを紹介する。次の条件を満たす問題のクラスＱを前提とする。

(1) 答ａを先に選び，ａを正答とする問題$q \in Q$を作成するのは容易である。

(2) $q \in Q$に対する答a'が正答か否かは誰でも容易に判定できるが，ｑの作成者以外がｑの正答を求めるのは困難である。

利用者Ａは，一つの$q_A \in Q$を作成して，Ｓに登録し，q_Aの正答a_Aを秘密に保持し，要求を出すとき，名前と自由に選んだパラメータをＳに渡す。Ｓは名前Ａ，パラメータｒの要求に対して，q_Aとｒに依存して，次の条件Ｔを満たす質問のなかから無作為に１つｔを選んで，要求者Ａ′に出す。ｒにも依存させるのは，Ｓによる一方的誘導尋問を避けるためである。

Ｔ：ｔに対する答が正答か否かは誰でも容易に判定できるが，q_Aの正答を知らずにｔに正答するのは困難である。

Ａ′が正答を与えたら，要求を認める。一方，利用者ＡがＳの質問に正直に答えることによって，秘密a_Aを知っていることのみをＳに示し，a_Aについての“実質的な情報”をもらすおそれがないことが必要である。

なお，２．では計算が容易，困難，同程度に困難，実質的な情報をもらさない，などあいまいな表現を使ってきたが，それらの厳密な定義については，例えば，文献４）～６）を参照されたい。

参考文献

１）原島　博：“知的情報理論の課題”，ＳＡ-4-1，昭和63年電子情報通信学会春季全国大会。

2）笠原正雄：“符号化変調方式”Ⅰ～Ⅲ，電子情報通信学会誌，72巻1～3号（1989）.

3）土居範久，小山謙二：コンピュータ・セキュリティ 共立出版（1986）.

4）U. Feige, A. Fiat and A. Shamir：“Zero knowledge proofs of identity”, J. of Cryptology 1，pp77—94（1988）.

5）嵩　忠雄，藤原　融：“暗号アルゴリズムと計算量の理論”情報処理，vol．25，No．6，pp547—553（1984）.

6）嵩　忠雄：“問題の難しさをどのように定義するか”コンピュータソフトウェア，vol．3，No．5．5 pp．92—100（1986）.