I240 暗号理論 2019

2019/10/11 講師 藤崎

1 用語と定義

- $\mathbb{N}$: 自然数の集合（この講義では 0 を含むとする）
- $\mathbb{Z}$: 整数の集合
- $\mathbb{Z}/n\mathbb{Z}$: n を正の整数として、0 以上 n 未満の整数の集合
- $(\mathbb{Z}/n\mathbb{Z})^\times$: $\mathbb{Z}/n\mathbb{Z}$ の元で、n と素なものの集合. すなわち、

$$(\mathbb{Z}/n\mathbb{Z})^\times := \{x \in \mathbb{Z}/n\mathbb{Z} \mid \gcd(x, n) = 1\}.$$

- 0^k: 0 の k ビット列. すなわち、$0^k := \overbrace{0 \dots 0}^{k}$.
- 1^k: 1 の k ビット列. すなわち、$1^k := \overbrace{1 \dots 1}^{k}$.
- $\{0,1\}^k$: k ビット長のビット列の集合.
- $\{0,1\}^*$: 有限長のビット列の集合. すなわち、$\{0,1\}^* := \bigcup_{k \in \mathbb{N}} \{0,1\}^k$. ただし、$\{0,1\}^0 = \{\varepsilon\}$ とし、ε は、長さが 0 のビット列を表す. $\varepsilon 1011 = 1\varepsilon 011 = 1011\varepsilon = 1011$.

■剰余類環 $\mathbb{Z}/n\mathbb{Z}$

- n は正の整数.$a, b \in \mathbb{Z}/n\mathbb{Z}$ に対して、

$$a + b := (a + b) \bmod n$$
$$a \cdot b := (a \cdot b) \bmod n$$

  と演算を定義すると、$\mathbb{Z}/n\mathbb{Z}$ 上で加法 $+$ と乗法 $\cdot$ が閉じる（右辺の加法 $+$ と乗法 $\cdot$ は $\mathbb{Z}$ 上の演算）
- a^{-1} を、$a \cdot a^{-1} = 1 \bmod n$ となる元で定義し、a の逆元と呼ぶ. $(\mathbb{Z}/n\mathbb{Z})^\times$ の元は全て逆元を持つ.
- n が素数の時、$a \neq 0$ に対して、a^{-1} は常に存在する。n が素数でない時、存在条件は $\gcd(a, n) = 1$（a と n の最大公約数が 1
- n が素数のとき、$\mathbb{Z}/n\mathbb{Z}$ は体（四則演算について閉じている）.

■排他的論理和 (Exclusive-OR, XOR) $\{0,1\}$ 上の二項演算である排他的論理和 $\oplus : \{0,1\} \times \{0,1\} \to \{0,1\}$ は、以下のように定義される。

$\oplus$	0	1
0	0	1
1	1	0

$a \oplus b = (a + b) \bmod 2$ $(a, b \in \{0,1\})$ が成立することに注意。

$\{0,1\}^n$ 上の排他的論理和 $\oplus : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ は、bit ごとの排他的論理和 (bit-wise XOR) で定義される。すなわち、$a = a_1 \dots a_n \in \{0,1\}^n$, $b = b_1 \dots b_n \in \{0,1\}^n$, $a \oplus b = c_1 \dots c_n \in \{0,1\}^n$ とすると、$c_i = a_i \oplus b_i \in \{0,1\}$. よって、(bit-wise) XOR を $\mathbb{Z}/2\mathbb{Z}$ 上の、$+$ 演算（の直積）と考えて良い。

2 拡張ユークリッドの互除法 (Extended Euclidian Algorithm)

- ユークリッドの互除法：$a, b \in \mathbb{Z}$ の最大公約数 (Greatest Common Divisor) d を求めるアルゴリズム
- 拡張ユークリッドの互除法：$a, b \in \mathbb{Z}$ に対して、

$$aX + bY = d$$

となる $X, Y \in \mathbb{Z}$ と最大公約数 d を同時に求めるアルゴリズム．$a \in \mathbb{Z}/n\mathbb{Z}$ の逆元 a^{-1} が存在する時、それを求めるアルゴリズムになっている．なぜならば、$aX + bn = 1$ なる X は、$aX \equiv 1 \pmod{n}$ であるから。

- ユークリッドの互除法を考える時、$(a, b) := a\mathbb{Z} + b\mathbb{Z} (= \{ax + by \,|\, x, y \in \mathbb{Z}\})$ と考えて変形を見て見ると理屈がよくわかる。

2.1 ユークリッドの互除法 (Euclidean Algorithm)

$(a, b) := a\mathbb{Z} + b\mathbb{Z} (= \{ax + by \,|\, x, y \in \mathbb{Z}\})$ と定義。すると $a - kb \geq 0$ となる $k \in \mathbb{Z}$ に対して、

$$(a, b) = (a - kb, b).$$

が常に成り立つ。小さな数で大きな数の余りとり変形し続ければ最終的に必ず $(d, 0) = (0, d) = (d) = d\mathbb{Z}$ の形になる。d が a, b の最大公約数 $\gcd(a, b)$ になることは自分で確かめてみる。

Algorithm 1 Euclid Algorithm

Require: $a \geq b \geq 0$.

Ensure: output the greatest common divisor of a and b.

```
Take a, b as input.
if b > 0 then
  (a, b) := (b, a mod b).
end if
if (a, b) = (d, 0) then
  return output d.
else
  return output ⊥.
end if
```

2.2 拡張ユークリッドの互除法 (Extended Euclidean Algorithm)

ユークリッドの互除法は、$a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$ となるような、d を探してくるアルゴリズムであった。集合としてイコールであるから、$aX + bY = d$ となる、$X, Y \in \mathbb{Z}$ が存在する。実際ユークリッドの互除法は、d と同時に X, Y も計算している。それを明示的に求めるアルゴリズムが拡張ユークリッドの互除法である。

このアルゴリズムは常に $aX + bY = \gcd(a, b)$ となる整数解 (X, Y) を出力して停止する。実際に計算してみるとわかりやすい。

Algorithm 2 Extended Euclidean Algorithm

Require: $a \geq b \geq 0$.

Ensure: output $d = \gcd(a, b)$ with $(X, Y) \in \mathbb{Z}^2$ such that $aX + bY = d$.

```
Take a, b as input.
(X_0, Y_0, Z_0) := (1, 0, a); (X_1, Y_1, Z_1) := (0, 1, b).
i := 0.
if Z_{i+1} > 0 then
  Compute quotient q := ⌊Z_i / Z_{i+1}⌋ and remainder r := Z_i mod Z_{i+1}.
  (X_{i+2}, Y_{i+2}, Z_{i+2}) := (X_i − qX_{i+1}, Y_i − qY_{i+1}, r).
  i++ (i.e., i := i + 1)
end if
if Z_{i+1} = 0 then
  return output (X_i, Y_i, Z_i).
else
  return output ⊥.
end if
```

i	q	Z_i	X_i	Y_i	$Z_i = 10X_i + 7Y_i$
0		10	1	0	$10 = 10 \cdot 1 + 7 \cdot 0$
1		7	0	1	$7 = 10 \cdot 0 + 7 \cdot 1$
2	1	3	1	-1	$3 = 10 \cdot 1 + 7 \cdot (-1)$
3	2	1	-2	3	$1 = 10 \cdot (-2) + 7 \cdot 3$
4	3	0	7	-10	$0 = 10 \cdot 7 + 7 \cdot (-10)$

図 1　拡張 Euclid による $10X + 7Y = \gcd(10, 7)$ の X, Y の求め方.

問題 1　解を求めよ.

- $6^{-1} \in \mathbb{Z}/31\mathbb{Z}$.
- $5^{-1} \in \mathbb{Z}/127\mathbb{Z}$.

3　One-Time Pad (OTP) と完全秘匿性 (Perfect Secrecy)

■共通鍵暗号 (Symmetric-Key Encryption) の定義　共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ とは、二つの集合と二つのアルゴリズムの組であり，次のように定義される：

- 鍵生成空間 $\mathcal{K}$: 秘密鍵 s の集合.
- 平文空間 $\mathcal{M}$: 平文 m の集合.
- 暗号化アルゴリズム $\mathbf{E}$: 秘密鍵 $s \in \mathcal{K}$, $m \in \mathcal{M}$ を入力としてとり，暗号文 c を出力するアルゴリズム この試行を $c \leftarrow \mathbf{E}_s(m)$ と書く. 出力は確率的 (probablistic) でもよい $c = \mathbf{E}_s(m; r)$（r は $\mathbf{E}$ に入力される乱数を表す）.
- 復号アルゴリズム $\mathbf{D}$: 秘密鍵 s と暗号文 c を入力としてとり, 平文 m を出力する確定的 (deterministic)

アルゴリズム。この試行を $m \leftarrow \mathbf{D}_s(c)$ と書く。

さらに、$\mathbf{D}$ は全ての鍵 $s \in \mathcal{K}$, 全ての平文 $m \in \mathcal{M}$ に対して、常に $\mathbf{D}_s(\mathbf{E}_s(m)) = m$ を満たす。この条件を共通鍵暗号の Correctness 条件という。すなわち、

$$\text{Correctness:} \quad \forall m \in \mathcal{M} \quad \Pr_{s \leftarrow \mathcal{K}}[\mathbf{D}_s(\mathbf{E}_s(m)) = m] = 1.$$

復号アルゴリズムを確率的なアルゴリズムと定義しても良いが、Correctness を満たすのであれば、復号アルゴリズムは必ず確定的になる。

■One-Time Pad (Vernam Cipher)　共通鍵暗号の定義に基づいて記述すると次のようになる。

- 鍵空間 $\mathcal{K} = \{0,1\}^\kappa$.
- 暗号化アルゴリズム $\mathbf{E}$. 秘密鍵 s と平文 $m \in \mathcal{M} := \{0,1\}^\kappa$ を受け取り、暗号文 $c = \mathbf{E}_s(m) := m \oplus s$ を出力する。
- 復号アルゴリズム $\mathbf{D}$. 秘密鍵 s と暗号文 c を受け取り、平文 $m = \mathbf{D}_s(c) := c \oplus s$ を出力する.

Correctness を満たすことを確認せよ。

■完全秘匿性 (Perfect Secrecy)　共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性を満たすとは、平文に対する情報量が、暗号文 c を見る前と後で全く変わらないときを言う。

定義 2（Shannon）　共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性 (perfect secure) を満たすとは、$\mathcal{M}$ 上定義されるの任意の確率分布 X, 任意の平文 $m \in \mathcal{M}$, 任意の暗号文 c に対して、

$$\Pr_{X,K}[X = m \wedge \mathbf{E}_K(X) = c] = \Pr_X[X = m] \cdot \Pr_{X,K}[\mathbf{E}_K(X) = c].$$

が成り立つときを言う。ここで、K は $\mathcal{K}$ から s を一様ランダムに選ぶ分布に従う確率変数とする。

すなわち、$\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性とは、任意の確率変数 X と $\mathbf{E}_K(X)$ が独立であること。

SKE : perfectly secret $\overset{\text{def}}{\iff}$

$$\forall X\, \forall m\, \forall c : \Pr_{X,K}[X = m \wedge \mathbf{E}_K(X) = c] = \Pr_X[X = m] \cdot \Pr_{X,K}[\mathbf{E}_K(X) = c].$$

命題 3　以下は全て等しい。

- 共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性を満たす。
- $\mathcal{M}$ 上定義されるの任意の確率分布 X, 任意の平文 $m \in \mathcal{M}$, 任意の暗号文 $c \in \mathsf{supp}(\mathbf{E}_K(X)) := \{c \mid \Pr_{X,K}[\mathbf{E}_K(X) = c] > 0\}$ に対して、

  $$\Pr_{X,K}[X = m \mid \mathbf{E}_K(X) = c] = \Pr_X[X = m].$$

- $\mathcal{M}$ 上定義されるの任意の確率分布 X に対して、

  $$H(X) = H(X \mid \mathbf{E}_K(X))$$

  が成り立つ。ここで $H(X)$ は確率変数 X の Shannon entropy を表す（後述)。
- $\mathcal{M}$ 上定義されるの任意の確率分布 X に対して、

  $$H(X,Y) = H(X) + H(\mathbf{E}_K(X)).$$

定理 4　共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性を満たすならば、$|\mathcal{K}| \geq |\mathcal{M}|$ が成り立つ。

$$\mathsf{SKE}: \text{ perfectly secret } \Longrightarrow |\mathcal{K}| \geq |\mathcal{M}|$$

暗号化アルゴリズム $\mathbf{E}$ が確定的 (deterministic) でない、つまり確率的 (probablistic) でも成り立つ。

(証明) SKE が、$|\mathcal{K}| < |\mathcal{M}|$ にも関わらず、完全秘匿性を満たしたとすると矛盾であることを証明する。

$m_0 \in \mathcal{M}$, $s_0 \in \mathcal{K}$ を選び、$c \leftarrow \mathbf{E}_{s_0}(m_0)$ を計算する。$S(c) := \{\mathbf{D}_s(c) \,|\, s \in \mathcal{K}\}$ と定義する。すると、復号アルゴリズムの確定性から $S(c) \leq |\mathcal{K}|$. 一方、仮定より $|\mathcal{K}| < |\mathcal{M}|$ であるので、

$$S(c) < |\mathcal{M}|$$

が成り立つ。よって、$\mathcal{M} \backslash S(c) \neq \emptyset$ であるから、$m \in \mathcal{M} \backslash S(c)$ が選べる。その選び方より、全ての $s \in \mathcal{K}$ に対して、$\mathbf{E}_s(m) \neq c$. よって、($\Pr[X = m_0] > 0$ なる) 全ての確率変数 X に対して、

$$\Pr_{X,K}[X = m \,|\, E_K(X) = c] = 0. \tag{1}$$

いま、X を $\mathcal{M}$ 上の一様分布に従う確率変数とする。すると、

$$\Pr_{X}[X = m] = \frac{1}{|\mathcal{M}|}. \tag{2}$$

ここで、$c \in \mathsf{supp}(E_K(X))$ であり、SKE は完全秘匿性を満たすので、式 (1) と式 (2) の値は一致しなければならないが、一致しないので矛盾。(よって、対偶により、) SKE が完全秘匿性を満たすなら、$|\mathcal{K}| \geq |\mathcal{M}|$ となる。 ■

■Shannon Entropy　確率変数 X に対して $X = x$ なる事象の起こりにくさの尺度を示す値として情報量 (information content) がある。確率変数 X における $X = x$ を表す情報量は $\mathsf{Info}(x) = -\log_2(p(x))$ で定義される ($p(x) := \Pr[X = x]$). $p(x, y) := \Pr[X = x \wedge Y = y]$ と $p(x|y) := \Pr[X = x \,|\, Y = y]$ とし、$\mathsf{Info}(x, y) := -\log_2(p(x, y))$, $\mathsf{Info}(x|y) := -\log_2(p(x|y))$ と定義する。その時、次が成立することは容易にわかる。$\mathsf{Info}(x, y) = \mathsf{Info}(x) + \mathsf{Info}(y|x) = \mathsf{Info}(y) + \mathsf{Info}(x|Y)$. X, Y が独立ならば、$\mathsf{Info}(x, y) = \mathsf{Info}(x) + \mathsf{Info}(y)$.

確率変数 X に対する Shannon Entropy (または平均情報量) を、

$$H(X) := \sum_{x \in \mathcal{X}} p(x) \cdot \mathsf{Info}(x)$$

で定義する。

確率変数 (X, Y) の (X と Y の結合) Shannon Entropy は、

$$H(X, Y) := \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \cdot \mathsf{Info}(x, y)$$

で定義される。一方、条件付き Shannon Entropy は

$$H(X|Y) := \sum_{y \in \mathcal{Y}} H(X|Y = y)$$

で定義される。$H(X|Y = y) = \sum_{x \in \mathcal{X}} p(x|y) \cdot \mathsf{Info}(x|y)$ より、

$$\begin{aligned} H(X|Y) &= \sum_{y \in \mathcal{Y}} p(y) \sum_{x \in \mathcal{X}} p(x|y) \cdot \mathsf{Info}(x|y) \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \cdot \mathsf{Info}(x|y). \end{aligned}$$

命題 5（Chain Rule）

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y).$$

(証明) 式変形から明らか。

$$\begin{aligned}
H(X,Y) &:= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x,y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot (\mathsf{Info}(x) + \mathsf{Info}(x|y)) \qquad *** \ \mathsf{Info}(x,y) = \mathsf{Info}(x) + \mathsf{Info}(x|y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x) + \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x|y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x) + H(X|Y) \qquad *** \ H(X|Y) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x|y) \\
&= \sum_{x\in\mathcal{X}} p(x)\cdot \mathsf{Info}(x) + H(X|Y) \qquad *** \ p(x) = \sum_{y\in\mathcal{Y}} p(x,y) \\
&= H(X) + H(X|Y).
\end{aligned}$$

命題 6 X,Y が独立な確率変数であれば、

$$H(X,Y) = H(X) + H(Y).$$

(証明) X,Y が独立なら $\mathsf{Info}(x,y) = \mathsf{Info}(x) + \mathsf{Info}(y)$. さらに、$p(x) = \sum_{y\in\mathcal{Y}} p(x,y)$ に注意すれば式変形から明らか。

$$\begin{aligned}
H(X,Y) &:= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x,y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot (\mathsf{Info}(x) + \mathsf{Info}(y)) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x) + \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(y) \\
&= H(X) + H(Y)
\end{aligned}$$

命題 7 以下は全て等しい。

- X と Y が独立な確率変数
- 全ての $x\in\mathcal{X}$, $y\in\mathcal{Y}$ に対して、

$$\Pr[X = x \wedge Y = y] = \Pr[X = x]\cdot\Pr[Y = y].$$

- 全ての $x\in\mathcal{X}$, $y\in\mathcal{Y}$ に対して、

$$\Pr[X = x] = \Pr[X = x|Y = y].$$

- X,Y の結合エントロピーが各エントロピーの和

$$H(X,Y) = H(X) + H(Y).$$

- X の Shannon entropy が、Y が起きた後にも減少しない

$$H(X) = H(X|Y).$$