情報理論 第 5 回
– 情報源符号化と情報源符号化定理 –
(教科書: 6.4 節, 8 章, 9 章)

野崎 隆之

0. 復習と導入

復習

- (離散) 定常無記憶情報源 X で情報源アルファベットの出現頻度が与えられる

$$X : \begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$$

- 瞬時符号 $=$ 語頭符号
- クラフトの不等式 (瞬時符号の符号長の満たすべき条件)

導入

$a_i = x$	p_i	$f(a_i)$	ℓ_i
a_1 =a	0.5		
a_2 =b	0.25		
a_3 =c	0.15		
a_4 =d	0.1		

- どのように符号語を割り当てれば平均符号語長の短い語頭符号が構成できるか?
- 平均符号語長の短さの限界は?

今日の目的と流れ

概要

- 平均符号語長の短さの限界
  $\Rightarrow$ 情報源符号化定理
- 平均符号語長の短い語頭符号の構成 (良い情報源符号の構成法)
  $\Rightarrow$ ハフマン符号

今日の目標

- 平均符号語長の計算ができる
- 情報源符号化定理の意味を説明できる
- ハフマン符号の構成ができる

今日の流れ

1. 平均符号語長
2. 情報源符号化定理
3. ハフマン符号

1. 平均符号語長 (1: 定義)

(定義) 平均符号語長

- $\{a_1, a_2, \ldots, a_k\}$: 情報源アルファベット
- p_i : シンボル a_i の生起確率 (定常無記憶情報源 X から $p_X(a_i) = p_i$)
- ℓ_i : シンボル a_i の符号語 $f(a_i)$ の長さ

符号 $f(x)$ の平均符号語長 $\bar{L}$ は　　短いほうがいい

$$\bar{L} := \sum_{i=1}^{k} p_i \ell_i$$

意味：符号長の期待値 $\Rightarrow$ 符号化の効率の指標

1. 平均符号語長 (2: 例題)

(例) 平均符号語長

$a_i = x$	p_i	$f_1(a_i)$	l_i
$a_1 =$ a	0.5	0	1
$a_2 =$ b	0.25	10	2
$a_3 =$ c	0.15	110	3
$a_4 =$ d	0.1	111	3

$$\bar{L} = \sum_{i=1}^{k} p_i \ell_i$$
$$= 0.5 \times 1 + 0.25 \times 2 + 0.15 \times 3 + 0.1 \times 3$$
$$= 1.75$$

(例題 1)

$a_i = x$	p_i	$f_2(a_i)$	l_i
$a_1 =$ a	0.5	00	2
$a_2 =$ b	0.25	01	2
$a_3 =$ c	0.15	10	2
$a_4 =$ d	0.1	11	2

平均符号語長 $\bar{L}$ **は**? 2

(例題 2)

$a_i = x$	p_i	$f_3(a_i)$	l_i
$a_1 =$ a	0.5	111	
$a_2 =$ b	0.25	110	
$a_3 =$ c	0.15	10	
$a_4 =$ d	0.1	0	

平均符号語長 $\bar{L}$ **は**?

1. 平均符号語長 (3: 考察)

前のページの例から以下がわかる

- **(例と例題 1,2 より)**
  生起確率が同じでも，符号によって平均符号語長が異なる
- **(例と例題 1 より)**
  固定長符号よりも可変長符号の方が平均符号語長が短い場合がある
- **(例と例題 2 より)**
  短い符号語を生起確率が高いシンボルに割り当てると平均符号語長が短くなる

1. **平均符号語長はどこまで短く出来るか?**
2. **どうやれば平均符号語長の短い符号を構成できるか?**

2. 情報源符号化定理 (1: 情報源のエントロピー)

復習

- (離散) 定常無記憶情報源 X で情報源アルファベットの出現頻度が与えられる

$$X : \begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ p_1 & p_2 & \cdots & p_k \end{pmatrix}$$

- 確率分布が $p_1, p_2, \ldots, p_k$ で与えられる確率変数 X のエントロピーは

$$H(X) = -\sum_{i=1}^{k} p_i \log p_i$$

定常無記憶情報源のエントロピー

定常無記憶情報源 X のエントロピー $H(X)$ は

$$H(X) = -\textstyle\sum_{i=1}^{k} p_i \log p_i$$

2. 情報源符号化定理 (1: 概要)

これはかなり大事

情報源符号化定理の意味 (重要!!)

定常無記憶情報源 X に対して，

逆定理と呼ばれている

1. 平均符号語長が $H(X)$ より小さい語頭符号は作れない (逆定理)
2. 平均符号語長 $\bar{L}$ が次の不等式を満足する語頭符号が存在 (順定理 1)

$$H(X) \leq \bar{L} < H(X) + 1$$

3. 次を満足する語頭符号が存在する (順定理 2)
   - 情報源の出力系列を n シンボル毎に区切って符号化
   - $\bar{L}_n$: 1 シンボルあたりの平均符号語長

$$H(X) \leq \bar{L}_n < H(X) + \frac{1}{n}$$

$\Rightarrow$ n を十分を大きくとれば，平均符号語長がエントロピーに漸近

2-1. 情報源符号化逆定理 (1: 逆定理)

(定理) 情報源符号化定理 (逆定理)

離散定常無記憶情報源 X に対して，
任意の語頭符号の平均符号語長 $\bar{L}$ は次式を満たす

$$\bar{L} \geq H(X)$$

(等号成立は任意の i に対して $p_i = 2^{-\ell_i}$)

意味：
平均符号語長の下界がエントロピー $H(X)$ で与えられる．
$\Longleftrightarrow$ 平均符号語長をエントロピーより小さくすることはできない．

2-1. 情報源符号化逆定理 (2: 逆定理の証明の道具)

(定理) クラフトの不等式 (1)

2 元語頭符号が k 個の符号語をもち，それらの符号長が $\ell_1, \ell_2, \ldots, \ell_k$ であれば，次の不等式を満足する

$$\sum_{i=1}^{k} 2^{-\ell_i} \leq 1.$$

(補題) シャノンの補助定理 (第 2 回資料より)

確率分布 $(p_1, p_2, \ldots, p_k)$ と $\sum_{i=1}^{k} q_k \leq 1$ を満たす $(q_1, q_2, \ldots, q_k)$ に対して，次の不等式が成り立つ

$$-\sum_{i=1}^{k} p_i \log p_i \leq -\sum_{i=1}^{k} p_i \log q_i \tag{1}$$

等号成立条件は全ての i に対して $p_i = q_i$.

2-1. 情報源符号化定理 (3: 逆定理の証明)

情報源符号化逆定理の証明：

語頭符号なので，クラフトの不等式を満足する

$$\sum_{i=1}^{k} 2^{-\ell_i} \leq 1. \qquad (2)$$

$q_i := 2^{-\ell_i}$ **と定義すると，**

$$\sum_{i=1}^{k} q_i \leq 1 \qquad (\text{式 (2) より}), \qquad (3)$$

$$\begin{aligned} L &= \sum_{i=1}^{k} p_i \ell_i && (\text{平均符号語長の定義}) && (4) \\ &= -\sum_{i=1}^{k} p_i \log q_i && (\ell_i = -\log q_i \iff q_i := 2^{-\ell_i}) && (5) \\ &\geq -\sum_{i=1}^{k} p_i \log p_i && (\text{シャノンの補助定理}) && (6) \\ &= H(X) && (\text{エントロピーの定義}) && (7) \end{aligned}$$

2-2. 情報源符号化定理 1 (1: 順定理 1)

(定理) 情報源符号化定理 (順定理 1)

離散定常無記憶情報源 X に対して，
平均符号語長 $\bar{L}$ が以下の条件を満たす語頭符号が存在する

$$H(X) \leq \bar{L} < H(X) + 1$$

意味：

- **エントロピーに接近する語頭符号が存在**
- **情報源の出力に偏りが大きい (エントロピーが小さい) 場合，平均符号語長は短く出来る．**

注意：

- **左側の不等号は逆定理より明らか (右側だけ証明すれば良い)**
- **証明の手順を追うとエントロピーに接近する語頭符号が構成可能 (シャノン・ファノ符号と呼ばれる)**

2-2. 情報源符号化定理 (2: 順定理 1 の証明の道具)

(定理) クラフトの不等式 (2)

$$\sum_{i=1}^{k} 2^{-\ell_i} \leq 1.$$

を満たす符号長の組 $\ell_1, \ell_2, \ldots, \ell_k$ が与えられた時に，これらの符号長を有する 2 元語頭符号を構成できる．

天井関数 (ceiling function)

実数 x に対して，x 以上で最小の整数を $\lceil x \rceil$ と表し，天井関数と呼ぶ (直感的には“切り上げ” のこと)

$\lceil 2.5 \rceil = 3,\ \lceil 4 \rceil = 4,\ \lceil -2.5 \rceil = -2$

$$x \leq \lceil x \rceil < x + 1$$

2-2. 情報源符号化定理 1 (3: 順定理 1 の証明)

順定理 1 の証明：
$\ell_i := \lceil -\log p_i \rceil$ **とする．**

$$-\log p_i \leq \ell_i < -\log p_i + 1 \tag{8}$$

$$\begin{aligned} \textstyle\sum_{i=1}^k 2^{-\ell_i} &\leq \textstyle\sum_{i=1}^k 2^{\log p_i} && \textbf{(式 (8) の左側)} \\ &= \textstyle\sum_{i=1}^k p_i = 1 \end{aligned}$$

クラフトの不等式が成立するので，符号長が $(\ell_1, \ell_2, \ldots, \ell_k)$ **となる語頭符号が存在する．**
この語頭符号の平均符号語長は

$$\begin{aligned} \bar{L} &= \textstyle\sum_{i=1}^k p_i \ell_i \\ &< \textstyle\sum_{i=1}^k p_i (-\log p_i + 1) && \textbf{(式 (8) の右側)} \\ &= H(X) + 1 \end{aligned}$$

2-2. 情報源符号化定理 (4: 順定理 1 の確認)

(例)

$$X : \begin{pmatrix} a_1 & a_2 & a_3 \\ 0.5 & 0.3 & 0.2 \end{pmatrix}$$

i	1	2	3
p_i	0.5	0.3	0.2
$-\log p_i$	1	1.736	2.322
$\ell_i = \lceil -\log p_i \rceil$	1	2	3
$f(a_i)$	0	10	110

クラフトの不等式 (2) の証明での符号を構成法を利用

- **平均符号語長** $\bar{L} = 1.7$
- **エントロピー** $H(X) = 1.4852$

シャノン・ファノ (Shannon-Fano) 符号 **テストで出てた**

この手順で得られる語頭符号をシャノン・ファノ符号と呼ぶ

2-3. 情報源符号化定理 2 (1: 拡大情報源)

(定義) n 次拡大情報源

定常無記憶情報源 X に対して，

$$P_{X^n}(x_1x_2\cdots x_n)=\prod_{i=1}^{n}P_X(x_i) \tag{9}$$

で与えられる情報源 X^n を n 次拡大情報源と呼ぶ

意味： 情報源 X の出力系列を n シンボル毎に区切ったものの発生確率

(例)

$$X:\begin{pmatrix}0 & 1\\ 0.6 & 0.4\end{pmatrix}$$

に対する 2 次拡大情報源 X^2 は

$$X^2:\begin{pmatrix}00 & 01 & 10 & 11\\ 0.36 & 0.24 & 0.24 & 0.16\end{pmatrix}$$

2-3. 情報源符号化定理 2 (2: 拡大情報源の性質)

拡大情報源のエントロピー

定常無記憶情報源 X に対する n 次拡大情報源 X^n のエントロピーは

$$H(X^n) = nH(X) \tag{10}$$

を満たす.

2-3. 情報源符号化定理 2 (3: 拡大情報源の性質の証明)

証明：数学的帰納法を用いる．

1. $n = 1$ **のとき，明らか．**
2. $n = t$ **のときに，式 (10) が成り立つと仮定する．**
   $\underline{\boldsymbol{x}}^{(t)} := x_1 x_2 \cdots x_t, \quad \underline{\boldsymbol{x}}^{(t+1)} := x_1 x_2 \cdots x_t x_{t+1}$
   $P_{X^{t+1}}(\underline{\boldsymbol{x}}^{(t+1)}) = P_{X^t}(\underline{\boldsymbol{x}}^{(t)}) P_X(x_{t+1})$ **に注意．**

$$
\begin{aligned}
&H(X^{t+1}) \\
&\quad = -\textstyle\sum_{\underline{\boldsymbol{x}}^{t+1}} P_{X^{t+1}}(\underline{\boldsymbol{x}}^{t+1}) \log P_{X^{t+1}}(\underline{\boldsymbol{x}}^{t+1}) \\
&\quad = -\textstyle\sum_{\underline{\boldsymbol{x}}^t} \sum_{x_{t+1}} P_{X^t}(\underline{\boldsymbol{x}}^t) P_X(x_{t+1}) [\log P_{X^t}(\underline{\boldsymbol{x}}^t) + \log P_X(x_{t+1})] \\
&\quad = -\textstyle\sum_{\underline{\boldsymbol{x}}^t} \sum_{x_{t+1}} P_{X^t}(\underline{\boldsymbol{x}}^t) P_X(x_{t+1}) \log P_{X^t}(\underline{\boldsymbol{x}}^t) \\
&\qquad - \textstyle\sum_{\underline{\boldsymbol{x}}^t} \sum_{x_{t+1}} P_{X^t}(\underline{\boldsymbol{x}}^t) P_X(x_{t+1}) \log P_X(x_{t+1}) \\
&\quad = -[\textstyle\sum_{x_{t+1}} P_X(x_{t+1})] \sum_{\underline{\boldsymbol{x}}^t} P_{X^t}(\underline{\boldsymbol{x}}^t) \log P_{X^t}(\underline{\boldsymbol{x}}^t) \\
&\qquad - [\textstyle\sum_{\underline{\boldsymbol{x}}^t} P_{X^t}(\underline{\boldsymbol{x}}^t)] \sum_{x_{t+1}} P_X(x_{t+1}) \log P_X(x_{t+1}) \\
&\quad = 1 \cdot H(X^t) + 1 \cdot H(X) = tH(X) + H(X) = (t+1)H(X)
\end{aligned}
$$

2-3. 情報源符号化定理 2 (4: 順定理 2)

(定理) 情報源符号化定理 2

- X^n : **定常無記憶情報源 X の n 次拡大情報源**
- $\bar{L}_n$: **1 シンボルあたりの平均符号語長**

$\bar{L}_n$ **が次の不等式を満足する X^n に対する語頭符号が存在**

$$H(X) \leq \bar{L}_n < H(X) + \frac{1}{n}$$

証明：

情報源符号化定理 1 より，X^n に対して平均符号語長 $\bar{L}$ の語頭符号が存在

$$H(X^n) \leq \bar{L} < H(X^n) + 1. \tag{11}$$

- $\bar{L} = n\bar{L}_n$ 　($\bar{L}$ **が n シンボル毎の平均符号語長だから**)
- $H(X^n) = nH(X)$

$$nH(X) \leq n\bar{L}_n < nH(X) + 1 \tag{12}$$

3. ハフマン符号 (1:概要)

ハフマン符号の性質

- 平均符号語長が最小となる語頭符号を構成できる
  ＝ コンパクト符号

注意： シャノン・ファノ符号はコンパクト符号とは限らない

(定義) コンパクト符号

与えられた無記憶通信路に対して，平均符号語長が最小となる語頭符号

ここから先の流れ

(3-1) ハフマン符号の構成 (重要!!) ←こっちだけでいい

(3-2) ハフマン符号がコンパクト符号であることの証明 (時間があれば…)
丁寧な (わかりやすい) 証明法は下記参照

- 植松 "イラストで学ぶ 情報理論の考え方" pp.122–128
- 山本ら "情報理論 基礎と広がり" pp.90–93

3-1. ハフマン符号の構成 (1:構成例)

手続きをきちんと書くとわかりづらい
$\Rightarrow$ **まず例で構成の仕方を覚える**

$$X : \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ 0.35 & 0.2 & 0.2 & 0.15 & 0.1 \end{pmatrix}$$

写真で撮った
・上から確率の高いやつ
・下のやつを足し合わせたら
順序を入れ替える
・足し合わせた点の上を0,下を1として、
富豪の木を作る
・みぎから読み上げる

3-1. ハフマン符号の構成 (2:構成法)

1. **情報源アルファベットを生起確率が高いものから順に上から並べる．**
2. **生起確率の最も低い** 2 **つのシンボルを合併させる．**
   - **新たな接点を追加して** 2 **つのシンボルと枝で結ぶ**
   - **新たな接点の生起確率を** 2 **つのシンボルの生起確率にする**
3. **生起確率の高さの順に上から並び替える**
4. 1 **つの木になるまで** Step 2,3 **を繰り返す**
5. **枝に** $\{0, 1\}$ **でラベル付けをする**

3-1. ハフマン符号の構成 (3: 例題)

(例題)

1. **以下の定常無記憶通信路に対して，ハフマン符号を構成せよ**

$$X : \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ 0.35 & 0.30 & 0.10 & 0.10 & 0.10 & 0.05 \end{pmatrix}$$

2. **得られたハフマン符号の平均符号語長を求めよ**

平均符号語長は符号の長さと確率をかけたやつを足す

3-2. 証明 (1: 概要と定義)

(定義) コンパクト符号 (再掲)

与えられた無記憶通信路に対して，平均符号語長が最小となる語頭符号

(定理)

ハフマン符号はコンパクト符号である

証明の概略

1. コンパクト符号の性質
2. 縮約情報源 S_i の定義
3. S_i に対するハフマン符号がコンパクト符号ならば S_{i+1} に対するハフマン符号もコンパクト符号 (数学的帰納法)

3-2. 証明 (2: コンパクト符号の性質 1)

性質 1

コンパクト符号のとき，$p_i \geq p_j$ ならば $\ell_i \leq \ell_j$.
すなわち，$p_1 \geq p_2 \geq \cdots \geq p_k$ ならば $\ell_1 \leq \ell_2 \leq \cdots \leq p_k$

証明の概略： 背理法を用いる．
($\ell_i > \ell_j$ を仮定するとコンパクト性に矛盾)

性質 2

生起確率の最も低い 2 つのシンボルの符号長は等しい．
すなわち，$p_1 \geq p_2 \geq \cdots \geq p_{k-1} \geq p_k$ のとき $\ell_{k-1} = \ell_k$

証明の概略：背理法を用いる．
($\ell_{k-1} < \ell_k$ を仮定すると，符号語 $f(a_k)$ の最後の $\ell_k - \ell_{k-1}$ シンボルを削ることで，より短い語頭符号が構成可能)

3-2. 証明 (3: コンパクト符号の性質 2)

性質 3

コンパクト符号の木では，内部接点に必ず 2 本の枝が存在する．

証明の概略： **背理法を用いる．**
(1 **本の枝しか持たない内部接点があれば，さらに短い語頭符号ができる**)

性質 4

$p_1 \geq p_2 \geq \cdots \geq p_k$ **ならば** $f(a_{k-1})$ **と** $f(a_k)$ **は最後の** 1 **シンボルのみが異なる．**

証明の概略： **性質** 1,2,3 **を用いる．**
符号長が最も長くなる符号語の数が 2 **の時と** 3 **以上の時で場合分け**

3-2. 証明 (4: コンパクト符号の性質の確認)

(例)

$$X : \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \\ 0.5 & 0.3 & 0.15 & 0.05 \end{pmatrix}$$

に対する次の符号 f はコンパクト符号の性質を満たす

x	a_1	a_2	a_3	a_4
$f(x)$	0	10	110	111

3-2. 証明 (5: 縮約情報源 1)

ハフマン符号の第 2 ステップ

2. **生起確率の最も低い 2 つのシンボルを合併させる．**
   (2 つのシンボルを合わせて新しい 1 つのシンボルを作る)
   - **新たな接点を追加して 2 つのシンボルと枝で結ぶ**
   - **新たな接点の生起確率を 2 つのシンボルの生起確率にする**

ハフマン符号は 1 回の操作毎に 1 つずつシンボルが減少
$\Rightarrow$ 1 回の操作毎にシンボル数の減った新しい情報源ができる

縮約情報源

ハフマン符号の構成によってシンボル数が m 個になった情報源を
縮約情報源と呼び，S_m と表す．

3-2. 証明 (6: 縮約情報源 2)

(例)

$$S_4 : \begin{pmatrix} a_1^{(4)} & a_2^{(4)} & a_3^{(4)} & a_4^{(4)} \\ 0.5 & 0.3 & 0.15 & 0.05 \end{pmatrix}$$

$$S_3 : \begin{pmatrix} a_1^{(3)} & a_2^{(3)} & a_3^{(3)} \\ 0.5 & 0.3 & 0.2 \end{pmatrix}, \quad S_2 : \begin{pmatrix} a_1^{(2)} & a_2^{(2)} \\ 0.5 & 0.5 \end{pmatrix} \quad S_1 : \begin{pmatrix} a_1^{(1)} \\ 1 \end{pmatrix}$$

符号の木は板書

3-2. 証明 (7: コンパクト性の証明 1)

(定理) (再掲)

ハフマン符号はコンパクト符号である．

証明の概要；縮約情報源に対する数学的帰納法

(補題 1)

$$S_2 : \begin{pmatrix} a_1^{(2)} & a_2^{(2)} \\ p_1^{(2)} & p_2^{(2)} \end{pmatrix}$$

に対するハフマン符号はコンパクト符号である．
すなわち，$f(a_1^{(2)}) = 0, f(a_2^{(2)}) = 1$ はコンパクト符号である．

証明： 符号長は少なくとも 1 以上になるので，平均符号語長は 1 以上．
ハフマン符号によって与えられる平均符号語長は 1 なので，最小．

(補題 2)

縮約情報源 S_i に対するハフマン符号 f_i がコンパクト符号ならば
縮約情報源 S_{i+1} に対するハフマン符号 f_{i+1} もコンパクト符号

3-2. 証明 (8: 補題 2 の証明 1)

縮約情報源の表現法

$$S_{i+1} : \begin{pmatrix} a_1^{(i+1)} & a_2^{(i+1)} & \cdots & a_i^{(i+1)} & a_{i+1}^{(i+1)} \\ p_1^{(i+1)} & p_2^{(i+1)} & \cdots & p_i^{(i+1)} & p_{i+1}^{(i+1)} \end{pmatrix},$$

$$\tilde{S}_i : \begin{pmatrix} a_1^{(i+1)} & a_2^{(i+1)} & \cdots & A \\ p_1^{(i+1)} & p_2^{(i+1)} & \cdots & p_i^{(i+1)} + p_{i+1}^{(i+1)} \end{pmatrix}, \quad \text{(合併)}$$

$$S_i : \begin{pmatrix} a_1^{(i)} & a_2^{(i)} & \cdots & a_i^{(i)} \\ p_1^{(i)} & p_2^{(i)} & \cdots & p_i^{(i)} \end{pmatrix} \quad \text{(確率の順に並び替え)}$$

$\tilde{S}_i$ と S_i は本質的に同じ (並び替えをしているだけ)

3-2. 証明 (9: 補題 2 の証明 2)

平均符号語長の性質

f_{i+1}	シンボル	$a_1^{(i+1)}$	$a_2^{(i+1)}$	$\cdots$	A	
	符号長	$\ell_1^{(i+1)}$	$\ell_2^{(i+1)}$	$\cdots$	$\ell_i^{(i)}$	
	確率	$p_1^{(i+1)}$	$p_2^{(i+1)}$	$\cdots$	$p_i^{(i+1)} + p_{i+1}^{(i+1)}$	
f_i	シンボル	$a_1^{(i+1)}$	$a_2^{(i+1)}$	$\cdots$	$a_i^{(i+1)}$	$a_{i+1}^{(i+1)}$
	符号長	$\ell_1^{(i+1)}$	$\ell_2^{(i+1)}$	$\cdots$	$\ell_i^{(i)} + 1$	$\ell_i^{(i)} + 1$
	確率	$p_1^{(i+1)}$	$p_2^{(i+1)}$	$\cdots$	$p_i^{(i+1)}$	$p_{i+1}^{(i+1)}$

- $\bar{L}_{i+1}$: S_{i+1} **に対するハフマン符号** f_{i+1} **の平均符号語長**
- $\bar{L}_i$: $\tilde{S}_i$ **に対するハフマン符号** f_i **の平均符号語長**

$$\bar{L}_{i+1} = \ell_1^{(i+1)} p_1^{(i+1)} + \ell_2^{(i+1)} p_2^{(i+1)} + \cdots + (\ell_i^{(i)} + 1)(p_i^{(i+1)} + p_{i+1}^{(i+1)})$$

$$\bar{L}_i = \ell_1^{(i+1)} p_1^{(i+1)} + \ell_2^{(i+1)} p_2^{(i+1)} + \cdots + \ell_i^{(i)} (p_i^{(i+1)} + p_{i+1}^{(i+1)})$$

$$\bar{L}_{i+1} - \bar{L}_i = p_i^{(i+1)} + p_{i+1}^{(i+1)} \tag{13}$$

3-2. 証明 (10: 補題 2 の証明 3)

S_{i+1} に対するハフマン符号がコンパクト符号でないとする (背理法).
すなわち，平均符号語長 $\bar{L}'_{i+1}$ となるコンパクト符号 f'_{i+1} が存在して，

$$\bar{L}'_{i+1} < \bar{L}_{i+1} \tag{14}$$

コンパクト符号の性質 4 より，
$f'_{i+1}(a_i^{(i+1)})$ と $f'_{i+1}(a_{i+1}^{(i+1)})$ は最後のシンボルのみが異なる.
S_i の新しい符号 f'_i を次の表のとおりに構成

f'_{i+1}	シンボル	$a_1^{(i+1)}$	$a_2^{(i+1)}$	$\cdots$	A	
	符号語	$\boldsymbol{y}_1^{(i+1)}$	$\boldsymbol{y}_2^{(i+1)}$	$\cdots$	$\boldsymbol{y}_i^{(i)}$	
	符号長	$\ell_1^{(i+1)}$	$\ell_2^{(i+1)}$	$\cdots$	$\ell_i^{(i)}$	

f'_i	シンボル	$a_1^{(i+1)}$	$a_2^{(i+1)}$	$\cdots$	$a_i^{(i+1)}$	$a_{i+1}^{(i+1)}$
	符号語	$\boldsymbol{y}_1^{(i+1)}$	$\boldsymbol{y}_2^{(i+1)}$	$\cdots$	$\boldsymbol{y}_i^{(i)}0$	$\boldsymbol{y}_i^{(i)}1$
	符号長	$\ell_1^{(i+1)}$	$\ell_2^{(i+1)}$	$\cdots$	$\ell_i^{(i)}+1$	$\ell_i^{(i)}+1$

$$\bar{L}'_{i+1} - \bar{L}'_i = p_{i+1}^{(i+1)} + p_{i+1}^{(i)} \tag{15}$$

3-2. 証明 (10: 補題 2 の証明 4)

平均符号語長についての式をまとめると

$$\bar{L}_{i+1} - \bar{L}_i = p_i^{(i+1)} + p_{i+1}^{(i+1)} \tag{16}$$

$$\bar{L}'_{i+1} < \bar{L}_{i+1} \tag{17}$$

$$\bar{L}'_{i+1} - \bar{L}'_i = p_{i+1}^{(i+1)} + p_{i+1}^{(i)} \tag{18}$$

以上から，

$$\bar{L}'_i = \bar{L}'_{i+1} - p_{i+1}^{(i+1)} - p_{i+1}^{(i)} < \bar{L}_{i+1} - p_{i+1}^{(i+1)} - p_{i+1}^{(i)} = \bar{L}_i \tag{19}$$

したがって，f_i よりも f'_i の方が平均符号語長が短い
f_i がコンパクト符号であることに反する．

今回のまとめ

情報源符号化定理の意味 (重要!!)

定常無記憶情報源 X に対して，

1. 平均符号語長が $H(X)$ より小さい語頭符号は作れない (逆定理)
2. 平均符号語長 $\bar{L}$ が次の不等式を満足する語頭符号が存在 (順定理 1)

$$H(X) \leq \bar{L} < H(X) + 1$$

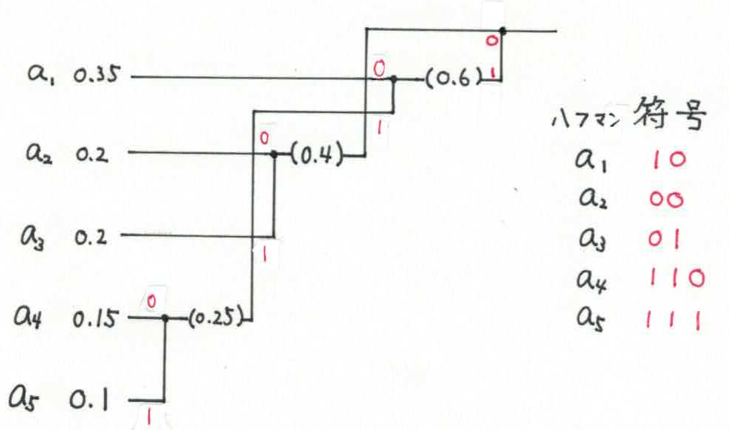
3. 次を満足する語頭符号が存在する (順定理 2)
   - 情報源の出力系列を n シンボル毎に区切って符号化
   - $\bar{L}_n$：1 シンボルあたりの平均符号語長

$$H(X) \leq \bar{L}_n < H(X) + \frac{1}{n}$$

$\Rightarrow$ n を十分を大きくとれば，平均符号語長がエントロピーに漸近

今回のまとめ 2

ハフマン符号の構成



ハフマン符号はコンパクト符号