



情報理論

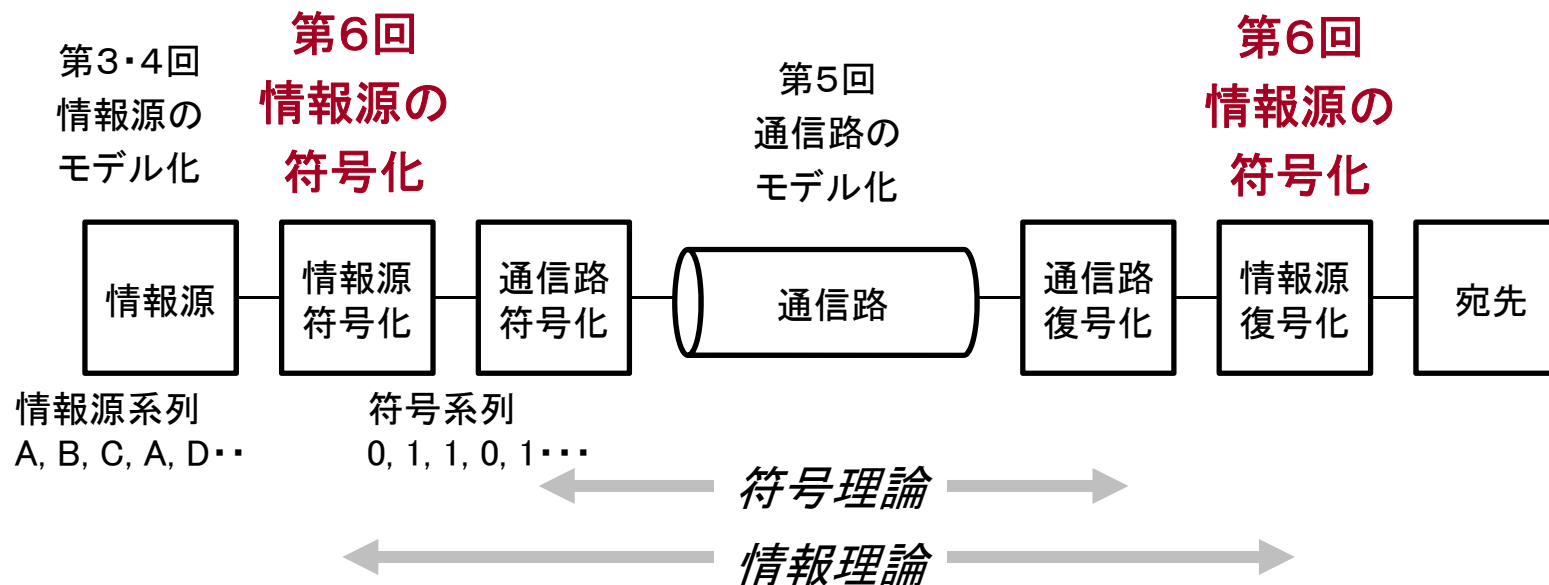
第6回 講義 情報源の符号化

2015. 5. 27

植松 芳彦

本日の講義範囲と内容

1. 情報源符号化に必要な条件
2. 瞬時符号と符号の木
3. 瞬時符号と符号長の関係





情報源符号化とは

- 情報源符号化とは、情報源から発生する情報源記号列を、一定の規則に基づき伝送や記録ができる符号系列に変換すること.
- 情報源復号化とは、受信した符号系列から元々の情報源記号列を再生すること.

情報源記号例) アルファベット, スペース, カンマ...

符号系列例) 0,1の並び

情報源符号化とは

- 各アルファベットに対する 0, 1 の列の割り当て方と捉ええると, やり方は無限に存在.
- どのような割り当て方がよいか考える.

表 4.1 情報源符号化の例

情報源記号	確率	C_1	C_2	C_3	C_4	C_5	C_6
A	0.6	00	0	0	0	0	0
B	0.25	01	10	10	01	10	10
C	0.1	10	110	110	011	11	11
D	0.05	11	1110	111	111	01	0
平均符号長		2.00	1.60	1.55	1.55	1.40	1.35

第一回講義で考えた



情報源符号化に必要な条件

- 各情報源記号に割り当てる符号長の平均値は短いほどよい(平均符号長)
- 受信側(復号側)で解釈を誤ったり, 解釈を迷ったりしない程よい(一意復号性)
- (符号化, 復号装置があまり複雑にならない)

情報源符号化則 C I

- 平均符号長 : $0.6 \times 2 + 0.25 \times 2 + 0.1 \times 2 + 0.05 \times 2 = 2$
- 一意復号性 : 一意復号可能(瞬時)
- その他 : **等長符号**

情報源記号	発生確率	符号則C I
A	0.6	0 0
B	0.25	0 1
C	0.1	1 0
D	0.05	1 1

初めの切れ目だけは
分っているとする



0 0 0 1 1 0 1 1
└─┘ └─┘ └─┘ └─┘
A B C D

2ビット受信すれば
直ちに元の情報源記号を判定可能

情報源符号化則 C II

- 平均符号長 : $0.6 \times 1 + 0.25 \times 2 + 0.1 \times 3 + 0.05 \times 4 = 1.6$
- 一意復号性 : 一意復号可能(瞬時)
- その他特徴 : **コンマ符号**

情報源記号	発生確率	符号則C II	
A	0.6	0	0
B	0.25	1 0	1 0
C	0.1	1 1 0	1 1 0
D	0.05	1 1 1 0	1 1 1 0

0 1 0 1 1 0 1 1 1 0...

A B C D

0を受信した段階で
直ちに元の情報源記号を判定可能

情報源符号化則 CⅢ

- 平均符号長 : $0.6 \times 1 + 0.25 \times 2 + 0.1 \times 3 + 0.05 \times 3 = 1.55$
- 一意復号性 : 一意復号可能(瞬時)
- その他特徴 : 特になし

情報源記号	発生確率	符号則CⅢ
A	0.6	0
B	0.25	1 0
C	0.1	1 1 0
D	0.05	1 1 1

0 1 1 1 1 1 1 0...
 □ □ □ □ □ □ □ □
 A D D A

・0を受信した段階 または
 ・1を3つ受信した段階 で
 直ちに元の情報源記号を判定可能

情報源符号化則 CIV

- 平均符号長 : $0.6 \times 1 + 0.25 \times 2 + 0.1 \times 3 + 0.05 \times 3 = 1.55$
- 一意復号性 : 一意復号可能(非瞬時)
- その他特徴 : 特になし

情報源記号	発生確率	符号則CIV
A	0.6	0
B	0.25	0 1
C	0.1	0 1 1
D	0.05	1 1 1

$\boxed{0} \quad \boxed{1 \ 1 \ 1} \quad \boxed{1 \ 1 \ 1} \quad \boxed{0} \dots$
 A? D? D? ←あたり

$\boxed{0 \ 1} \quad \boxed{1 \ 1 \ 1} \quad \boxed{1 \ 1} \quad \boxed{0} \dots$
 B? D? ?? はずれ

$\boxed{0 \ 1 \ 1} \quad \boxed{1 \ 1 \ 1} \quad \boxed{1} \quad \boxed{0} \dots$
 C? D? ??

情報源符号化則 CV

- 平均符号長 : $0.6 \times 1 + 0.25 \times 2 + 0.1 \times 2 + 0.05 \times 2 = 1.40$
- 一意復号性 : 一意復号不可能
- その他特徴 : 特になし

情報源記号	発生確率	符号則 CV
A	0.6	0
B	0.25	1 0
C	0.1	1 1
D	0.05	0 1

0 1 1 0 ..
 □ □ □

A C A

0 1 1 0 ..
 □ □

D B

この後何を受信しても
 どちらか決まらない

情報源符号化則 CVI

- 平均符号長 : $0.6 \times 1 + 0.25 \times 2 + 0.1 \times 2 + 0.05 \times 1 = 1.35$
- 一意復号性 : 一意復号不可能
- その他特徴 : 特異符号

情報源記号	発生確率	符号則 CVI
A	0.6	0
B	0.25	1 0
C	0.1	1 1
D	0.05	0

0 0 0 0 . .
□ □ □ □

それぞれがAかDか
全く分からない

異なる情報源記号に同じ符号を割当

最もよい情報源符号化則は？

- 平均符号長, 一意復号性のバランスでCⅢ.
- 瞬時符号であるには何か条件が必要そう.

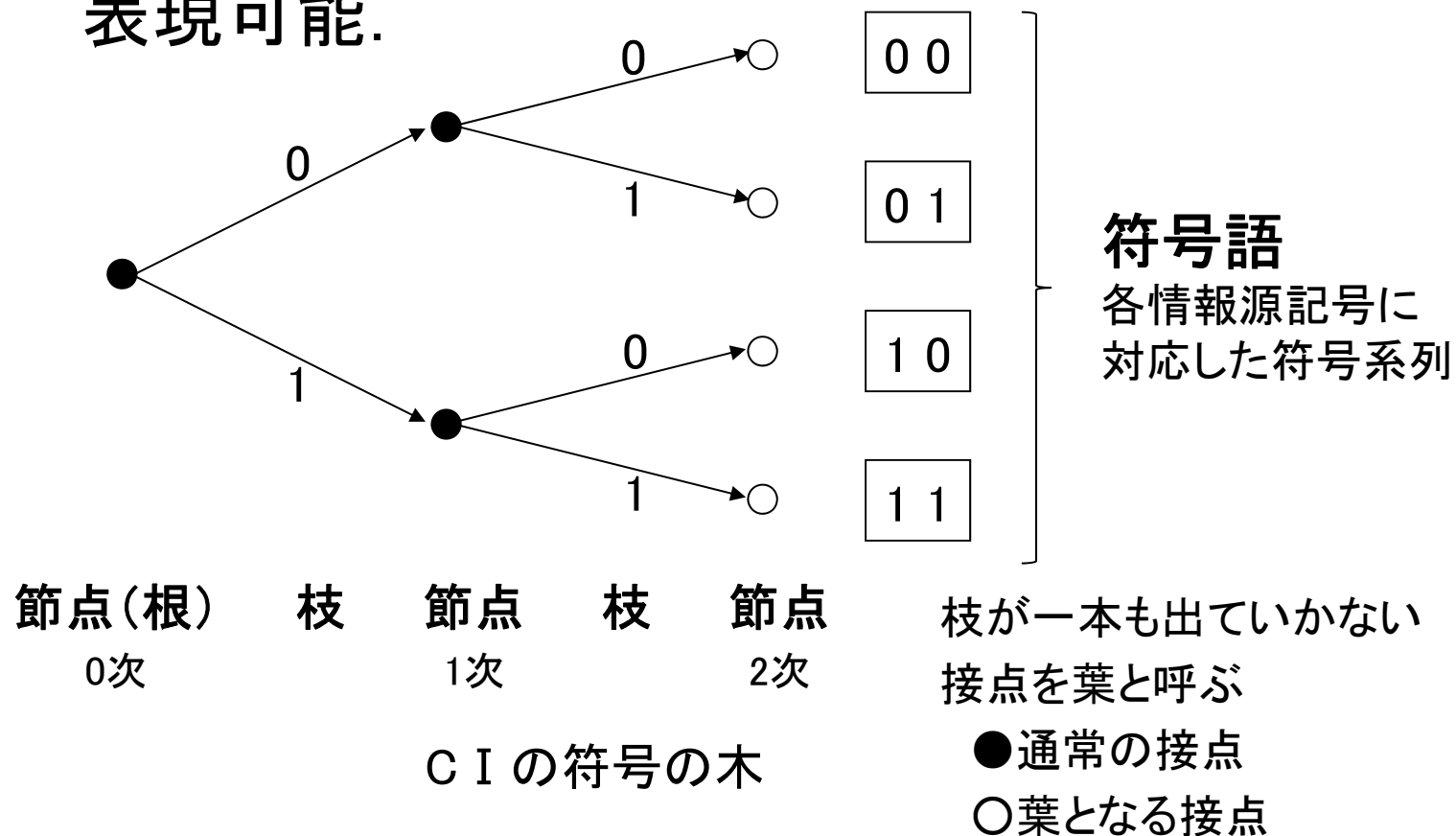
表 4.1 情報源符号化の例

情 報 源 号	確 率	C_1	C_2	C_3	C_4	C_5	C_6
A	0.6	00	0	0	0	0	0
B	0.25	01	10	10	01	10	10
C	0.1	10	110	110	011	11	11
D	0.05	11	1110	111	111	01	0
平 均 符 号 長		2.00	1.60	1.55	1.55	1.40	1.35

←一意復号可能 (瞬時) → ←一意復号可能 (非瞬時) → ←一意復号不可能 →

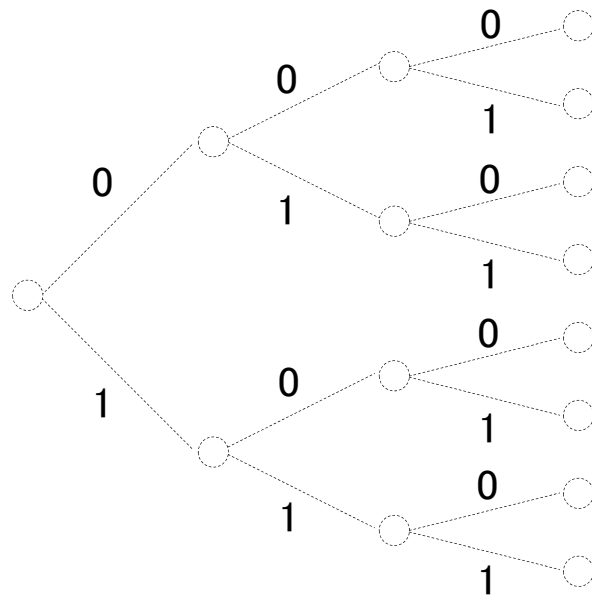
符号の木

- 各情報源記号に対応した符号系列を樹形図で表現. どのような符号語も枝を伸ばしていけば表現可能.



【演習1】符号の木

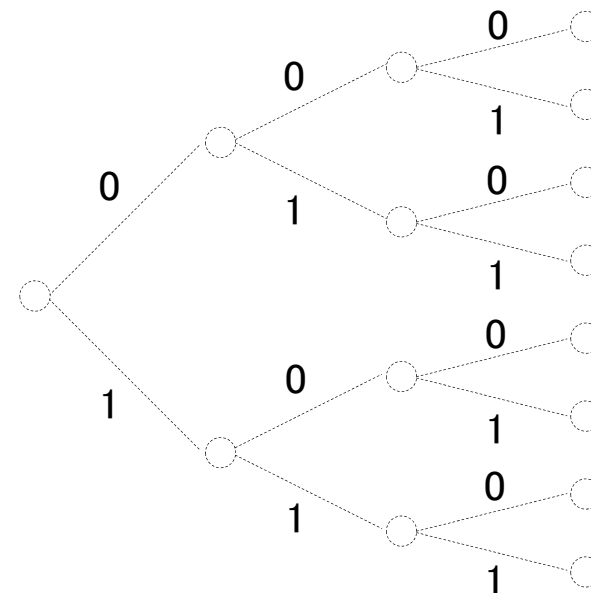
- 符号則CⅢ, CIVに対応する符号の木を書く.
- どの接点が符号語に対応するか, 接点や枝として残るのは点線のうちどこか書いてみよう.



CⅢの符号の木

符号語

0	1 0	1 1 0	1 1 1
---	-----	-------	-------



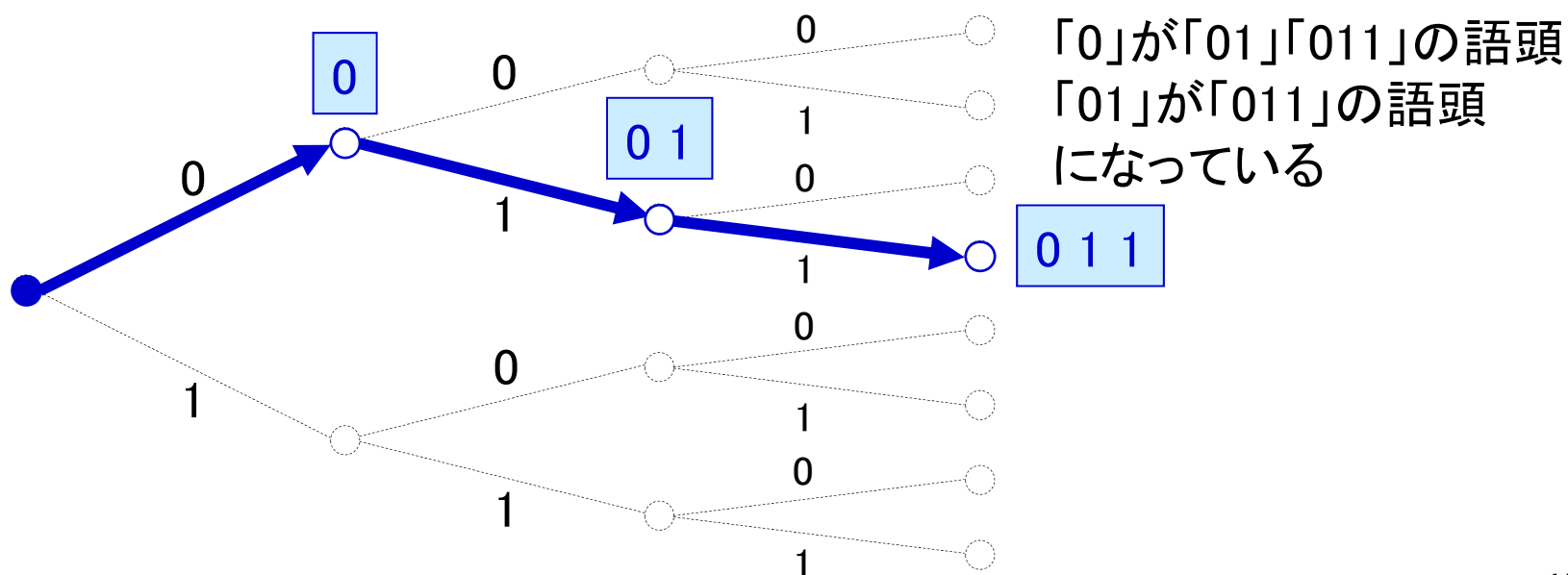
CIVの符号の木

符号語

0	0 1	0 1 1	1 1 1
---	-----	-------	-------

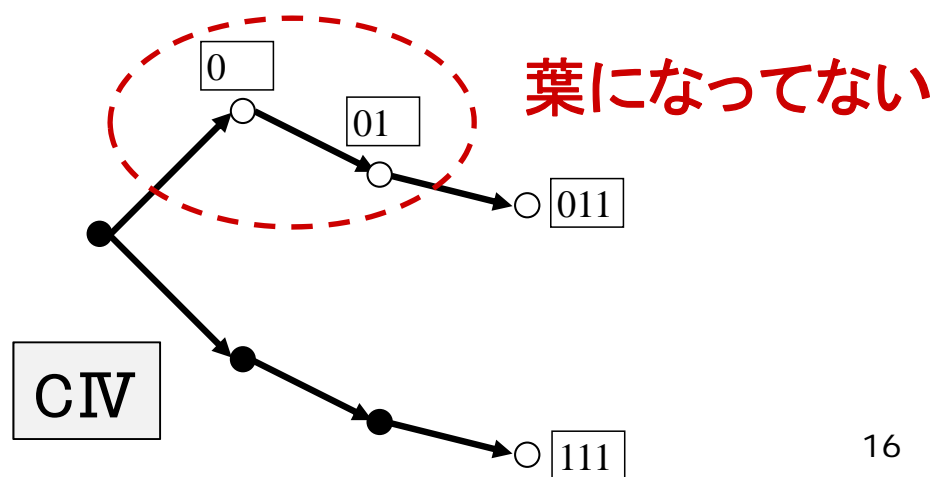
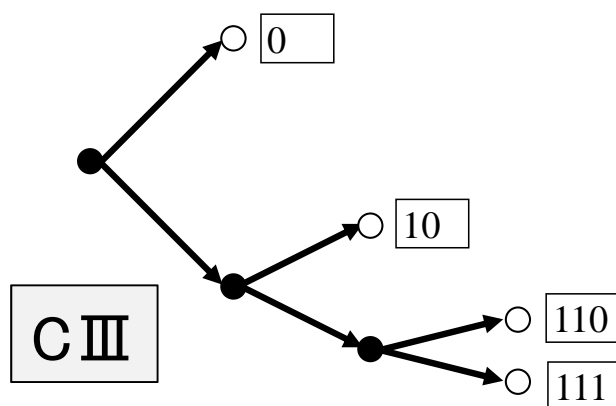
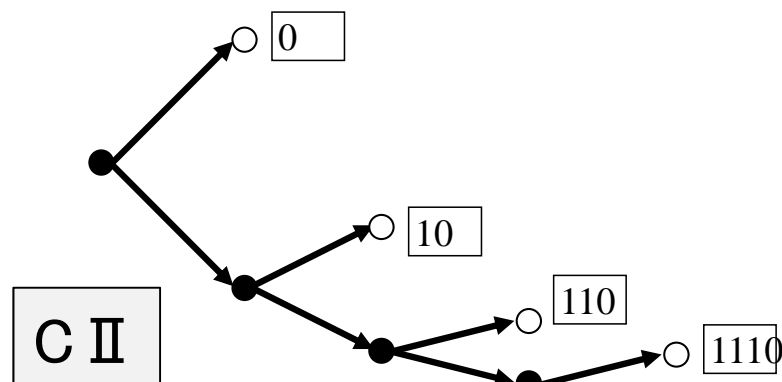
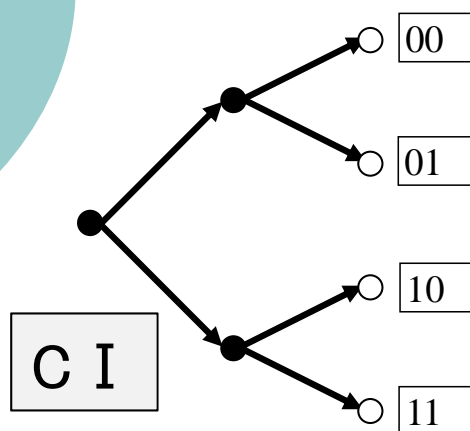
符号の木と瞬時符号の条件

- CIVはなぜ瞬時符号ではないか.
- 符号「0」を受信した時, 「0」か, 「01」の始まりか, 「011」の始まりか, 判別できない.
- 瞬時符号では, ある符号語が他の符号語の語頭になってはならない(語頭条件)



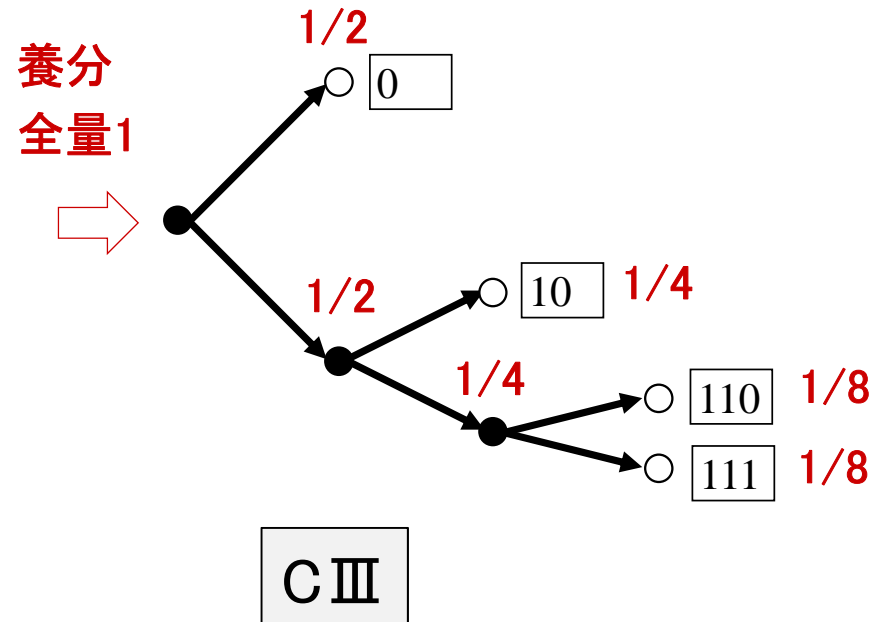
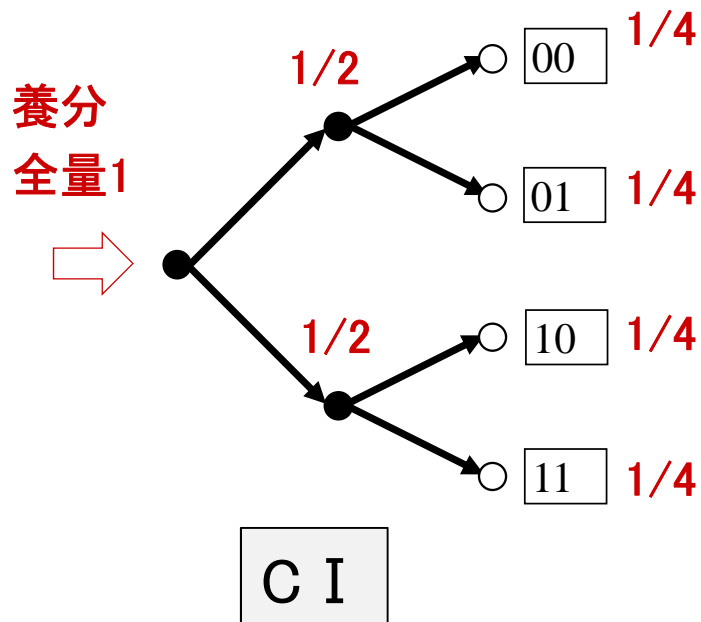
符号の木と瞬時符号の条件

- ある符号語が他の符号語の語頭になってはならない
⇒全ての符号語が「葉」に対応しないといけない



クラフトの不等式

- 符号の木の幹から全量1の養分を流し込んだ時に, 符号語の「葉」にどれだけ行きわたるか考える.
- 「葉」に行きわたった養分の総和は1を超えないはず.



クラフトの不等式

- これを一般化して、長さが l_1, l_2, \dots, l_M からなる M 個の符号語を持つ符号が瞬時符号となる条件を考える.
- それぞれの符号語が「葉」に対応したとき、長さ l_i の符号語に行きわたる養分は 2^{-l_i} . 養分の総和は1を超えないので、以下が成り立つ.
- 符号語の数 (M) が大きいとき、 l_1, l_2, \dots, l_M はあまり小さい値にできない.

$$2^{-l_1} + 2^{-l_2} + \dots + 2^{-l_M} \leq 1 \quad (\text{式4.2})$$

【演習2】以前の公務員試験問題

【No. 36】 表は、情報源アルファベット $S_1 \sim S_4$ を、4通りの方法で符号㉠～符号㉤に符号化したものであり、各符号は四つのビット列(符号語)で構成されている。符号㉠～符号㉤のうちから、瞬時に復号可能な符号のみをすべて選び出しているのはどれか。

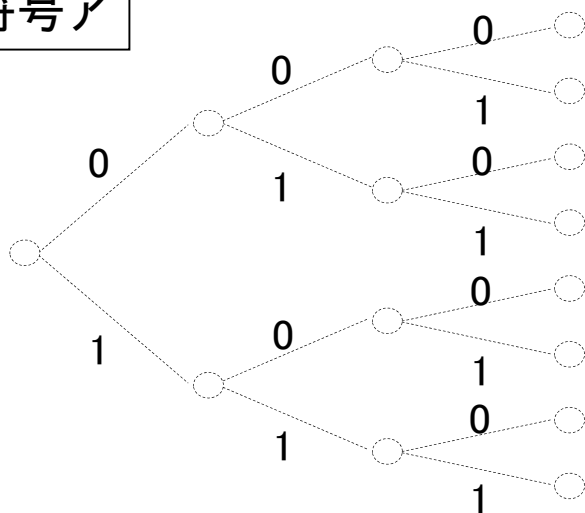
ただし、瞬時に復号可能な符号とは、一つのビット列を受け取れば直ちに情報源アルファベットに復号できるものをいう。

情報源アルファベット	符号㉠	符号㉡	符号㉢	符号㉤
S_1	0	0	0	00
S_2	00	01	10	01
S_3	01	011	110	10
S_4	11	0111	111	11

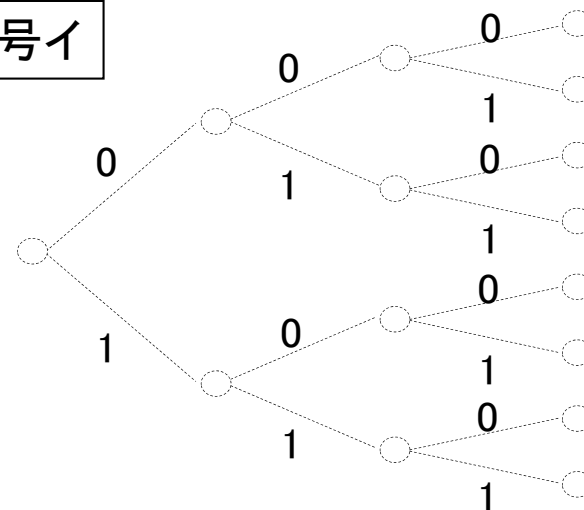
1. 符号㉠, 符号㉡
2. 符号㉠, 符号㉤
3. 符号㉡, 符号㉢
4. 符号㉡, 符号㉤
5. 符号㉢, 符号㉤

【演習2】以前の公務員試験問題

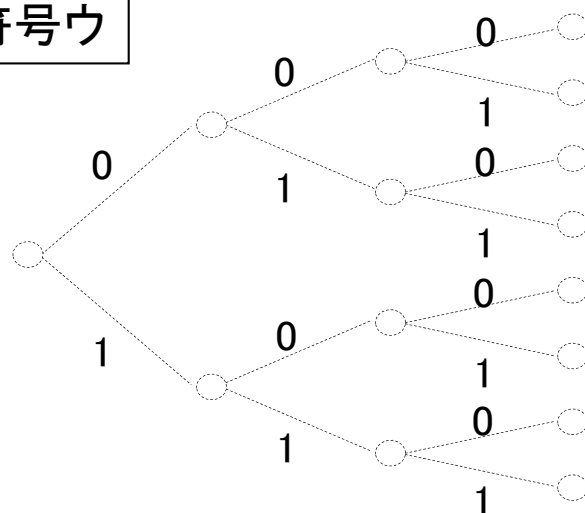
符号ア



符号イ



符号ウ



符号エ

