・・・・3・・・・

暗号入門じみたもの

60 回生 山本真吾

いつしか部誌を書かねばならない季節がやってきて、過ぎて行きつつある[1]。昨年予告したように、暗号にまつわる小話を書いてみようと思う。ほんの少しの暇と数学の勘をポケットに突っ込んで駆け出そう。夜明けまでまだ三時間ある。

アリスとボブには伝えたい言葉がある

とりあえず、状況をわかりやすく説明するために、登場人物をでっち上げておく。名前はアリスとボブ[2]。人間である、以外の設定は勝手に考えてくれて構わないが、本題とは何の関係もない。

アリスは、ボブに伝えたいことがあるとしよう。それは明日の天気に関する国家機密かもしれないし、コロネをどこから食べるか、みたいなつまらない話かもしれないが、そのあたりは重要ではない。とにかく他人に聞かれたくない内容なのであるが、同時に今すぐ伝えたい内容でもある。電話をすればいい？なにを言っているんだ、今この瞬間も彼らの秘密を暴こうと全世界の諜報機関が盗聴を試みているというのに、そんな無防備な真似ができるものか。

こういった非常に特殊な状況で利用されるのが、暗号通信である (冗談だ)。本稿の目的は、いくつかのよく知られた暗号アルゴリズムをちょっぴり数学風に解説することである。数学は苦手なので、厳密さに関しては目をつぶってもらいたいところだが、本当に目をつぶってしまっては読めない。ディレンマである。

アリスとボブは秘密の手紙を書く

さて、現代では暗号は個人の機密保持などにも利用されているが、本来暗号は軍事機密保持のために利用されるものであった。その歴史は古く、紀元前十九世紀くらいまでさかのぼることができるらしい。ヒエログリフで書かれた文章の中に、標準以外の文字を使って書かれたものがあるそうだ。

●● Caesar 暗号 ●●

古代に利用された暗号の中でとりわけ有名なのは古代ギリシャだかローマだかの指導者 Julius Caesar[3]が利用した「Caesar 暗号」だろう。彼は時に、アルファベットを 3 文字ずらして[4]手紙を書いた。こうすることで、万一手紙が敵の手に落ちても簡単に内容を知られるということはない。が、あらかじめ 3 文字ずらされていることを知ってい

[1] まったく申し訳ない。

[2] こういう場合に伝統的に使われる名前というものがあって、それが Alice, Bob, Chris, David... なのだ (諸説あるけれど)。頭文字によるのだろう。

[3] ジュリアス・シーザー。「ユリウス・カエサル」の方が一般的かもしれないが、「カエサル」ってかっこ悪いから嫌いだ。

[4] `a`→`D`, `b`→`E`, ..., `z`→`C` という具合に。

る味方には復号[5]することが可能だ。このとき、ずらされた文字数 3 は暗号鍵としての意味を持つ。

残念ながらこのタイプ[6]の暗号は解読が非常に容易だ。何しろ暗号鍵は 26 通りしか存在しないし、うち 1 通りは平文と同じになるから除外できるので実質 25 通りであるから、人間の手でその全てを調べることが十分可能で、しかも容易だ。

なお、3 文字ではなく 13 文字ずらしたものが、現在でも ROT13 などと呼ばれて使われることがある。もちろん、機密保持のためなどではなく、主にネタバレ防止などのためである。なぜ 13 なのかというと、アルファベットが 26 文字なので、もう一度暗号化すると元に戻るからである。暗号化と復号を同じ作業で行えるということだ。

●● 換字式暗号 ●●

Caesar 暗号では文字の置き換えに規則性があったが、もう少し複雑にして、文字をまったくばらばらに置き換えてしまう。このように文字を一文字または複数文字ごとに置き換えるタイプの暗号を総称して換字式暗号[7]と呼び、一文字ずつ置き換えるものを特に単一換字式暗号と呼ぶ[8]。一般の換字式暗号は Caesar 暗号よりも解読がやや困難であるが、まったく残念なことに現在ではやはり非常に容易である。

単一換字式暗号の場合について説明しよう。文字を置き換えるだけという性質から、平文中の同じ文字は暗号文中でも同じ文字に置き換えられる。ということは、それぞれの文字の出現頻度は一定に保たれるということだ。標準的な英文における各文字の出現頻度はすでに調べられているから、あとは暗号文中の文字の出現頻度と照らし合わせて見当をつけていけばいい[9]。このような手法を使うと素人でもすぐに解読できてしまう。

さらに、換字式暗号にはもうひとつ弱点が存在する。それは、鍵が複雑になるということだ。Caesar 暗号ではずらした文字数だけでよかったが、この場合はそういった単純な規則性が存在しないから、どのように文字を置き換えたのか示す必要がある。それはランダムに決められた文字の並びだったり本に書かれた文章の一節だったりしたわけだが、とにかくこういった特徴が鍵配送問題 (後述) をより困難にしてしまった。

アリスとボブは数学の女王と遊ぶ

●● 数学してみる ●●

さて、前節であげた Caesar 暗号を数式で表してみよう。数式で表すということはつまり暗号を数学的に取り扱うということだ。26 文字のアルファベットを 0 から 25 までの数字で表すことにする[10]。と、Caesar 暗号は次のように書き表せる。

$$C = (P + 3) \bmod 26$$
$$P = (C - 3) \bmod 26$$

---

[5] 暗号文を平文 (ひらぶん) へと戻す作業。ただし、鍵を知らない第三者が無理矢理行う場合は「解読」と呼ばれて区別されることが多い。

[6] Caesar が利用したのは 3 文字ずらすものだが、それ以外の文字数のものも Caesar 暗号と呼ばれる。

[7] かえじしきあんごう

[8] Caesar 暗号も単一換字式暗号である。

[9] たとえば暗号文中でもっとも出現頻度の高い文字は標準的な英文で最も出現頻度の高い文字である E を置き換えたものではないか、など。ただし、あえて E をまったく使わずに書かれた小説も存在するから、常に有効とは限らない。

[10] a→0, b→1, ..., z→25

P は平文、C は暗号文をあらわす。あと、謎の記号 mod は、剰余を意味する。たとえば $7 \bmod 3 = 1$ だし、$-2 \bmod 5 = 3$ だ[11]。つまりこれらの式は

- 暗号化するときは 3 を加えて 26 で割った余りをとる
- 復号するときは 3 を引いて 26 で割った余りをとる

ということを表している。もうちょっと数学してしまおう。暗号化のプロセスを関数 $f(x)$ だと考えると

$$f(x) = (x + 3) \bmod 26$$
$$f^{-1}(x) = (x - 3) \bmod 26$$

となる[12]。

●● mod と合同式 ●●

実は、mod にはもうひとつ使い方がある。どういう風に表現したものかわからないので実例を挙げておく。

$$3 \equiv 8 \pmod 5$$

意味は伝わるだろう。3 も 8 も mod5 した値が等しいということ[13]で、「3 と 8 は 5 を法として合同である」と読む。

法が等しい合同式に関する規則をいくつか並べておこう。なんだか本筋から外れ気味に見えるかもしれないが、そもそも本筋など気にしたためしがないし、あとで数学的解説をするときに必要になるので続ける。

加算と減算

$$3 \equiv 8 \pmod 5$$
$$1 \equiv 6 \pmod 5$$

合同式を複数並べて加算することができる。上記の例だと

$$4 \equiv 14 \pmod 5$$

となり、正しいことがわかるだろう。同様に減算も可能なのだが、いちいち式を書くのが面倒なので自分で確かめてほしい。あと、証明も割愛する。

乗算と累乗　さらに乗算も可能だ。

$$\begin{array}{rrrl} & 7 \equiv & 12 & \pmod 5 \\ \times & 4 \equiv & 9 & \pmod 5 \\ \hline & 28 \equiv & 108 & \pmod 5 \end{array}$$

---

[11] 二つ目の式はいささか直感的ではないかもしれないが、-2 に 5 を加算してみると納得できるだろう。

[12] $f^{-1}(x)$ は逆関数、つまり復号を表す。それにしても、-1 乗とは気が利いていると思う。

[13] 難しい言葉では「3 と 8 は 5 を法とする同じ剰余類に属している」と表現するらしい。

ということで、累乗も可能だ。

$$7^4 = 2401 \equiv 20736 = 12^4 \pmod 5$$

これは少しばかり便利な性質だ。たとえば、10^{100} を 7 で割った剰余を求める場合を考える[14]。これは 10^{100} 日後の曜日を知りたい場合に役に立つ (冗談)。

まず、

$$\begin{aligned} 10 &\equiv 3 \pmod 7 \\ 10^6 \equiv 3^6 &= 729 \pmod 7 \\ &\equiv 1 \pmod 7 \end{aligned}$$

を求めておく。6 という数字は右辺を 1 にすべく天啓に導かれて得た。あとはこれをさらに 16 乗して

$$(10^6)^{16} = 10^{96} \equiv 1^{16} = 1 \pmod 7$$

さらに足りない分を計算し、掛ける。

$$\begin{array}{rl} & 10^{96} \equiv 1 \pmod 7 \\ \times & 10^4 \equiv 4 \pmod 7 \\ \hline & 10^{100} \equiv 4 \pmod 7 \end{array}$$

ということで、今日が火曜日なら 10^{100} 日後は土曜日である[15]。やった、週末だ!

原始元　素数 7 があったとする。このとき、7 を法としたときの 3 の累乗を考えると、

$$\begin{aligned} 3^0 &\equiv 1 \\ 3^1 &\equiv 3 \\ 3^2 &\equiv 2 \\ 3^3 &\equiv 6 \\ 3^4 &\equiv 4 \\ 3^5 &\equiv 5 \\ 3^6 &\equiv 1 \pmod 7 \end{aligned}$$

となる。1 から 6 まで一巡しているのがわかるだろう。こういう場合、3 は $\mathbb{Z}_7$ の原始元である、という[16]。p が素数ならば、$\mathbb{Z}_p$ には必ず原始元が存在するらしい。ああそうだ、5 も $\mathbb{Z}_7$ の原始元だ。

逆数　たとえば、3 に 1/3 をかけると 1 になる。こういう場合、1/3 は 3 の逆数である、という。ふつう逆数といえば、分子と分母を逆転させたものであるが、さて、mod$p$ になると話は別になってくる。たとえば、

$$(3 \times 5) \bmod 7 = 1$$

---

[14] ちなみに、この値 10^{100} を googol と呼ぶ。検索サービス google の名はこれに由来するが、開発者がスペルを間違えたために微妙に異なっているらしい。本当だろうか。

[15] グレゴリオ暦が使用されていることを前提としているから、もしかするとあまり意味のない計算なのかもしれない。

[16] 本当はもう少し厳密に書きたかったのだけれど、ここには十分な余白がないし、決闘前夜なので時間がない。

となるから、5 は 3 の逆数となっている。

●● もっと数学してみる ●●

さて、Caesar が使ったのは 3 文字ずらすものだったが、それ以外の文字数のことも考えてやりたい。ということで、

$$f(x) = (x + K) \bmod 26$$
$$f^{-1}(x) = (x - K) \bmod 26$$

としよう。ここで K は鍵である[17)]。

ところで、このままでは暗号として心もとない。鍵が 26 種類しかないのなら、誰だって復号できる。そこで、複数の文字をまとめて暗号化することを考えよう。たとえば二文字ずつの場合、文字の組を aa, ab, ac, ..., zy, zz という順に並べて、0 から $26^2 - 1 = 675$ までの番号をつける。この新しい暗号は $f'(x) = (x + K) \bmod 676$ と表され、鍵 K は 0 から 676 までの値をとることができる。鍵の可能性が増えたということは第三者による復号が困難になったということである。

しかし残念なことに、このタイプの暗号はこの程度では簡単に解読されてしまう[18)]。かといって鍵の種類をもっと増やすと計算に時間がかかりすぎる[19)]。ということで、シーザー暗号は現在ではほとんど探偵ごっこか上述の ROT13 のような用途にしか使われていない。

アリスとボブは鍵を共有する

今度は問題を違った方向から考えてみよう。とりあえず、第三者にやすやすと解読されることのない暗号があったとする。と、どのように鍵を安全に伝えるかが重要な問題になってくる。これは鍵配送問題といって長い間解決されなかった問題であるが、戦後になって暗号が数学的に研究されるようになり、暗号界に「公開鍵暗号」という革命が起こった。

そんなわけで 1976 年に提案されたのが Diffie-Hellman 鍵共有[20)]である。これは、盗聴されている可能性のある経路を通して二人の人間が同じ情報を共有するためのプロトコルである。盗聴している第三者はその情報を共有することができないようになっている。つまり、その情報を暗号鍵として利用することができるということだ。すばらしい。

それなりに大きな素数 p と、その原始元 n を用意する。これらは公開して構わない。札束の裏に書いて東京タワーからばら撒いても構わない。

次に、アリスとボブはそれぞれ秘密の数 A_a, A_b をランダムに選ぶ ($0 \leq A_a, A_b \leq p - 2$)。選んだら、それぞれ

$$B_a = n^{A_a} \bmod p$$
$$B_b = n^{A_b} \bmod p$$

を計算し、相手に送信する。一方で A_a や A_b は秘密にしておく。

最後に、アリスは

$$K_A = B_b^{A_a} \bmod p$$

[17)] 本来は暗号鍵と復号鍵を区別して K_E, K_D としなければならないところだが、Caesar 暗号の場合どうせ同じなのでまとめて書いた。

[18)] コンピュータのある現在ならなおさらだ。

[19)] と、思う。

[20)] ディフィー・ヘルマン

ボブは

$$K_B = B_a^{A_b} \bmod p$$

を計算する。生真面目に計算するとわかると思うけれど、

$$\begin{aligned} K_A &\equiv B_b^{A_a} \\ &\equiv \left(n^{A_b}\right)^{A_a} \equiv n^{A_a A_b} \equiv \left(n^{A_a}\right)^{A_b} \\ &\equiv B_a^{A_b} \equiv K_B \pmod p \end{aligned}$$

となる。これでアリスとボブは同じ値を共有することができた。鍵配送問題の解決である。

本当にこれで問題ないのだろうか？たとえば B_a, B_b が盗聴されていたとしよう。n と p も既知だとする。これらの値から K_A あるいは K_B を求めることはできないのだろうか？

実は、できなくは無い。のだが、これは離散対数問題といって、現時点では効率的な解法が存在しない[21]。あるいは、$n^{A_a} = B_a$ とわかっていて、しかも $0 \leq A_a \leq p-2$ だから、総当りすれば正しい値にたどり着くことはできる。ただし、p は非常に大きな数[22]なので、そうやすやすと破られたりはしない。

アリスとボブは鍵を公開する

1977 年に Rivest, Shamir, Adleman によって発明された暗号が RSA 暗号で、現在でもいろいろな形で使われている。このあたりまで説明して終わりにしたい。我ながらよくがんばったものだ (気が早い)。

●● 鍵の公開と暗号化の手順 ●●

公開鍵暗号は、従来の暗号とはすこしプロトコルが異なっている。通信の手順を簡単に説明する。アリスがボブにメッセージを送ることを想像してほしい。

1. アリスは、どこからかボブの公開鍵を探してくる。
2. アリスは、送信するメッセージをボブの公開鍵で暗号化する。
3. メッセージを送信する。
4. 受信したボブは、自分の秘密鍵でメッセージを復号する。
5. 黒ヤギさんたら読まずに食べた。

最後の手順は不要である (真顔)。

これを見ると、公開鍵暗号には「公開鍵」と「秘密鍵」の二種類の鍵が必要なことがわかると思う。二種類の鍵は対になっていて、公開鍵は暗号化、秘密鍵は復号を行う[23]。

●● もう少し数学 ●●

数学が嫌いな読者が読んでいるならば、この節は飛ばしたほうがいいかもしれない。が、そこまで難しくはないと思う。

---

[21] 量子コンピュータが実現すれば解けるらしいが、よく知らない。

[22] それこそ googol とかでも足りないくらいに大きい。

[23] 実は、署名をする際などには逆のことが行われているのだが、混乱するといけないので黙っておく。

Euler の ϕ 関数　自然数 n より小さく、n と互いに素な自然数の個数を $\phi(n)$ とあらわす。素数 p, q に対して $\phi(pq) = (p-1)(q-1)$ となる。

Fermat の小定理　Fermat[24]といえば最終定理しか思いつかない人も多いだろうが、別に彼がそれしか仕事をしていないわけではない。そんな彼の定理の一つに Fermat の小定理がある。素数に関する定理で、以下のようなものだ。

$$a^{p-1} \equiv 1 \pmod{p} \quad (p \text{ は素数、} a \text{ と } p \text{ は互いに素})$$

現在では、この定理の対偶が素数判定に利用されている。あと、下で述べるけれど Euler が拡張したものが RSA 暗号では重要な意味を持つ。

Euler の定理　a, n が互いに素な自然数であるとき、

$$a^{\phi(n)} \equiv 1 \pmod{n}$$

となる。n が素数ならば、$\phi(n) = n - 1$ からこれは Fermat の小定理と一致する (あたりまえだ)。

●●　鍵の算出と暗号化　●●

まず、それなりに大きな素数 p, q を選び、$n = pq$ と、$\phi(n) = n - (p + q) + 1$ を計算する。次に、$\phi(n)$ と互いに素で $\phi(n)$ より小さい自然数 e を選び[25]、$\bmod \phi(n)$ における e の逆数 d を計算する。このとき、

$$\text{公開鍵 } K_E = \langle n, e \rangle$$
$$\text{秘密鍵 } K_D = \langle n, d \rangle$$
$$f(x) = x^e \bmod n$$
$$f^{-1}(x) = x^d \bmod n$$

となる。

●●　確認してみる　●●

なんだか妙にきれいにまとまっているが、実際のところ正しく機能するのだろうか？確かめることにする。目指す場所は、$f^{-1}(f(x)) \equiv x \pmod{n}$ だ。とりあえず代入してみる。

$$f^{-1}(f(x)) \equiv x^{de} \pmod{n}$$

$de \equiv 1 \pmod{\phi(n)}$ から、

$$\equiv x^{(X\phi(n)+1)} \pmod{n}$$
$$\equiv x^{X\phi(n)}x \pmod{n}$$
$$\equiv (x^{\phi(n)})^X x \pmod{n}$$

n, x が互いに素ならば、Euler の定理から、

$$\equiv 1^X x \pmod{n}$$
$$f^{-1}(f(x)) \equiv x \pmod{n}$$

---

[24] フェルマー。弁護士にしてフェルマー予想の主犯。

[25] あまり大きいと面倒なので 17 だとか 65537 がよく使われる

となる。n, x が互いに素でない場合にも、いろいろとこじつけて証明できるのだけど、面倒なので、割愛する。

●● 安全性 ●●

$\langle n, e\rangle$ は公開鍵で、$\langle n, d\rangle$ は秘密鍵だと書いた。ならば、公開されている e と n から d を計算できれば秘密鍵を入手したことになり、暗号は破れたことになる。その可能性を考慮してみる。

d は $\mathrm{mod}\phi(n)$ における e の逆数である。逆数を求める計算自体は Euclid の互除法を用いて簡単にできる[26]。重要なのは、$\phi(n) = n - (p+q) + 1$ が計算できるかどうかだ。これはどうやら困難である。なぜなら、p, q が公開されていない以上、自分でそれを計算しなければならないのだが、それは何百桁という数 n を素因数分解しなければ得られないからである。そして、素因数分解に対して有効なアルゴリズムは今のところ存在しない[27]。したがって、RSA 暗号は解読が困難であるとされている[28]。

アリスとボブは締切に遅れる

時刻は午前五時半。本当はもう一日かけてゆっくり書きたかったのだけど、締切はもう過ぎているので、そんな悠長なことを言っているわけにもいきませんでした。そういいつつ悠長にあとがきなど書く僕です。

「暗号」というとなんだか縁の無いものにきこえがちですが、実際は生活の中にすっかり溶け込んでいます。たとえば Web サイトのアドレスが `https:`から始まったりしていたらそれは暗号化通信が行われているサインです (たぶん)。そういえば、何かのカードの暗号が破られたとかそういう話もあったような。

現在でも暗号は日々進歩していますが、とりあえず正しく利用すれば解読されない段階にはあると思います。量子コンピュータの実用化もまだしばらくかかりそうだし。しかし、それは暗号を正しく利用した場合の話であって、そうでない場合にはすぐに破られてしまいます。第二次世界大戦のドイツ軍も、強力な暗号を開発したにもかかわらず運用上のミスから解読されたりしています。それに、人間というのはダメな生き物なので、目の前に大金を積まれたらすぐに何でも話してしまいます。ソーシャルエンジニアリングとかいいます[29]。

そういったことを防ぐには、何より暗号を利用する人間がそれを正しく理解することが必要だと思うのです。本稿がその助けになればいいな、などと書きますがそんな価値のあるものだとは思っていません。良くてきっかけ程度でしょう。それでも、より多くの人が関心を持つことは大切だと思うので、関心を持ってください (懇願)。

なんだかうまくまとまったような気がして上機嫌です。学校に行くまでにすこしは眠りたいのでここらで終わりにしたいと思います。ありがとうございました。

---

[26] でなければ、鍵の生成に時間がかかりすぎる。

[27] 量子コンピュータによる効率的なアルゴリズムは存在する。これは脅威となりうる。

[28] しかし、RSA 問題が素因数分解とそっくり等価であることは示されていない。つまり、素因数分解以外の方法で解読できる可能性が無いと決まったわけではないということである。

[29] そういえば暗号に対する攻撃についてほとんど触れていない。OB 特別寄稿の機会があればやろうかとも思う。