6.符号化法（6章）

（情報源の）符号化定理と符号化法

情報源符号化定理（シャノンの第１定理）は、符号化の理論的限界を与えているだけであり、具体的な符号化の手法を与えてはいない。すなわち、ある情報源に対して、エントロピーは平均符号長の目標にすぎない。

ここでは、具体的な符号化法を与える。

代表的(情報源)符号化法

- シャノン・ファノ符号化(算術符号化)

確率を2進数化して符号化する。

- ハフマン符号化(コンパクト符号化)

最小の平均符号長を持つ符号
(コンパクト符号)を実現する符号。
符号の木の葉から符号を割り当てていく。

P進数と基数変換
（シャノン・ファノ符号化法の準備）

p進数

p 種類の記号 $\{0,1,\cdots,p-1\}$ を基に、
数を表現することができる。小数点以上n桁、小数点以下m桁の p 進数

$$(q_{n-1}q_{n-2}\cdots q_0 . q_{-1}q_{-2}\cdots q_{-m})_p$$

基数を明示する表記

は以下の値を持つ。ただし、$q_i \in \{0,1,\cdots,p-1\}$

$$\sum_{i=-m}^{n} q_i \times p^i$$

$$= q_{n-1}p^{n-1} + q_{n-2}p^{n-2} + \cdots + q_0 p^0$$

整数部（小数点の左の数、1以上）

$$+ q_{-1}p^{-1} + q_{-2}p^{-2} + \cdots + q_{-m}p^{-m}$$

小数部（小数点の右の数、1未満）

10進数と2進数

10^{n-1} 10 $10^0 = 1$ $10^{-1} = \frac{1}{10}$ $10^{-m} = \frac{1}{10^m}$

$$D = (d_{n-1}d_{n-2}\cdots d_1d_0.d_{-1}d_{-2}\cdots d_{-m})_{10}$$

$d_i \in \{0,1,2,3,4,5,6,7,8,9\}$

各桁の数字は、その位の値が何個あるのかを示している。

小数点の位置が重要

基数を明示する表記（通常は10の省略）

$b_i \in \{0,1\}$

$$B = (b_{s-1}b_{s-2}\cdots b_1b_0.b_{-1}b_{-2}\cdots b_{-t})_2$$

2^{s-1} 2 $2^0 = 1$ $2^{-1} = \frac{1}{2}$ $2^{-t} = \frac{1}{2^t}$

$$(19)_{10} = 1 \times 10^1 + 9 \times 10^0$$

$$= 10 + 9$$

$$\Updownarrow$$

$$16 + 2 + 1$$

$$= 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$$

$$= (10011)_2$$

練習

次の値を求めよ。

（１）

$$(11011.01)_2$$

（２）

$$(2102.2)_3$$

（３）

$$(43.21)_5$$

（４）

$$(72.3)_8$$

10進数から2進数への変換1

入力　　　　　　　　　　　　出力

$$D^{+} = (d_{n-1}d_{n-2}\cdots d_1 d_0)_{10} \Rightarrow B^{+} = (b_{s-1}b_{s-1}\cdots b_1 b_0)_2$$

2進数変換アルゴリズム(整数部分)

[step1]: $D_0^{+} := D^{+}, i = 0$とする。

初期設定

[step2]: $D_i^{+} \neq 0$ の間以下を繰り返す;

(2-1) $b_i := D_i^{+} \bmod 2$

2で割った余り

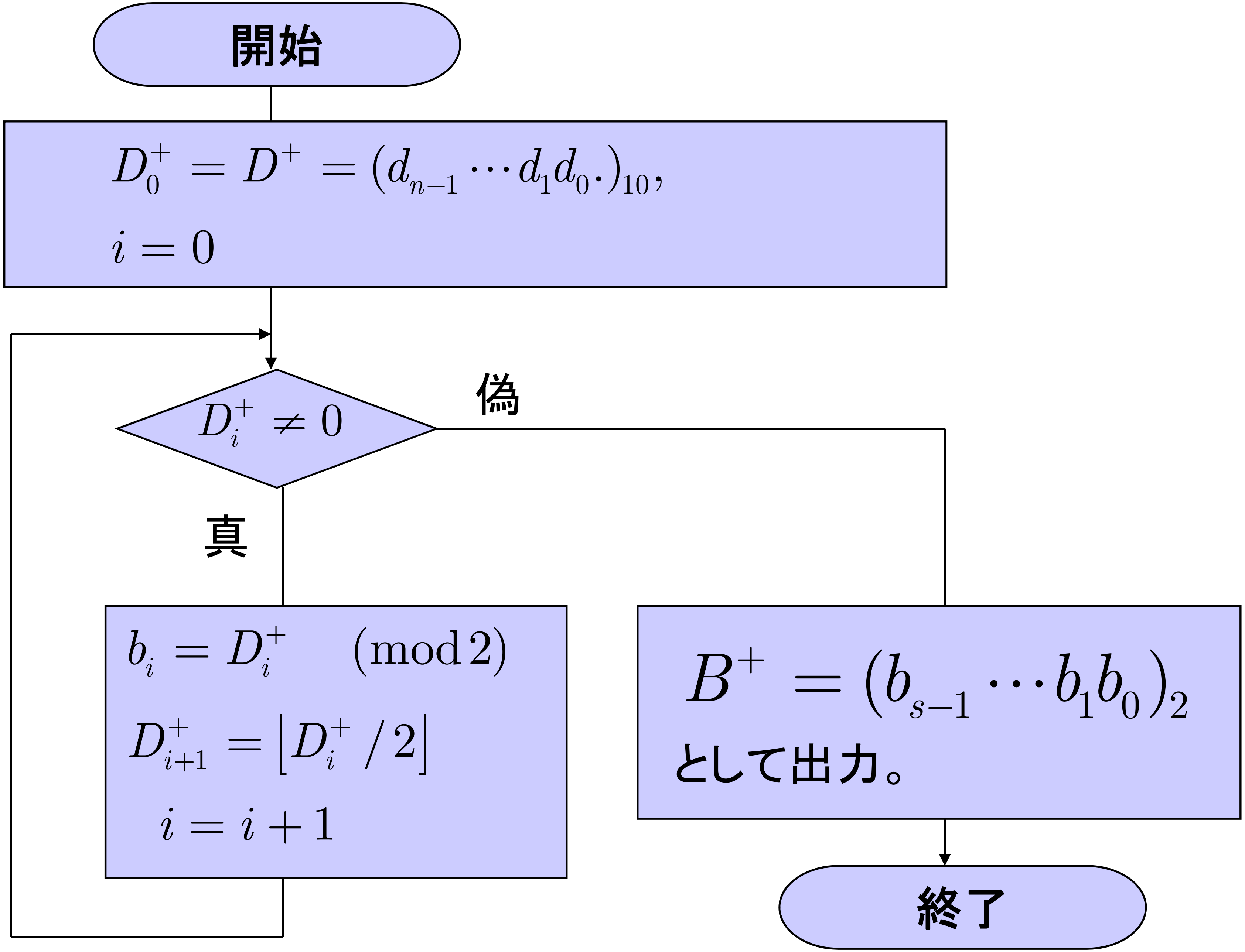
(2-2) $D_{i+1}^{+} := \lfloor D_i^{+} / 2 \rfloor$

2で割った商(切り捨て)

(2-3) $i := i + 1$

i を1増加させる。

[step3]: $B^{+} = (b_{s-1}\cdots b_1 b_0)_2$ を出力して終了する。


開始
$D_0^+ = D^+ = (d_{n-1} \cdots d_1 d_0.)_{10}$,
$i = 0$
$D_i^+ \neq 0$
偽
真
$b_i = D_i^+ \pmod 2$
$D_{i+1}^+ = \lfloor D_i^+ / 2 \rfloor$
$i = i + 1$
$B^+ = (b_{s-1} \cdots b_1 b_0)_2$
として出力。
終了

例

$(54)_{10}$ を2進数に変換せよ。

```
2)54
2)27   ・・・0
2)13   ・・・1
2) 6   ・・・1
2) 3   ・・・0
2) 1   ・・・1
   0   ・・・1
```

$2)\ D_0^+$

$2)\ D_1^+ \quad \cdots \quad b_0$

$2)\ D_2^+ \quad \cdots \quad b_1$

$2)\ D_{s-1}^+ \quad \cdots \quad b_{s-2}$

$0 \quad \cdots \quad b_{s-1}$

よって、

$$(54)_{10} = (110110)_2$$

練習

次の10進数を2進数に変換せよ。

（1）

$$(35)_{10}$$

（2）

$$(63)_{10}$$

（3）

$$(48)_{10}$$

（4）

$$(41)_{10}$$

アルゴリズムの正当性

$$D^{+} = D_0^{+} = b_{s-1} \times 2^{s-1} + \cdots + b_1 \times 2^1 + b_0 \times 2^0$$

$$= \underbrace{\underbrace{\left(\left(\underbrace{\left(\underbrace{\left(b_{s-1}\right)}_{D_{s-1}} \times 2 + b_{s-2}\right)}_{D_{s-2}} \times 2 + \cdots\right) \times 2 + b_1\right)}_{D_1} \times 2 + b_0}_{D_0}$$

$$= D_1 \times 2 + b_0$$

除算における商と余りの関係式

$$
\begin{aligned}
(19)_{10} &= (9)_{10} \times 2 + 1 \\
&= \Big((4)_{10} \times 2 + 1 \Big) \times 2 + 1 \\
&= \Big(\big((2)_{10} \times 2 + 0 \big) + 1 \Big) \times 2 + 1 \\
&= \Big(\Big(\big((1)_{10} \times 2 + 0 \big) \times 2 + 0 \Big) \times 2 + 1 \Big) \times 2 + 1 \\
&= \Big(\Big(\big((1)_{2} \times 2 + 0 \big) \times 2 + 0 \Big) \times 2 + 1 \Big) \times 2 + 1 \\
&= \Big(\big((10)_{2} \times 2 + 0 \big) \times 2 + 1 \Big) \times 2 + 1 \\
&= \big((100)_{2} \times 2 + 1 \big) \times 2 + 1 \\
&= (1001)_{2} \times 2 + 1 = (10011)_{2}
\end{aligned}
$$

奇数なので1余る

性質:整数部分基数変換アルゴリズムのループ不変条件

$B_i^+ \equiv (b_{s-1} b_{s-2} \cdots b_i)_2$ とおく。

ループの i 回の繰り返しにおいて次式が成り立つ。

このような式を
ループ不変条件という。

$$D_i^+ = B_i^+$$

アルゴリズムの途中で現れる10進数

先頭の$s-i$桁を抜き出した2進数。(途中では、求まっていないことに注意する。)

証明

繰り返し回数 i に関する帰納法で証明する。

($i = 0$ のとき。)

$$(d_{n-1} \cdots d_0) = D^+ = B^+ = (b_{s-1} \cdots b_0)$$

であるので成り立つ。

アルゴリズムの初期化により
正しい。

（$i > 0$ のとき。）

$i = k$ のとき $D_k^+ = B_k^+ = (b_{s-1} \cdots b_k)_2$ であると仮定する。

帰納法の仮定

$i = k + 1$ のとき

アルゴリズムの各繰り返しにより、$D_{k+1}^+ = \left\lfloor \frac{D_k^+}{2} \right\rfloor$ であるので、以下のように計算できる。

$$
\begin{aligned}
D_{k+1}^+ &= \left\lfloor \frac{D_k^+}{2} \right\rfloor \\
&= \left\lfloor \frac{B_k^+}{2} \right\rfloor \\
&= \left\lfloor \frac{(b_{s-1} \cdots b_k)}{2} \right\rfloor \\
&= (b_{s-1} \cdots b_{k+1})_2 \\
&= B_{k+1}^+
\end{aligned}
$$

アルゴリズムの動作

帰納法の仮定

定義

QED

性質：整数部分基数変換アルゴリズムの出力の正当性

$D^+ = B^+ = (b_{s-1} \cdots b_1 b_0)_2$ のとき
各 $i, 0 \leq i \leq s-1$ に対して、

$$b_i = D_i^+ \mod 2$$

証明

一般に、Y を X で割った商が Q で余りが R であるとき、
以下の式が成り立つ。

$$Y = QX + R \quad , 0 \leq R < X$$

$$\Leftrightarrow R = Y \pmod{X}$$

この関係より、各 $i, 0 \le i \le s-1$ に対して次ように計算できる。

ループ不変条件

$$D_i^+ = B_i^+$$

$$= (b_{s-1} \cdots b_i)_2$$

$$= b_{s-1} \times 2^{s-i-1} + \cdots + b_{i+1} \times 2^1 + b_i \times 2^0$$

$$= \left(b_{s-1} \times 2^{s-i-2} + \cdots + b_{i+1} \times 2^0\right) \times 2 + b_i$$

$$= B_{i+1}^+ \times 2 + b_i$$

$$\Leftrightarrow b_i = D_i^+ \pmod 2$$

アルゴリズムの動作

QED

10進数から2進数への変換2

入力

$$D^- = (0.d_{-1}d_{-2}\cdots d_{-m})_{10}$$

出力

$$B^- = (0.b_{-1}b_{-2}\cdots b_{-t})_2$$

2進数変換アルゴリズム（小数部分）

[step1]: $D^-_{-1} := D^-, i = -1$ とする。 初期設定

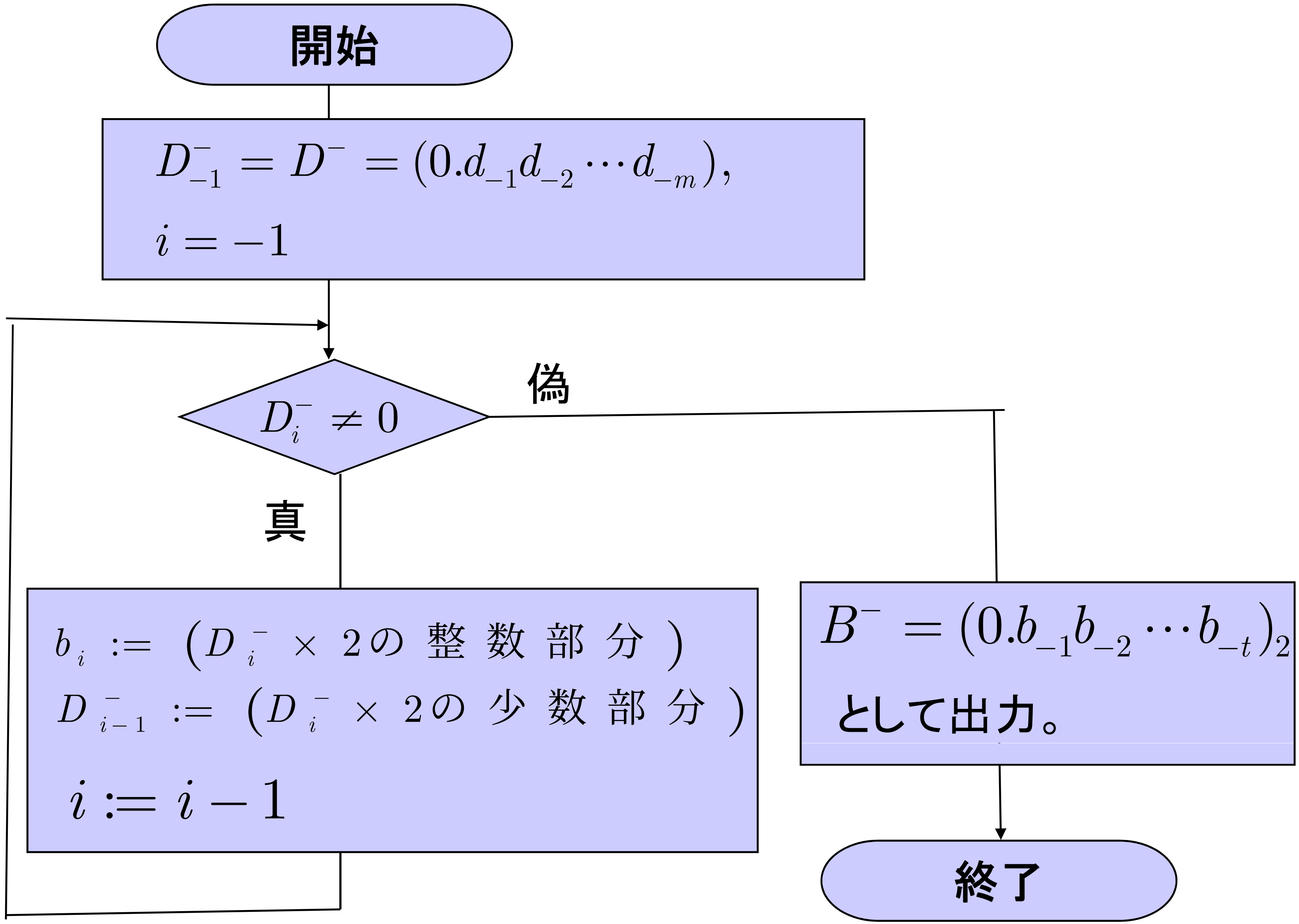
[step2]: $D^-_i \neq 0$ の間以下を繰り返す;

(2-1) $b_i :=$ ($D^-_i \times 2$の整数部分)

(2-2) $D^-_{i-1} :=$ ($D^-_i \times 2$の少数部分)

(2-3) $i := i - 1$

[step3]: $B^- = (0.b_{-1}b_{-2}\cdots b_{-t})_2$ を出力して終了する。


開始
$D^-_{-1} = D^- = (0.d_{-1}d_{-2}\cdots d_{-m})$,
$i = -1$
$D^-_i \neq 0$
偽
真
$b_i :=$ ($D^-_i \times 2$の整数部分)
$D^-_{i-1} :=$ ($D^-_i \times 2$の少数部分)
$i := i - 1$
$B^- = (0.b_{-1}b_{-2}\cdots b_{-t})_2$
として出力。
終了

例

$$(0.5625)_{10}$$ を2進数に変換せよ。

$$2 \times 0.5625 = 1.125 \cdots 1 + 0.125$$

$$2 \times 0.125 = 0.25 \cdots 0 + 0.25$$

$$2 \times 0.25 = 0.5 \cdots 0 + 0.5$$

$$2 \times 1 = 1.0 \cdots 1 + 0.0$$

$$2 \times \overline{D_{-1}} = b_{-1} + \overline{D_{-2}}$$

$$2 \times \overline{D_{-2}} = b_{-2} + \overline{D_{-3}}$$

$$\vdots$$

$$2 \times \overline{D_{-t}} = b_{-t} + 0.0$$

よって、

$$(0.5625)_{10} = (0.1001)_2$$

小数点に近い方から
順に求まる。

練習

次の10進数を2進数に変換せよ。

（1）

$$(0.625)_{10}$$

（2）

$$(0.53125)_{10}$$

（3）

$$(0.3)_{10}$$

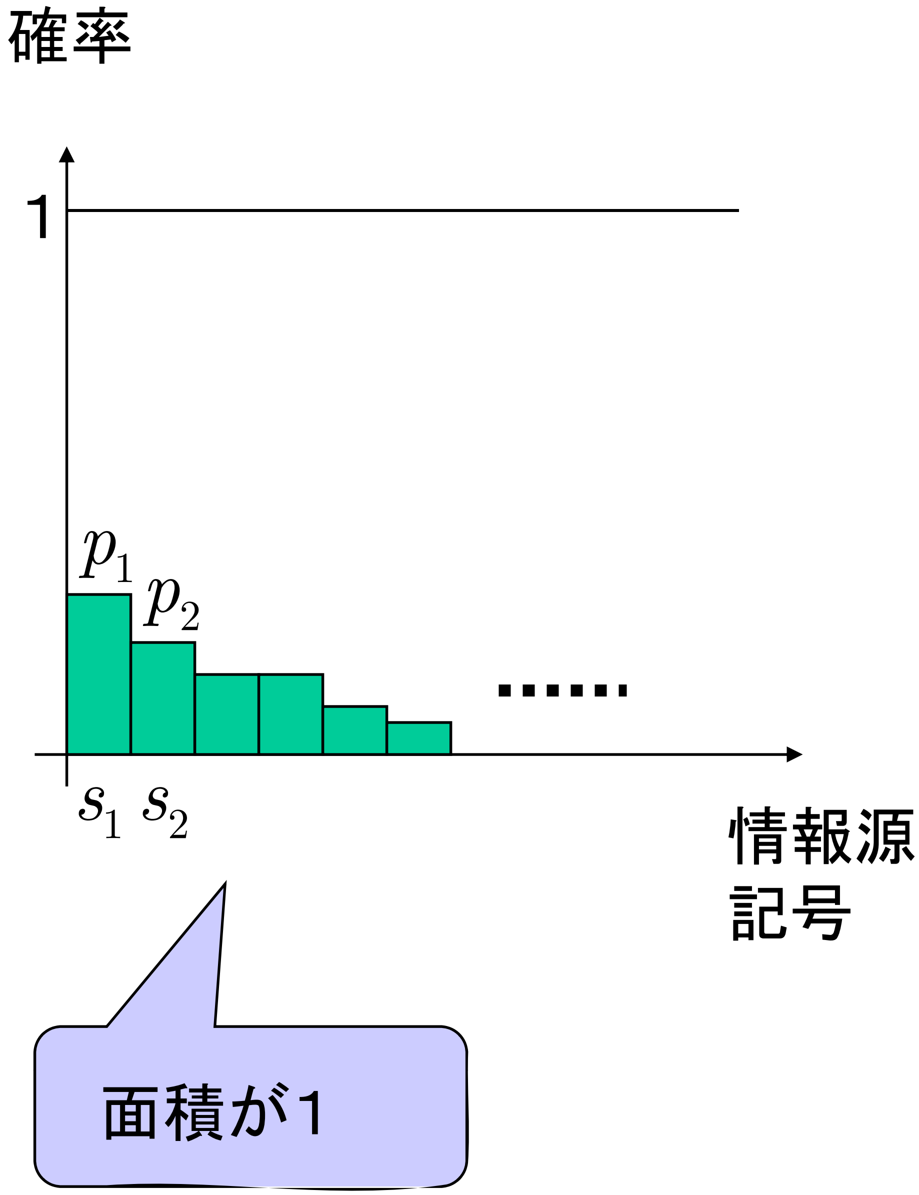
（4）

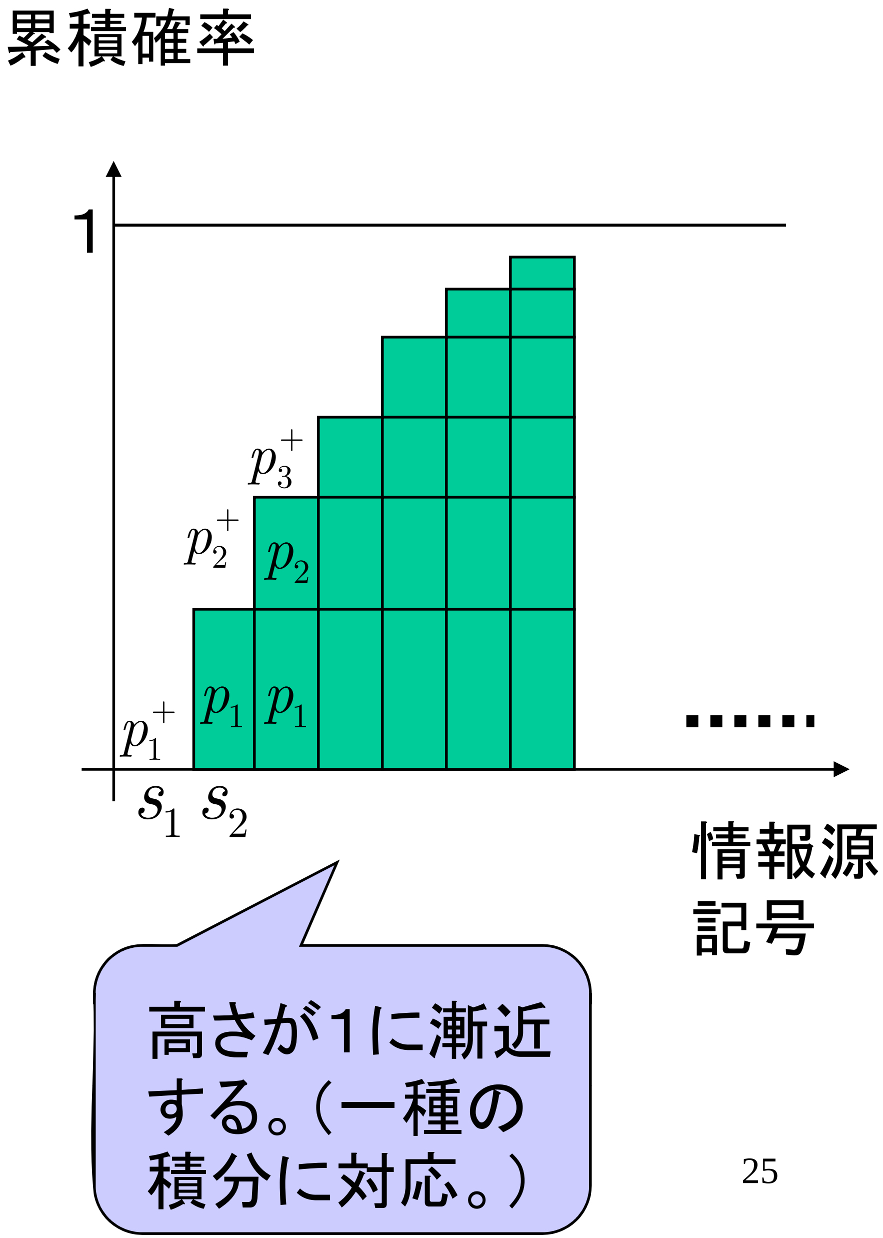
$$(0.59375)_{10}$$

練習

小数部分の基数変換アルゴリズムにおけるループ不変条件を示し、正当性を証明せよ。

累積確率


確率
1
p_1
p_2
s_1 s_2
情報源記号
面積が1


累積確率
1
p_3^+
p_2^+
p_2
p_1^+
p_1 p_1
s_1 s_2
情報源記号
高さが1に漸近する。(一種の積分に対応。)

シャノン・ファノ符号化法

シャノン・ファノ符号化

入力：情報源（情報源記号の集合とその発生確率）

$$S = \begin{Bmatrix} s_1 & , & \cdots & , & s_n \\ p_1 & , & \cdots & , & p_n \end{Bmatrix}$$

出力：符号（情報源記号に対応する符号語の集合）

$$\phi = \left\{ s_1 \mapsto \boldsymbol{c}_1, s_2 \mapsto \boldsymbol{c}_2, \cdots, s_n \mapsto \boldsymbol{c}_n \right\}$$

ステップ１：発生確率の大きい順に並べる。

（ここでは、添え字でこの順序が見たされるとする。）

ステップ２：各符号語長を次式で求める。

$$l_i = \lceil -\log p_i \rceil$$

切り上げ

ステップ３：次のような累積確率 p_i^+ を求める。

$$p_1^+ = 0 (i = 1), p_i^+ = \sum_{j=1}^{i-1} p_i (i \geq 2)$$

ステップ４：累積確率 p_i^+ を２進数 $(p_i^+)_2$ に変換する。

ステップ５：２進数 $(p_i^+)_2$ の上位 l_i 桁を符号語 c_i とする。

例

次の無記憶情報源 S に対して、シャノンファノ符号を構成する。

$$S = \begin{Bmatrix} a & , & b & , & c & , & d \\ 0.2 & , & 0.3 & , & 0.1 & , & 0.4 \end{Bmatrix}$$

ステップ１（降順に並び替え）

$$S = \begin{Bmatrix} d & , & b & , & a & , & c \\ 0.4 & , & 0.3 & , & 0.2 & , & 0.1 \end{Bmatrix}$$

$$\therefore s_1 = d, s_2 = b, s_3 = a, s_4 = c$$

ステップ2（符号語長の決定）

$$d : -\log 0.4 \simeq 1.322$$

$$\therefore l_1 = l_d = \lceil -\log 0.4 \rceil = 2$$

$$b : -\log 0.3 \simeq 1.737$$

$$\therefore l_2 = l_b = \lceil -\log 0.3 \rceil = 2$$

$$a : -\log 0.2 \simeq 2.322$$

$$\therefore l_3 = l_a = \lceil -\log 0.2 \rceil = 3$$

$$c : -\log 0.1 \simeq 3.322$$

$$\therefore l_4 = l_c = \lceil -\log 0.1 \rceil = 4$$

したがって、

$$L = \{l_1, l_2, l_3, l_4\}$$

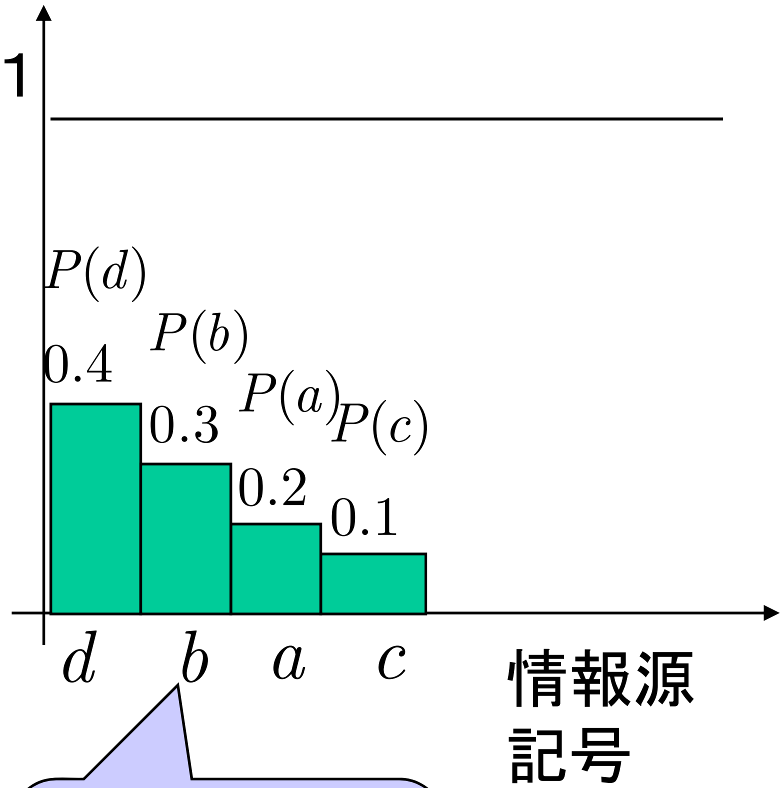
$$= \{2, 2, 3, 4\}$$

の符号語長を持つ符号を構成する。

ステップ3（累積確率の計算）

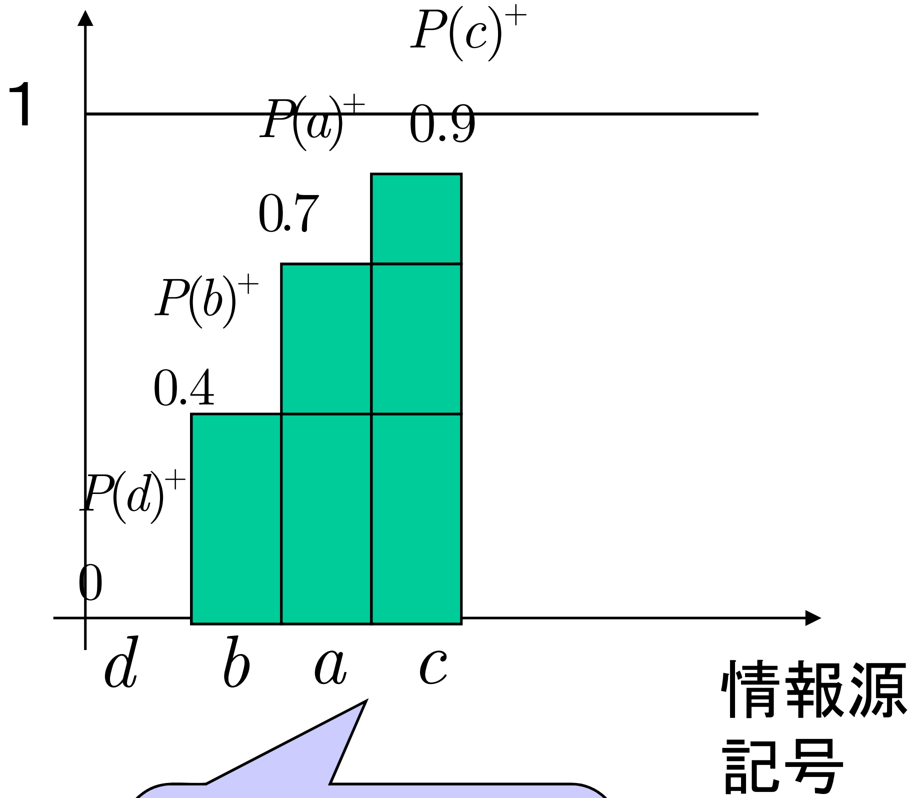
情報源記号	符号語長	確率	累積確率
d	$l_1 = 2$	$p_1 = P(d) = 0.4$	${p_1}^+ = 0.0$
b	$l_2 = 2$	$p_2 = P(b) = 0.3$	${p_2}^+ = 0.4$
a	$l_3 = 3$	$p_3 = P(a) = 0.2$	${p_3}^+ = 0.7$
c	$l_4 = 4$	$p_4 = P(c) = 0.1$	${p_4}^+ = 0.9$

確率



確率の降順
面積が1

累積確率



高さが1に漸近
する。（一種の
積分に対応。）

ステップ4（累積確率の２進数化）

$$\left(p_1{}^{+}\right)_{10} = (0.0)_{10} \simeq (0.00000)_2$$

$$\left(p_2{}^{+}\right)_{10} = (0.4)_{10} \simeq (0.01100)_2$$

$$\left(p_3{}^{+}\right)_{10} = (0.7)_{10} \simeq (0.10110)_2$$

$$\left(p_4{}^{+}\right)_{10} = (0.9)_{10} \simeq (0.11100)_2$$

ステップ５（符号の割り当て）

$$d \mapsto 00$$

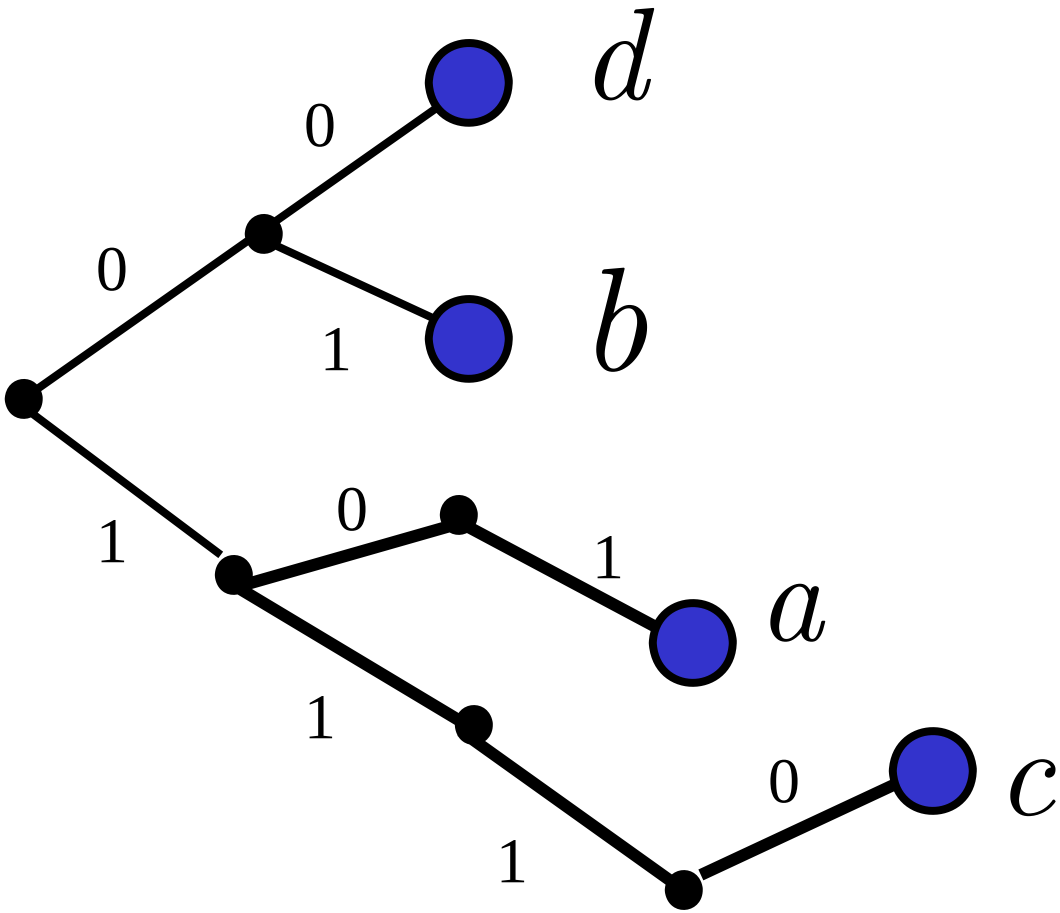
$$b \mapsto 01$$

$$a \mapsto 101$$

$$c \mapsto 1110$$

$$\therefore \phi = \left\{ d \mapsto 00, b \mapsto 01, a \mapsto 101, c \mapsto 1110 \right\}$$

符号の木



平均符号長

$$\overline{L} = 2 \times 0.4 + 2 \times 0.3 + 3 \times 0.2 + 4 \times 0.1$$
$$= 2.4$$

練習

次の情報源ジャンをシャノンファノの符号化法に従って符号化せよ。

$$\text{ジャン} = \begin{Bmatrix} \text{グー} & , & \text{チョキ} & , & \text{パー} \\ 0.35 & , & 0.25 & , & 0.4 \end{Bmatrix}$$

また、得られた符号に対する符号の木を示し、平均符号長、効率を求めよ。

シャノンファノ符号化法の平均符号長

$1 \le i \le n$ に対して、

$$l_i = \lceil -\log p_i \rceil$$

であるが、これは次式を満たす。

$$-\log p_i \le l_i < -\log p_i + 1$$

どこかで見た式

したがって、平均符号長は次式で求められる。

$$-\sum_{i=1}^{n} p_i \log p_i \le \sum_{i=1}^{n} p_i l_i < -\sum_{i=1}^{n} p_i \log p_i + \sum_{i=1}^{n} p_i$$

$$\therefore H(S) \le \overline{L} < H(S) + 1$$

シャノンファノ符号の非特異性

シャノンファノ符号化法で構成された符号
は非特異符号である。

証明

符号語長の選び方より、次式が成り立つ。

$$-\log p_i \le l_i < -\log p_i + 1$$

$$\therefore 2^{-l_i} \le p_i < 2^{-l_i+1}$$

一方、

記号 s_{i-1} の生成確率

$$p_i^+ - p_{i-1}^+ = p_{i-1}$$ に注意する。

累積確率の i 番目

累積確率の $i-1$ 番目

よって、２進数表現して、次式が成り立つ。

$$\left(p_i^+\right)_2 - \left(p_{i-1}^+\right)_2 = \left(p_{i-1}\right)_2$$

$$\geq 2^{\log p_{i-1}}$$

$$\geq 2^{-l_{i-1}}$$

第 $-l_{i-1}$ 桁で必ず異なることを意味する。

Q.E.D

縮退情報源
（ハフマン符号化法の準備）

縮退情報源

情報源　S 対して、
S　の2つ以上の情報源記号 $s_i, s_j \in S$　を一つにまとめた情報源を元の情報源の縮退情報源という。すなわち、

$$S = \begin{Bmatrix} s_1 & , & \cdots & , & s_i & , & s_j & , & \cdots & , & s_n \\ p_2 & , & \cdots & , & p_i & , & p_j & , & \cdots & , & p_n \end{Bmatrix}$$

$$S^- = \begin{Bmatrix} s_1 & , & \cdots & , & {}^*s_k & , & \cdots & , & s_n \\ p_1 & , & \cdots & , & {}^*p_k & , & \cdots & , & p_n \end{Bmatrix}$$

ここで、${}^*s_k = \{s_i, s_j\}$　　${}^*p_k = p_i + p_j$

$$\left|S^-\right| = \left|S\right| - 1$$

例

次の情報源の縮退情報源をいくつか示せ。

$$S = \begin{Bmatrix} a & , & b & , & c & , & d \\ 0.2 & , & 0.3 & , & 0.1 & , & 0.4 \end{Bmatrix}$$

解）

$$S_1^- = \begin{Bmatrix} A & , & c & , & d \\ 0.5 & , & 0.1 & , & 0.4 \end{Bmatrix}, A = \{a, b\}$$

$$S_2^- = \begin{Bmatrix} a & , & B & , & d \\ 0.2 & , & 0.4 & , & 0.4 \end{Bmatrix}, B = \{b, c\}$$

$$S_3^- = \begin{Bmatrix} C & , & b & , & d \\ 0.3 & , & 0.3 & , & 0.4 \end{Bmatrix}, C = \{a, c\}$$

練習

次の情報源に対して、2記号を1記号に縮退して得られる情報源をすべて示せ。

$$S = \begin{Bmatrix} a & , & b & , & c & , & d \\ 0.15 & , & 0.25 & , & 0.05 & , & 0.55 \end{Bmatrix}$$

ハフマン符号化法

ハフマン符号化

入力:情報源(情報源記号の集合とその発生確率)

$$S = \begin{Bmatrix} s_1 & , & s_2 & , & \cdots & , & s_n \\ p_1 & , & p_2 & , & \cdots & , & p_n \end{Bmatrix}$$

出力:符号(情報源記号と符号語の対応の集合)

$$\phi = \left\{ s_1 \mapsto \boldsymbol{c}_1, s_2 \mapsto \boldsymbol{c}_2, \cdots, s_n \mapsto \boldsymbol{c}_n \right\}$$

ステップ1: $k = 0$ とし、$S_0 = S = \begin{Bmatrix} s_1^0 & , & \cdots & , & s_n^0 \\ p_1^0 & , & \cdots & , & p_n^0 \end{Bmatrix}$ とする。

ここで、 $s_i^0 = s_i, p_i^0 = p_i, \quad (1 \leq i \leq n)$

ステップ2: 発生確率の大きい順に並べる。
(添え字の順序をこの順序とする。)

ステップ3：縮退情報源

$n-k$ 記号の情報源に縮退されている。

$$S_k = \begin{Bmatrix} s_1^k & , & \cdots & , & s_{n-k-1}^k & , & s_{n-k}^k \\ p_1^k & , & \cdots & , & p_{n-k-1}^k & , & p_{n-k}^k \end{Bmatrix}$$

に対して、確率の小さい2つの情報源記号 s_{n-k-1}^k, s_{n-k}^k に対して、対応する符号の末尾に0と1を割り当てる。さらに、s_{n-k-1}^k, s_{n-k}^k を縮退して、新たな縮退情報源

$$S_{k+1} = \begin{Bmatrix} s_1^{k+1} & , & \cdots & , & {}^*p_{n-(k+1)}^{k+1} \\ p_1^{k+1} & , & \cdots & , & {}^*p_{n-(k+1)}^{k+1} \end{Bmatrix}$$

を作成する。 ここで、

$$s_i^{k+1} = s_i^k, p_i^{k+1} = p_i^k, \quad (1 \le i \le n-k-2)$$

$$s_{n-(k+1)}^{k+1} = \left\{ s_{n-k-1}^k, s_{n-k}^k \right\}, p_{n-(k+1)}^{k+1} = p_{n-k-1}^k + p_{n-k}^k, \quad i = n-k-1$$

ステップ4： $k < n-1$ である限り、$k = k+1$ として
ステップ2に戻る。

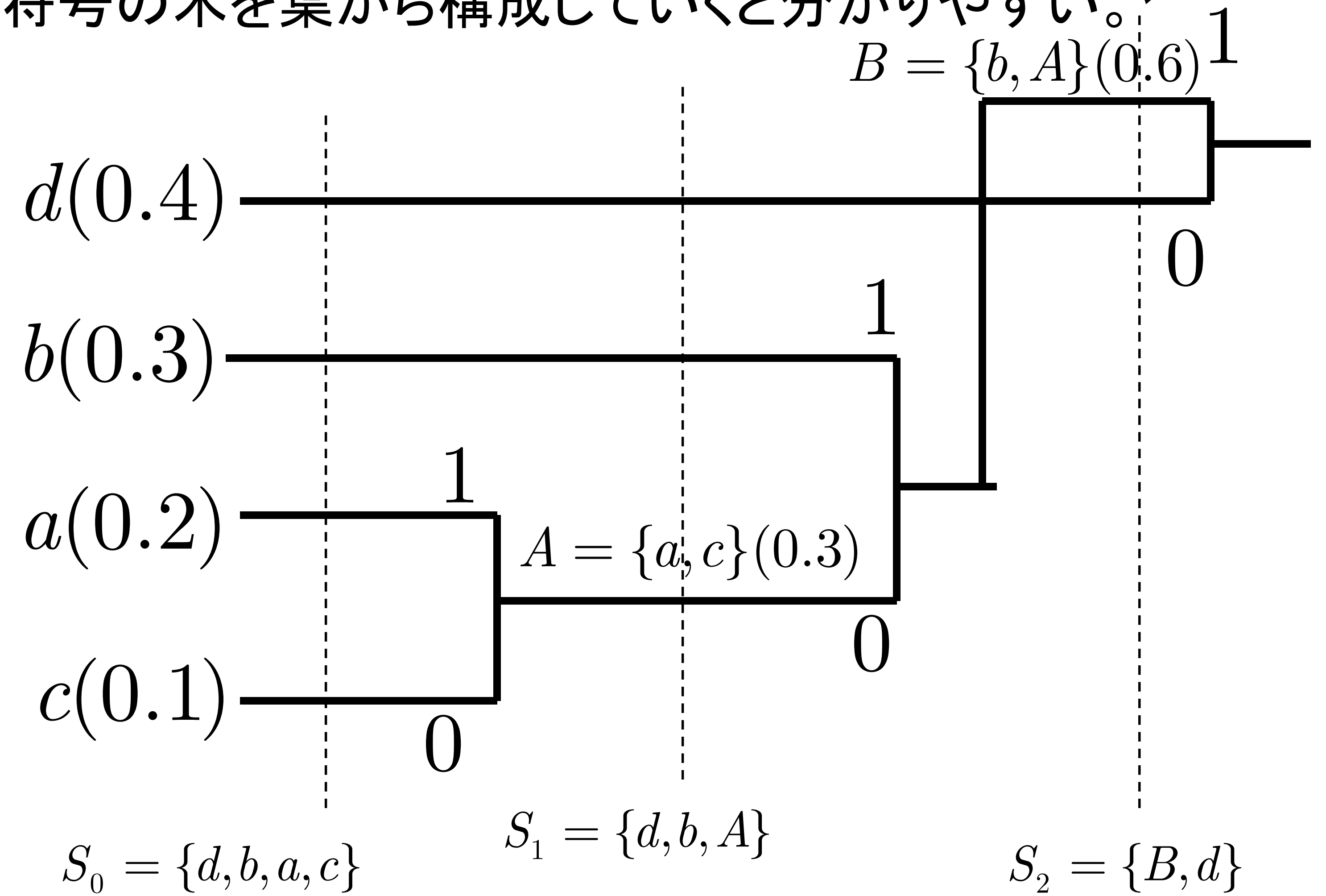
ハフマン符号化例

次の符号のハフマン符号を与える。

$$S = \begin{Bmatrix} a & , & b & , & c & , & d \\ 0.2 & , & 0.3 & , & 0.1 & , & 0.4 \end{Bmatrix}$$

符号の木を葉から構成していくと分かりやすい。

全ての点線で、縮退情報源の確率が降順になっていること。



前のスライドの符号の木より、

$$d \mapsto 0$$

$$b \mapsto 11$$

$$a \mapsto 101$$

$$c \mapsto 100$$

平均符号長 $\overline{L}$ は、次のように計算される。

$$\overline{L} = 0.4 \times 1 + 0.3 \times 2 + 0.2 \times 3 + 0.1 \times 3$$

$$= 1.9$$

練習

次の符号をハフマン符号化し、
符号の木、平均符号長、効率を求めよ。

$$S = \begin{Bmatrix} a & , & b & , & c & , & d \\ 0.15 & , & 0.25 & , & 0.05 & , & 0.55 \end{Bmatrix}$$

コンパクト符号

（定義）コンパクト符号

情報源に対して、符号語長が最短となる符号を
コンパクト符号という。

コンパクト符号であっても、効率が1
になるとは限らない。効率が1なら明
らかにコンパクト符号である。

ハフマン符号のコンパクト性

性質

ハフマン符号は、コンパクト符号である。

ハフマン符号化で得られる符号は、1通りとは限らない。しかし、どのようなハフマン符号もコンパクトとなる。

この証明のために、次の補題を示す。

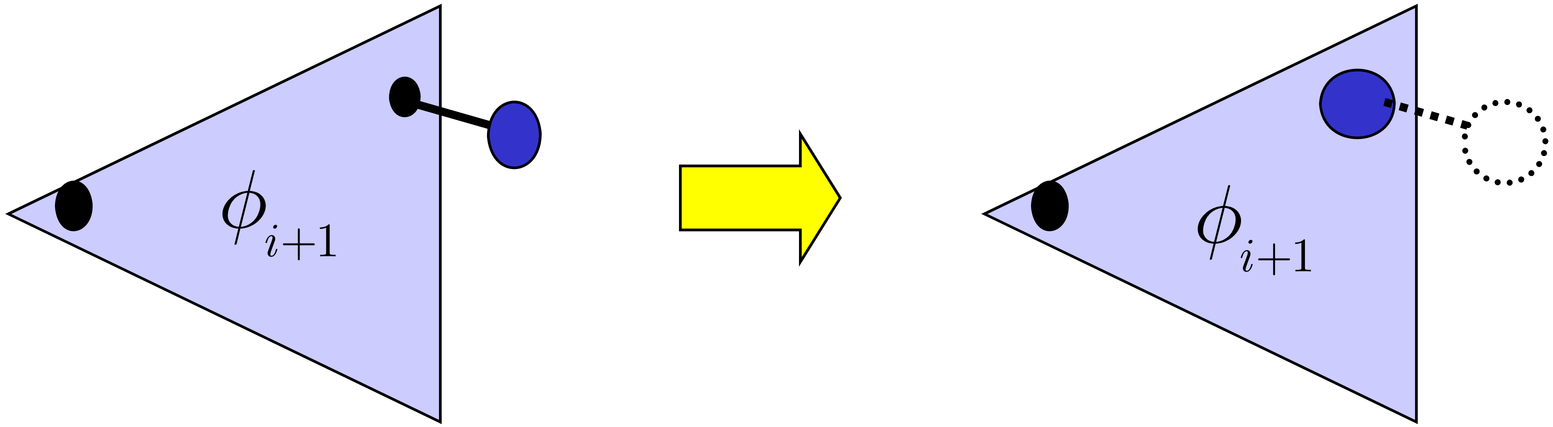
補題

情報源 S_i の最小の生成確率を持つ2記号を縮退して得られる情報源を S_{i+1} とする。情報源 S_i のハフマン符号を ϕ_i とし、情報源 S_{i+1} のハフマン符号を ϕ_{i+1} とする。このとき、ϕ_{i+1} がコンパクト符号ならば、ϕ_i もコンパクト符号である。

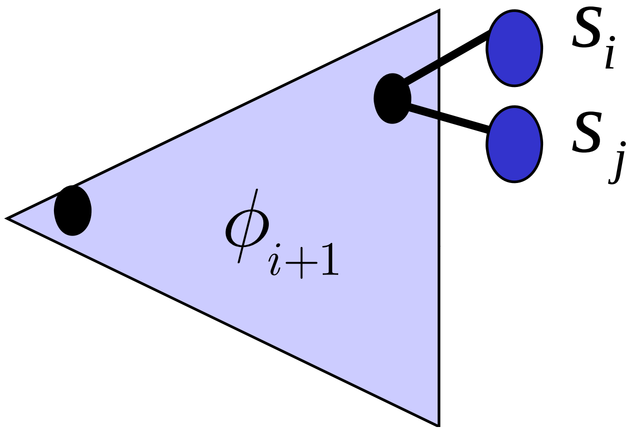
証明

ϕ_i の平均符号長を $\overline{L_i}$ とし、
ϕ_{i+1} の平均符号長を $\overline{L_{i+1}}$ とする。

まず、符号の木の経路長が最長の葉は2つある。
（もし、経路長が最長の葉が1つだけなら、さらに短くできる。）

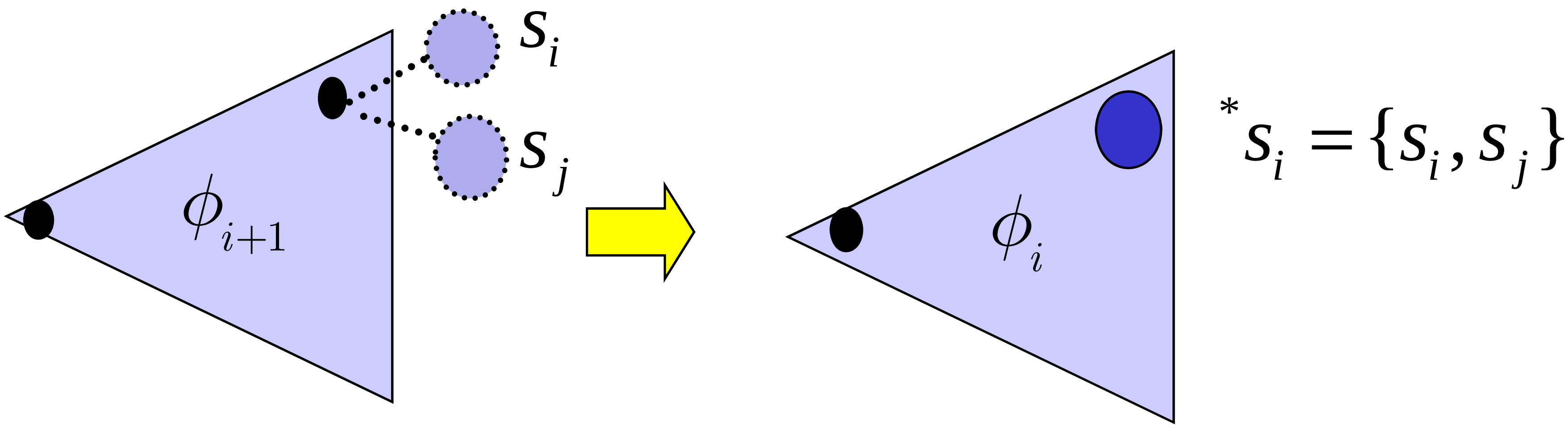


この2つの葉に発生確率最小の2記号　$s_i, s_j \in S_i$　を割り当てる。（もし、発生確率が最小以外の記号を割り当てるとより短い平均符号長が得られる。）



$$\overline{L_i} = \sum_{s_k \in S_i} p_k l_k$$

$$\le \sum_{s_k \in S_i} p_k l_k - p_i l_i - p_j l_j + p'_i \, l_i + p'_j \, l_j$$



$$\overline{L_{i+1}} = \overline{L_i} - p_i l_i - p_j l_i + (p_i + p_j)(l_i - 1)$$
$$= \overline{L_i} - p_i - p_j$$

今、S_{i+1}に対する任意の符号の平均符号長を $\overline{L'_{i+1}}$ と表す。

このとき、 $\overline{L_{i+1}}$ のコンパクト性より、以下が成り立つ。

$$\overline{L_{i+1}} \leq \overline{L'_{i+1}}$$

同じ値を両辺から引いているだけ。

$$\therefore \quad \overline{L_{i+1}} - p_i - p_j \le \overline{L'_{i+1}} - p_i - p_j$$

$$\therefore \quad \overline{L_i} \le \overline{L'_i}$$

情報源 S_{i+1} の平均符号長と等しい。

したがって、 ϕ_i もコンパクト符号になる。

Q.E.D

（ハフマン符号のコンパクト性の証明）

基礎

$i = n - 2$ 　のとき。

このとき、縮退情報源の記号の数は、

$$\left|S_i\right| = \left|S_{n-2}\right| = 2$$

であり、ハフマン符号は明らかにコンパクト符号である。

帰納

$n - 2 \geq i > 0$ のすべての i に対して、ハフマン符号 ϕ_i が情報源 S_i のコンパクト符号だと仮定する。（帰納法の仮定）

先の補題より、

ϕ_i がコンパクト符号ならば、ϕ_{i-1} もコンパクト符号である。

したがって、S_0 をハフマン符号化した ϕ_0 もコンパクト符号である。

Q.E.D