

準瞬時FV符号とその拡張符号

AIFV Codes

(Almost Instantaneous Fixed-to-Variable Length Codes)

and

Their Extended Codes

山本博資 (東京大学)

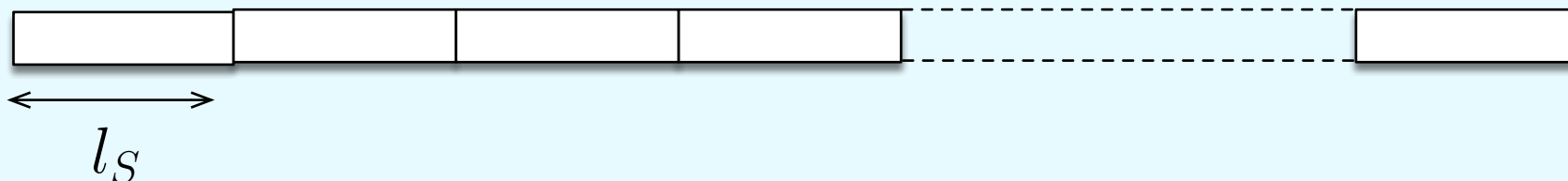


Outline

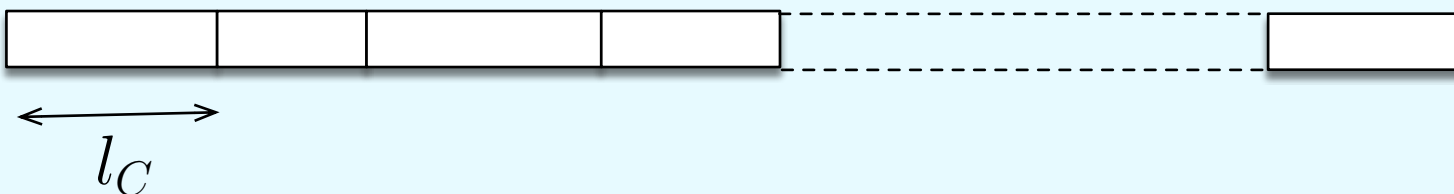
1. 背景 (瞬時符号, 一意復号可能符号, 準瞬時符号の関係)
2. 2元準瞬時FV符号 (符号木, 符号化, 復号化, 平均符号長, 最悪冗長度)
3. 2元準瞬時FV符号木の総数と符号木の符号化
4. 最適な2元準瞬時FV符号木の導出
5. 動的ハフマン符号と動的準瞬時FV符号
6. アルファベティック準瞬時FV符号
7. 2元準瞬時FV- m 符号
8. K 元準瞬時FV符号

FV(Fixed-to-Variable length)符号

情報源系列 (i.i.d., 有限離散アルファベット \mathcal{X})



符号語系列



| | FV符号 | VF符号 |
|-------|----------|----------|
| l_S | Fixed | Variable |
| l_C | Variable | Fixed |



FV符号

例.1

情報源アルファベット

$$\mathcal{X} = \{a, b, c, d, e\}$$

符号アルファベット $\mathcal{Y} = \{0, 1\}$

情報源確率分布

$$P_X(a) = 0.35, \quad P_X(b) = 0.1,$$

$$P_X(c) = 0.3 \quad P_X(d) = 0.15, \quad P_X(e) = 0.1$$

| x | codeword |
|-----|----------|
| a | 00 |
| b | 110 |
| c | 01 |
| d | 10 |
| e | 111 |

情報源系列: a c e d c a b a

符号語系列: 00 10 111 01 10 00 110 00

符号木(Code Tree)

例.1

情報源アルファベット

$$\mathcal{X} = \{a, b, c, d, e\}$$

符号アルファベット $\mathcal{Y} = \{0, 1\}$

情報源の確率分布

$$P_X(a) = 0.35, \quad P_X(b) = 0.1,$$

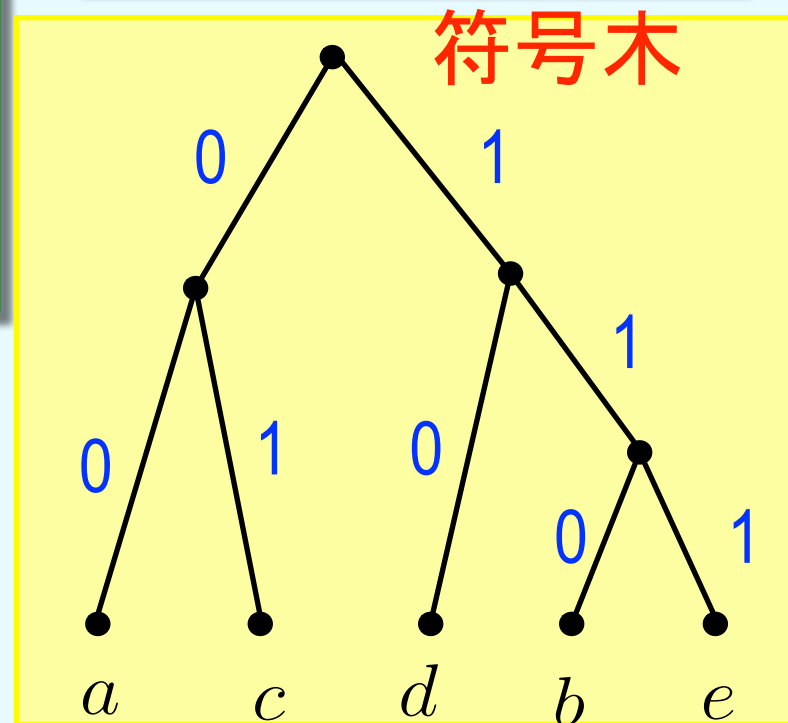
$$P_X(c) = 0.3 \quad P_X(d) = 0.15, \quad P_X(e) = 0.1$$

瞬時符号(Instantaneous code)

語頭符号

(Prefix code, Prefix-free code)

| x | codeword |
|-----|----------|
| a | 00 |
| b | 110 |
| c | 01 |
| d | 10 |
| e | 111 |





ハフマン符号, クラフトの不等式, マクミランの定理

ハフマン符号が最適
(Huffman, 1952)

クラフトの不等式

$$\sum_{x \in \mathcal{X}} 2^{-l(x)} \leq 1$$

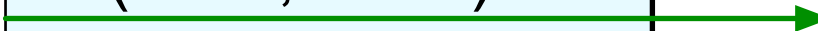
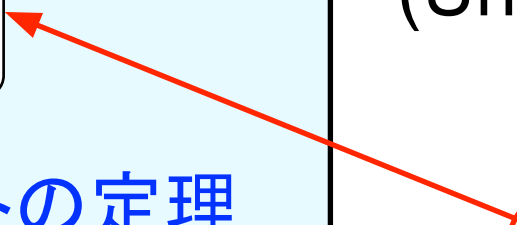
クラフトの定理
(Kraft, 1949)

マクミランの定理
(McMillan, 1956)

$l(x)$: 情報源シンボル x の符号長

一意復号可能符号
(Uniquely decodable
codes)

瞬時符号
(語頭符号)





ハフマン符号, クラフトの不等式, マクミランの定理

ハフマン符号が最適
(Huffman, 1952)

クラフトの不等式

$$\sum_{x \in \mathcal{X}} 2^{-l(x)} \leq 1$$

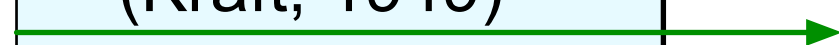
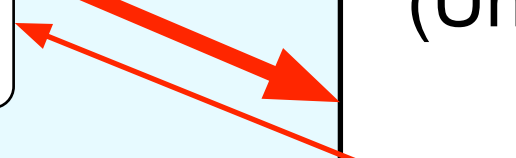
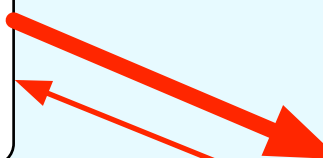
クラフトの定理
(Kraft, 1949)

マクミランの定理
(McMillan, 1956)

$l(x)$: 情報源シンボル x の符号長

一意復号可能符号
(Uniquely decodable
codes)

瞬時符号
(語頭符号)





ハフマン符号, クラフトの不等式, マクミランの定理

ハフマン符号が最適
(Huffman, 1952)

クラフトの不等式

$$\sum_{x \in \mathcal{X}} 2^{-l(x)} \leq 1$$

クラフトの定理
(Kraft, 1949)

マクミランの定理
(McMillan, 1956)

単一符号木の場合だけを取り扱っている.

一意復号可能符号
(Uniquely decodable codes)

瞬時符号
(語頭符号)

複数の符号木: ハフマン符号より圧縮率のよい一意復号可能符号が存在する可能性がある.

FV符号の階層

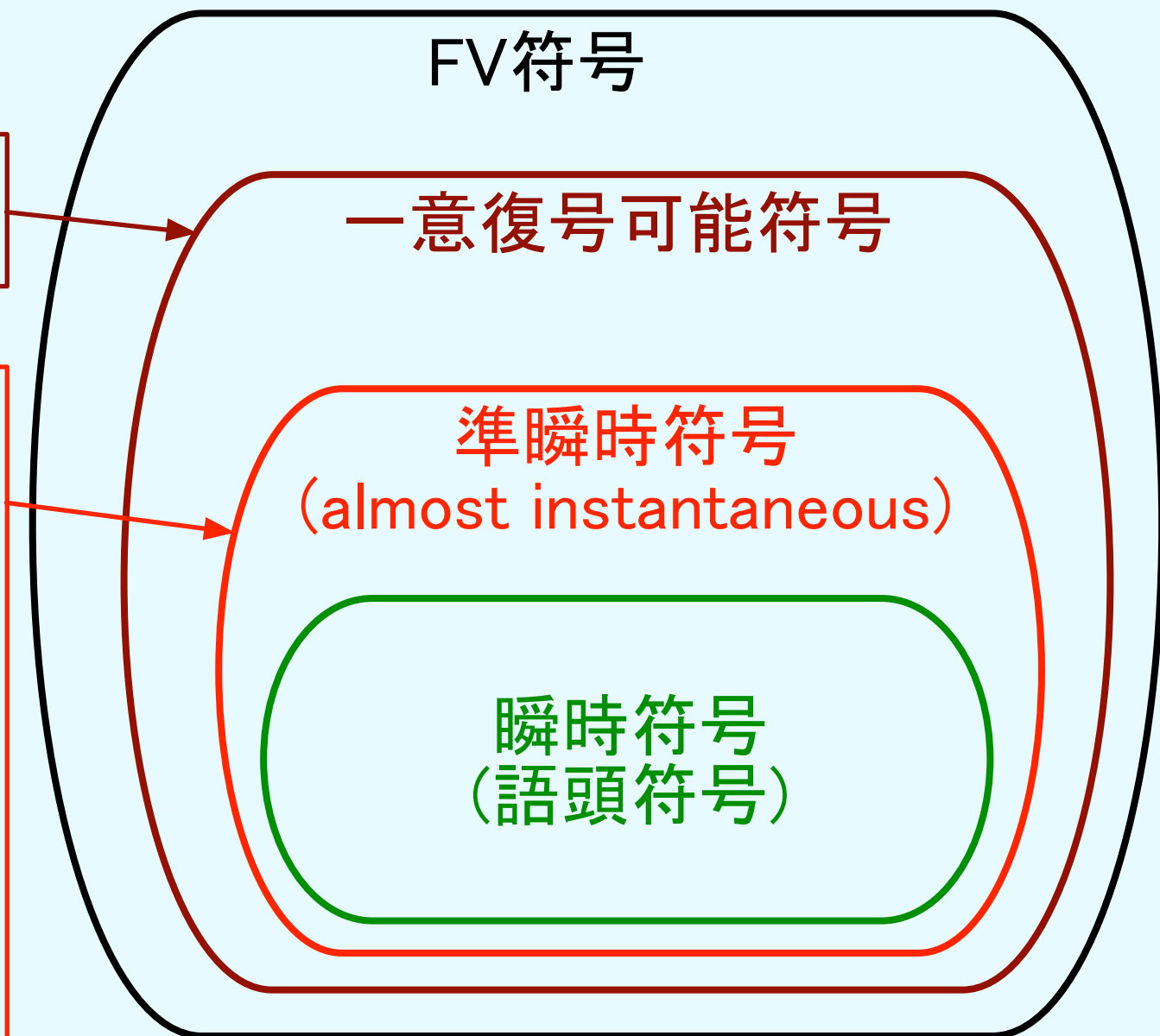
復号遅延が非常に大きくなる可能性がある

復号遅延:

2元準瞬時FV符号
最大2ビット

2元準瞬時FV- m 符号
最大 m ビット

K 元準瞬時FV符号
最大1符号シンボル

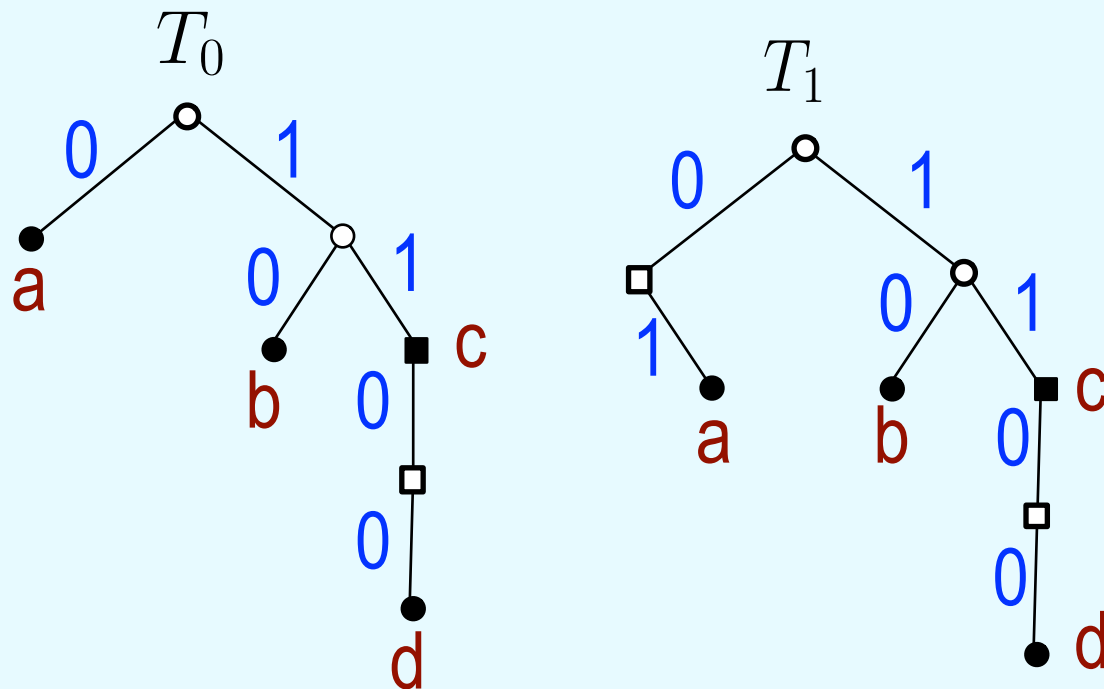


2元準瞬時FV符号

(Yamamoto, Wei, 2013)

(Yamamoto, Tsuchihashi, Honda, 2015)

例.2 情報源アルファベット: $\mathcal{X} = \{a, b, c, d\}$
符号アルファベット: $\mathcal{Y} = \{0, 1\}$



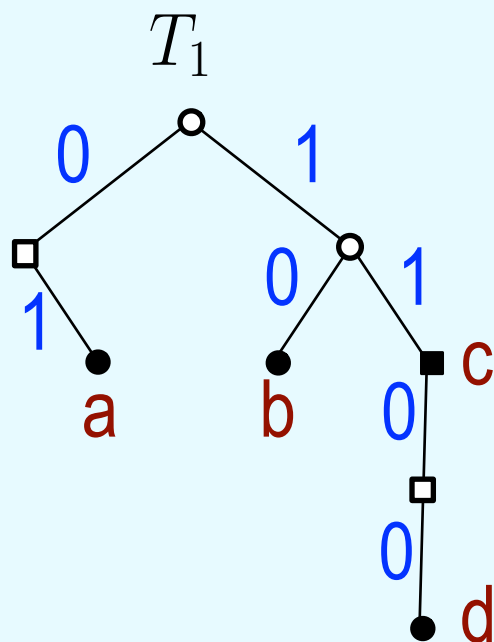
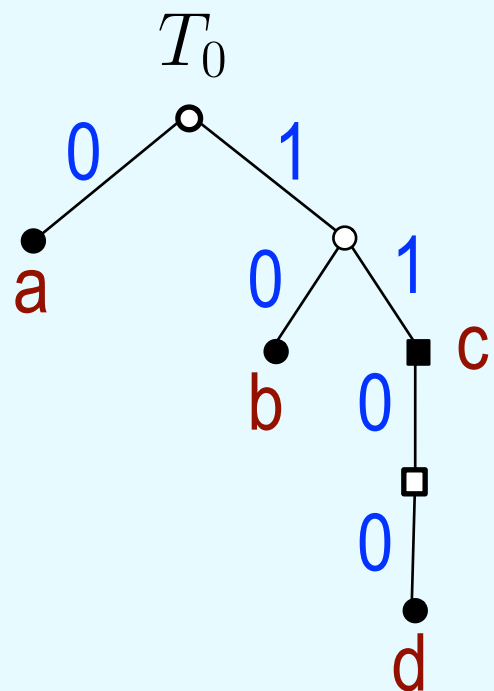
- 葉
- 完全内部節点
- 不完全内部節点
 - マスター節点 (Master node)
 - スレーブ節点 (Slave node)

2元準瞬時符号の符号化

情報源系列: c b d c a

11

T_0 : 初期符号木



符号木の遷移則

葉 $\rightarrow T_0$

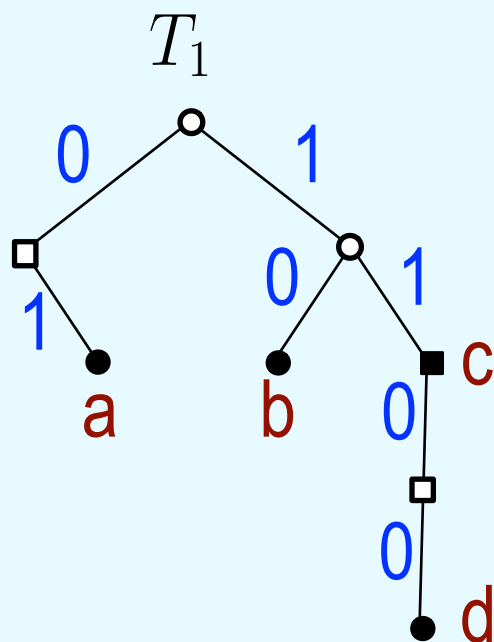
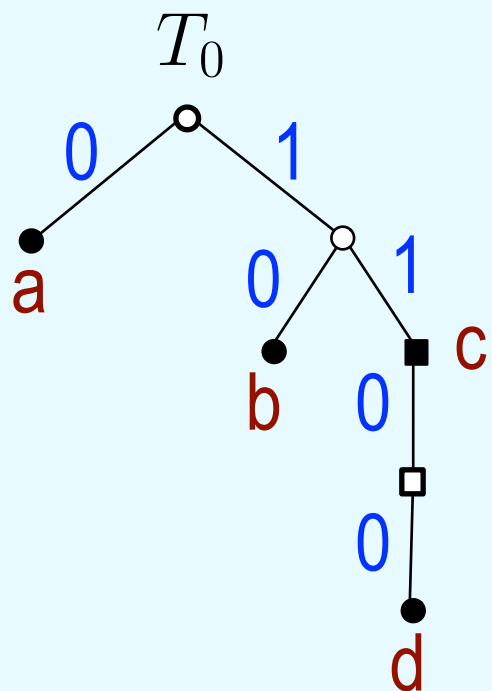
マスター
節点 $\rightarrow T_1$

2元準瞬時符号の符号化

情報源系列: c b d c a

符号語系列: 11 10 1100 11 01

T_0 T_1 T_0 T_0 T_1



符号木の遷移則

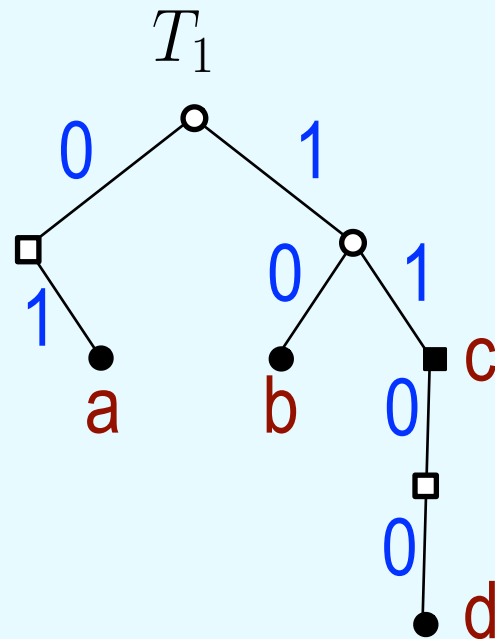
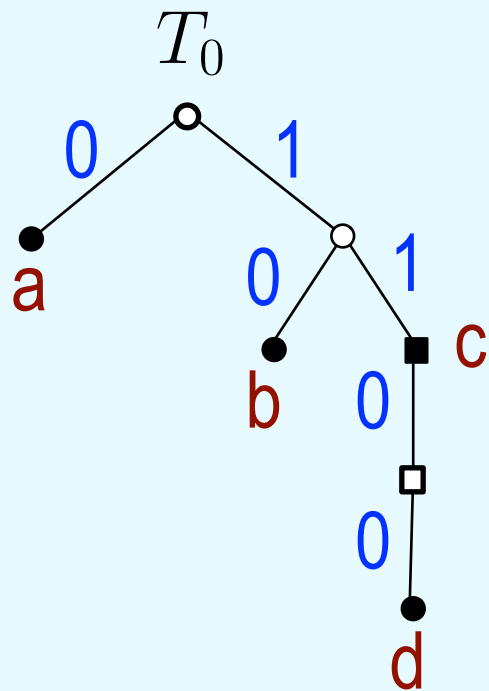
葉 $\rightarrow T_0$

マスター
節点 $\rightarrow T_1$

2元準瞬時符号の復号化

符号語系列: 111011001101

情報源系列: c



T_0 : 初期符号木

符号木の遷移則

葉 $\rightarrow T_0$

マスター
節点 $\rightarrow T_1$

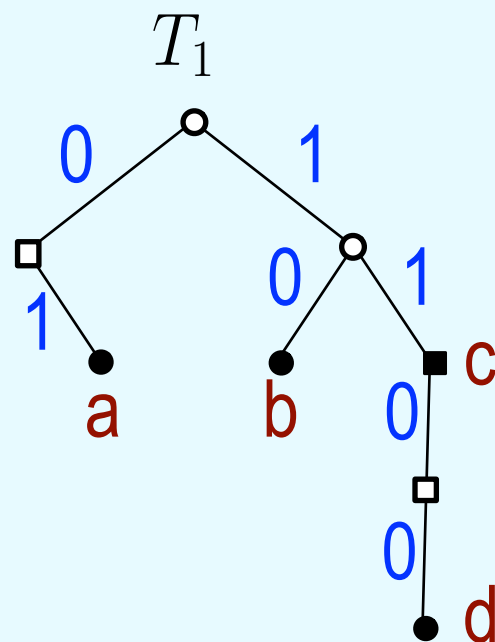
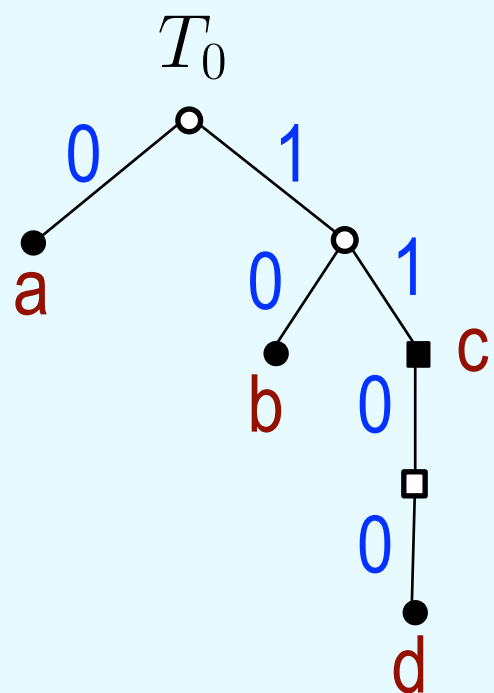
2元準瞬時符号の復号化

符号語系列: 111011001101

情報源系列: c b d c a

T_0 T_1 T_0 T_0 T_1

T_0 : 初期符号木



符号木の遷移則

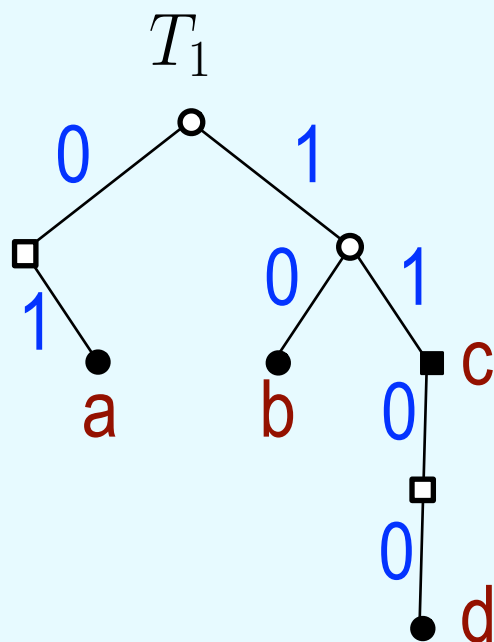
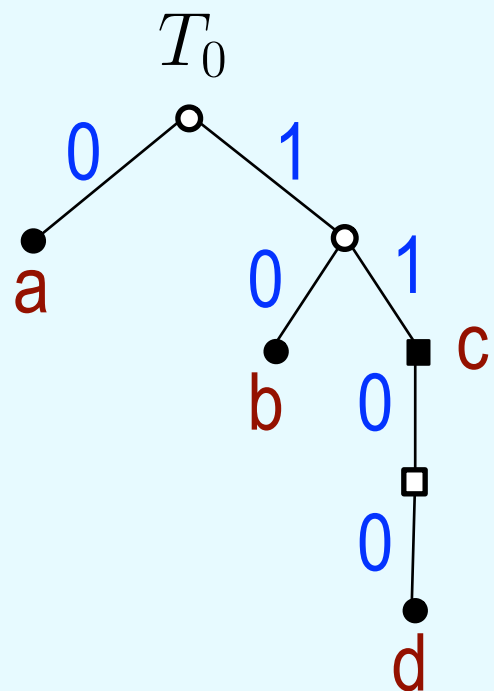
葉 $\rightarrow T_0$

マスター
節点 $\rightarrow T_1$

2元準瞬時符号の復号化

符号語系列: 111011001101

情報源系列: c b d c a



最大復号遅延2ビットで一意復号可能

平均符号長

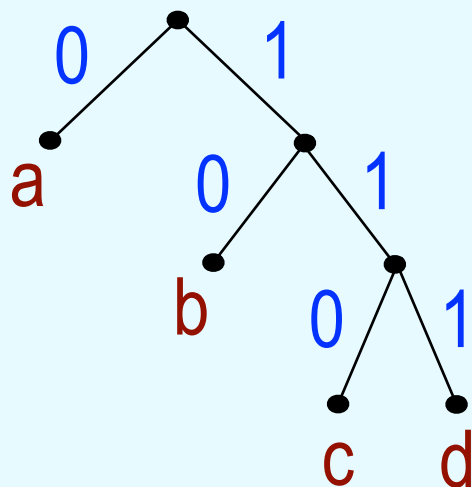
例.3 情報源アルファベット: $\mathcal{X} = \{a, b, c, d\}$

符号語アルファベット: $\mathcal{Y} = \{0, 1\}$

$P(a) = 0.45, P(b) = 0.3, P(c) = 0.2, P(d) = 0.05$

エントロピー: $H(X) = -\sum_{x \in \mathcal{X}} P(x) \log_2 P(x) \approx 1.7200$

ハフマン符号木



平均符号長

$$L_H = 1.8$$

平均符号長

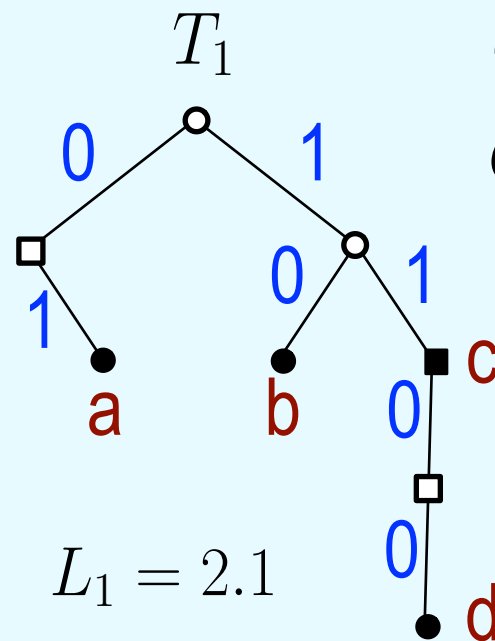
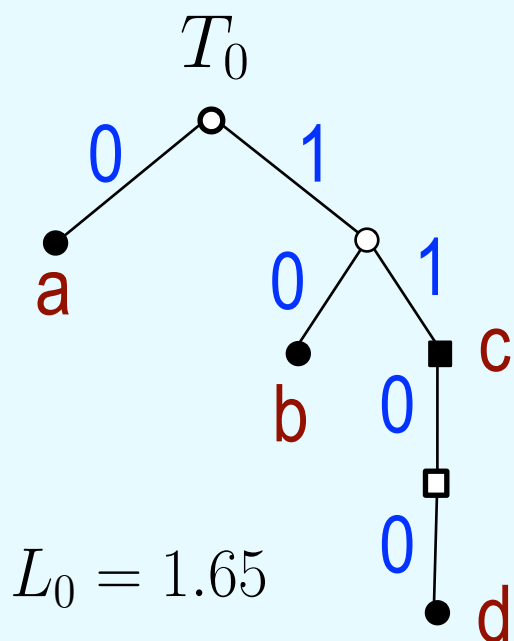
例.3 情報源アルファベット: $\mathcal{X} = \{a, b, c, d\}$

符号語アルファベット: $\mathcal{Y} = \{0, 1\}$

$P(a) = 0.45, P(b) = 0.3, P(c) = 0.2, P(d) = 0.05$

エントロピー: $H(X) = -\sum_{x \in \mathcal{X}} P(x) \log_2 P(x) \approx 1.7200$

ハフマン符号:
 $L_H = 1.8$



$$Q(T_1|T_0) = P(c) = 0.2$$

$$Q(T_0|T_1) = P(a) + P(b) + P(d) = 0.8$$

$$Q(T_0) = 0.8 \quad Q(T_1) = 0.2$$

$$\begin{aligned}
 L_{AIFV} &= Q(T_0)L_0 + Q(T_1)L_1 \\
 &= 0.8 \times 1.65 + 0.2 \times 2.1 \\
 &= 1.74
 \end{aligned}$$

平均符号長

例.4 情報源アルファベット: $\mathcal{X} = \{a, b, c\}$

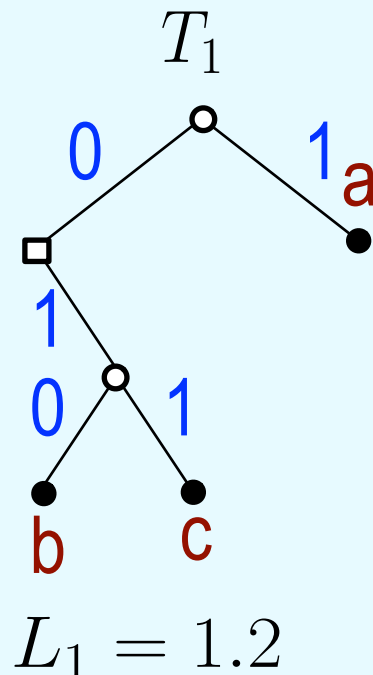
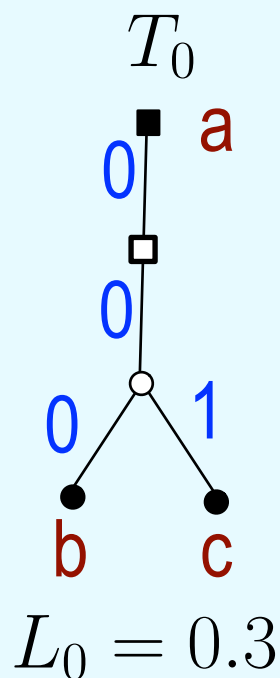
符号語アルファベット: $\mathcal{Y} = \{0, 1\}$

$P(a) = 0.9, P(b) = P(c) = 0.05$

エントロピー: $H(X) \approx 0.5690$

ハフマン符号:

$$L_H = 1.1$$



$$Q(T_1|T_0) = 0.9, Q(T_0|T_1) = 1$$

$$Q(T_0) = \frac{10}{19}, Q(T_1) = \frac{9}{19}$$

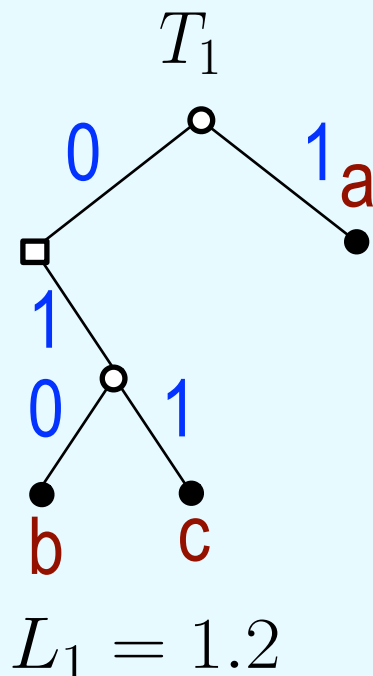
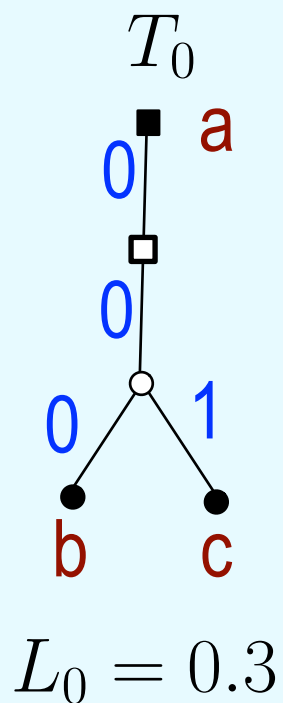
$$L_{AIFV} = Q(T_0)L_0 + Q(T_1)L_1 \approx 0.7263$$



空符号語 (Null codeword)

情報源系列: a a a b

符号語系列: λ 1 λ 010 (λ : null)
1010



$$Q(T_1|T_0) = 0.9, Q(T_0|T_1) = 1$$

$$Q(T_0) = \frac{10}{19}, Q(T_1) = \frac{9}{19}$$

$$L_{AIFV} = Q(T_0)L_0 + Q(T_1)L_1$$

$$\approx 0.7263$$



符号語長の冗長度(Redundancy)

$$\text{冗長度} = L - H(X)$$

$$p_{\max} = \max_{x \in \mathcal{X}} P(x)$$

ハフマン符号に対する冗長度の上界 (Ye, Young, 2002)

$$p_{\max} \geq \frac{1}{6} \approx 0.1667$$

のときタイト

IEEE Trans. Inform. Theory, vol.63,
no.8, p.5076, Aug. 2017のFig.3を
参照せよ

$$H(X) \leq L_H < H(X) + 1$$

$$h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$$

: 2元エントロピー関数



符号語長の冗長度(Redundancy)

準瞬時FV符号の冗長度の上界 (Hu, Yamamoto, Honda, 2017)

$p_{\max} \geq 0.5$ のときタイト

IEEE Trans. Inform. Theory, vol.63,
no.8, p.5079, Aug. 2017のFig.7を
参照せよ.

$$f(x) = \begin{cases} x^2 - 2x + 2 - h(x) & \text{if } \frac{1}{2} \leq x \leq \frac{-1+\sqrt{5}}{2}, \\ \frac{-2x^2+x+2}{1+x} - h(x) & \text{if } \frac{-1+\sqrt{5}}{2} \leq x < 1. \end{cases}$$

$p_{\max} \leq 0.5$ のとき

冗長度 ≤ 0.25

$$H(X) \leq L_{AIFV} < H(X) + 0.5$$



準瞬時FV符号とハフマン符号の比較 ($n = |\mathcal{X}|$)

| | ハフマン符号 | \mathcal{X}^2 に対するハフマン符号 | 準瞬時FV符号 |
|--------|--------|--|---------|
| 最悪冗長度 | 1 | 0.5 | 0.5 |
| 符号木サイズ | $O(n)$ | $O(n^2)$ | $O(2n)$ |
| 復号遅延 | 0 | $a_1 a_2 \in \mathcal{X}^2$ a_1 に対して大 | 最大2ビット |



数値実験例

$$\mathcal{X} = \{a_1, a_2, \dots, a_n\}$$

IEEE Trans. Inform. Theory, vol.61, no.12, p.6442, Dec. 2015のFig.12を参照せよ.

IEEE Trans. Inform. Theory, vol.61, no.12, p.6442, Dec. 2015のFig.13を参照せよ.

$$P_1(a_t) = \frac{t}{A_1}$$

$$P_2(a_t) = \frac{t^2}{A_2}$$

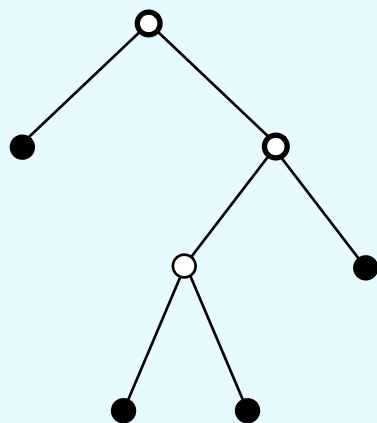
符号木の総数

ハフマン符号木の総数 ($n = |\mathcal{X}|$)

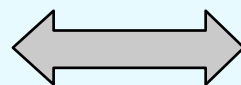
$$N_{H,n} = C_{n-1}$$

$$C_n \equiv \frac{1}{n+1} \binom{2n}{n} \text{ :Catalan number}$$

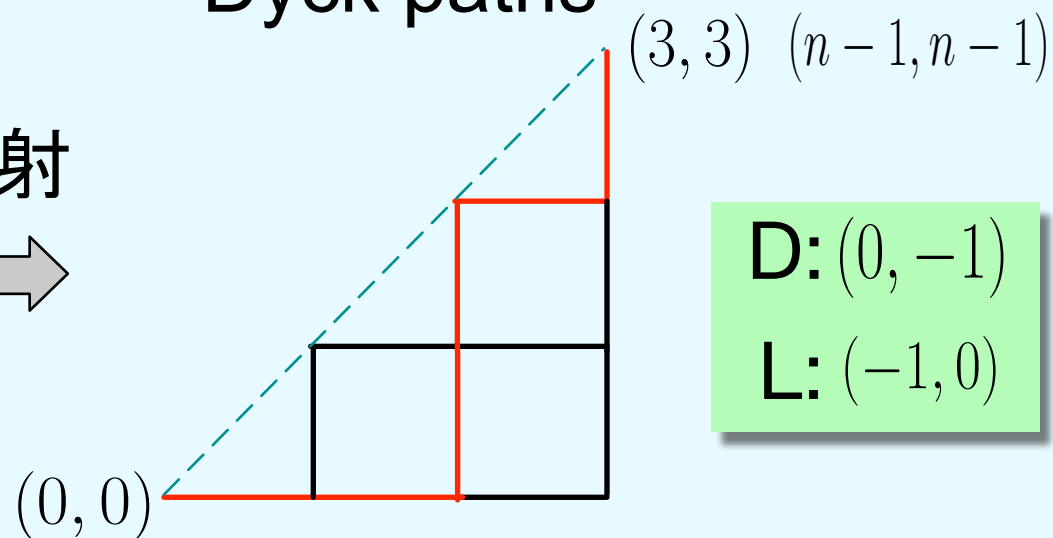
Binary full trees



全単射



Dyck paths



D: (0, -1)

L: (-1, 0)

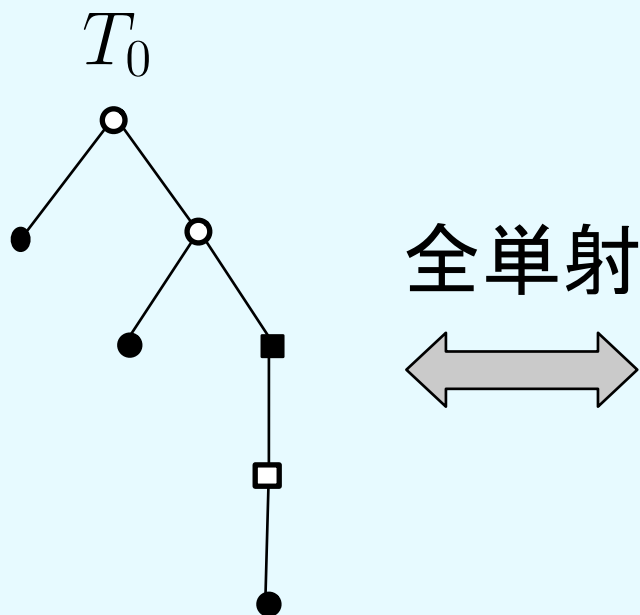
符号木の総数

準瞬時FV符号木の総数 ($n = |\mathcal{X}|$)

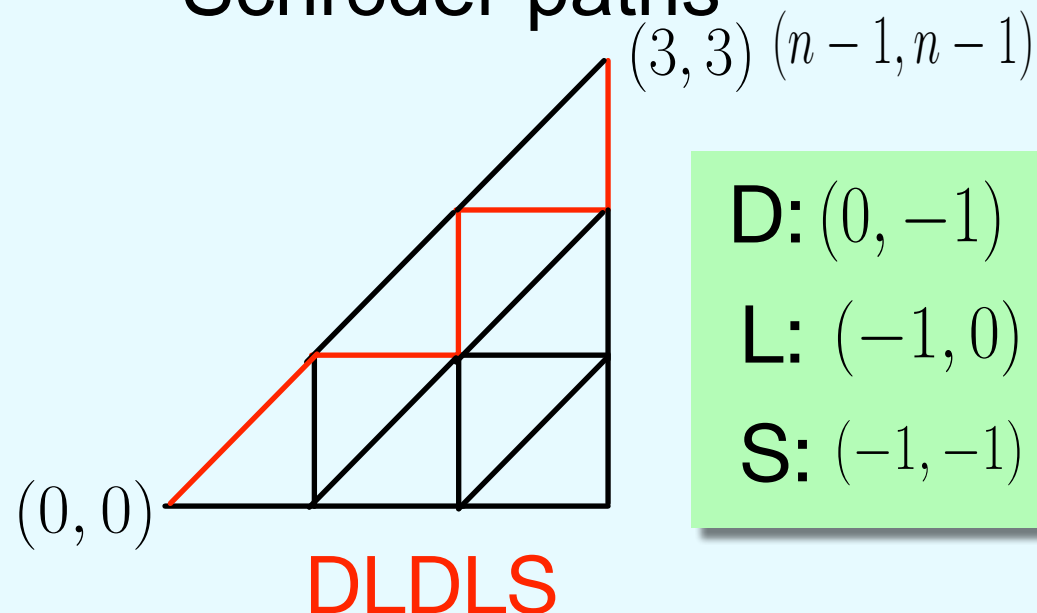
(Sumigawa, Yamamoto, 2017)

$$N_{T_0,n} = S_{n-1}, \quad N_{T_1,n} = S_{n-1} - S_{n-2}$$

$$S_n \equiv \sum_{k=0}^n \frac{1}{n-k+1} \binom{2n-2k}{n-k} \binom{2n-k}{k} : \text{Large Schröder number}$$



Schröder paths



D: $(0, -1)$

L: $(-1, 0)$

S: $(-1, -1)$



符号木の総数

$$n = |\mathcal{X}|$$

ハフマン符号木

$$N_{H,n} = C_{n-1}$$

C_n : Catalan number

$$\log_2 C_n = 2n + o(n)$$

準瞬時FV符号木

$$N_{T_0,n} = S_{n-1}$$

S_n : Large Schröder number

$$\begin{aligned} \log_2 S_n &= n \log_2 \left(3 + 2\sqrt{2} \right) + o(n) \\ &\approx 2.5431n + o(n) \end{aligned}$$

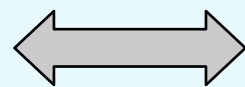
準瞬時FV符号木の符号化

(Sumigawa, Yamamoto,
2017)

$$n = |\mathcal{X}|$$

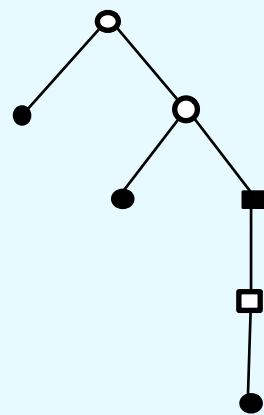
全単射

AIFV code tree



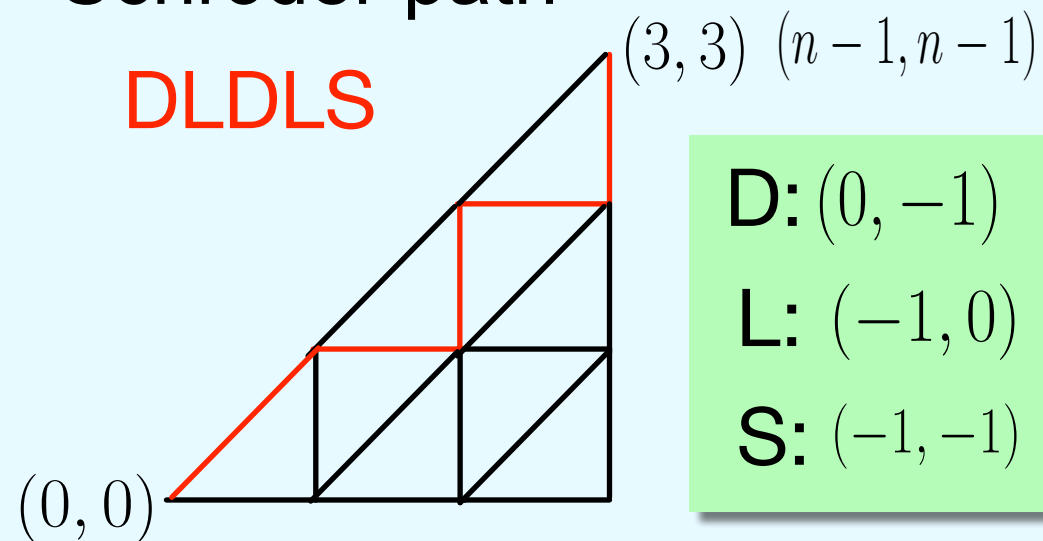
Schröder path

T_0



$O(n)$

DLDSL



D: $(0, -1)$

L: $(-1, 0)$

S: $(-1, -1)$

$\lceil \log_2 S_{n-1} \rceil$ bits

時間複雑度: $O(n^2)$ $O(n)$

メモリ量 $O(n^2)$ で事前計算

Total order: $L \prec S \prec D$

全てのSchröderパスに全順序
を与える。
対応するパスの順序で符号化



最適な準瞬時FV符号木の構成

ハフマン符号木

ハフマンアルゴリズム: シンプル

準瞬時FV符号木

反復最適化+ IP (整数計画法)

(Yamamoto, Tsuchihashi, Honda, 2015)

反復最適化+ DP (動的計画法)

(Iwata, Yamamoto, 2016)

Time complexity: $O(n^5)$, Space complexity: $O(n^3)$

$n = |\mathcal{X}|$

準瞬時FV符号木の反復最適化手法

(Yamamoto, Tsuchihashi, Honda, 2015)

平均符号長

$$\begin{aligned}L_{AIFV} &= Q(T_0)L_0 + Q(T_1)L_1 \\ &= \frac{Q_{0|1}L_0 + Q_{1|0}L_1}{Q_{0|1} + Q_{1|0}} \\ &= L_0 + CQ_{1|0} \\ &= L_1 - CQ_{0|1}\end{aligned}$$

$$Q_{0|1} \equiv Q(T_0|T_1)$$

$$Q_{1|0} \equiv Q(T_1|T_0)$$

$$C \equiv \frac{L_1 - L_0}{Q_{0|1} + Q_{1|0}}$$

: T_0 と T_1 から決まる.

$$Q(T_0) = \frac{Q_{0|1}}{Q_{0|1} + Q_{1|0}} \quad Q(T_1) = \frac{Q_{1|0}}{Q_{0|1} + Q_{1|0}}$$



準瞬時FV符号木の反復最適化手法

(Yamamoto, Tsuchihashi, Honda, 2015)

平均符号長

$$\begin{aligned} L_{AIFV} &= Q(T_0)L_0 + Q(T_1)L_1 \quad \leftarrow T_0 \text{ と } T_1 \\ &= \frac{Q_{0|1}L_0 + Q_{1|0}L_1}{Q_{0|1} + Q_{1|0}} \\ &= L_0 + CQ_{1|0} \quad \leftarrow T_0 \\ &= L_1 - CQ_{0|1} \quad \leftarrow T_1 \end{aligned}$$

$$C \equiv \frac{L_1 - L_0}{Q_{0|1} + Q_{1|0}}$$

準瞬時FV符号木の反復最適化手法

(Yamamoto, Tsuchihashi, Honda, 2015)

Step 1 : 初期化

$$C = C_{\text{init}} (= 2 - \log_2 3)$$

$$\begin{aligned} L_{AIFV} &= L_0 + CQ_{1|0} \\ &= L_1 - CQ_{0|1} \end{aligned}$$

Step 2 : 与えられた C に対して,

$$L_0 + CQ_{1|0}$$

を最小化する T_0 を求める.

$$L_1 - CQ_{0|1}$$

を最小化する T_1 を求める.

$$Q_{0|1} \equiv Q(T_0|T_1)$$

$$Q_{1|0} \equiv Q(T_1|T_0)$$

$$C \equiv \frac{L_1 - L_0}{Q_{0|1} + Q_{1|0}}$$

Step 3 : 得られた T_0 と T_1 から $Q_{0|1}$, $Q_{1|0}$, L_0 , L_1 , C を求める.

Step 4 : C が変化した場合, 新しい C を持って Step 2 に戻る.

C が変化する場合, L_{AIFV} は減少する.

C が変化しない場合, その時の T_0 と T_1 は最適.



準瞬時FV符号木の反復最適化手法

(Yamamoto, Tsuchihashi, Honda, 2015)

Step 1 : 初期化

$$C = C_{\text{init}} (= 2 - \log_2 3)$$

有限の反復で収束する.

Step 2 : 与えられた C に対して,

$$L_0 + CQ_{1|0}$$

を最小化する T_0 を求める.

$$L_1 - CQ_{0|1}$$

を最小化する T_1 を求める.

IP (Yamamoto, Tsuchihashi,
Honda, 2015)

DP (Iwata, Yamamoto, 2016)

A*アルゴリズム (山川, 西新,
2017)

Step 3 : 得られた T_0 と T_1 から $Q_{0|1}$, $Q_{1|0}$, L_0 , L_1 , C を求める.


Step 4 : C が変化した場合, 新しい C を持って Step 2 に戻る.

C が変化する場合, L_{AIFV} は減少する.

C が変化しない場合, その時の T_0 と T_1 は最適.

準瞬時FV符号木の簡易構成法

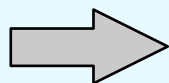
与えられた確率分布に対する最適な準瞬時FV符号木の構成

反復最適化 + IP, DP など  手間がかかる

必ずしも最適な準瞬時FV符号木でなくても構わないが、
ハフマン符号よりも圧縮率がよく、

$$H(X) \leq L_{AIFV} < H(X) + 0.5$$

を満たしたい。



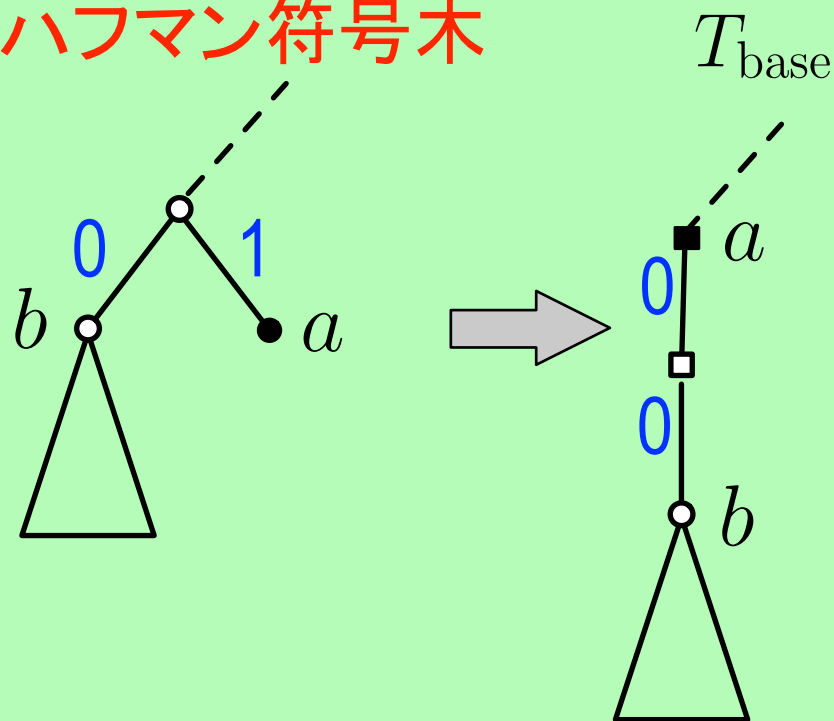
簡易構成法：

準瞬時FV符号木 T_0 と T_1 をハフマン符号木から作る。

ハフマン符号木を用いた準瞬時FV符号木の簡易構成法 (Hu, Yamamoto, Honda, 2017) (平岡, 山本, 2017)

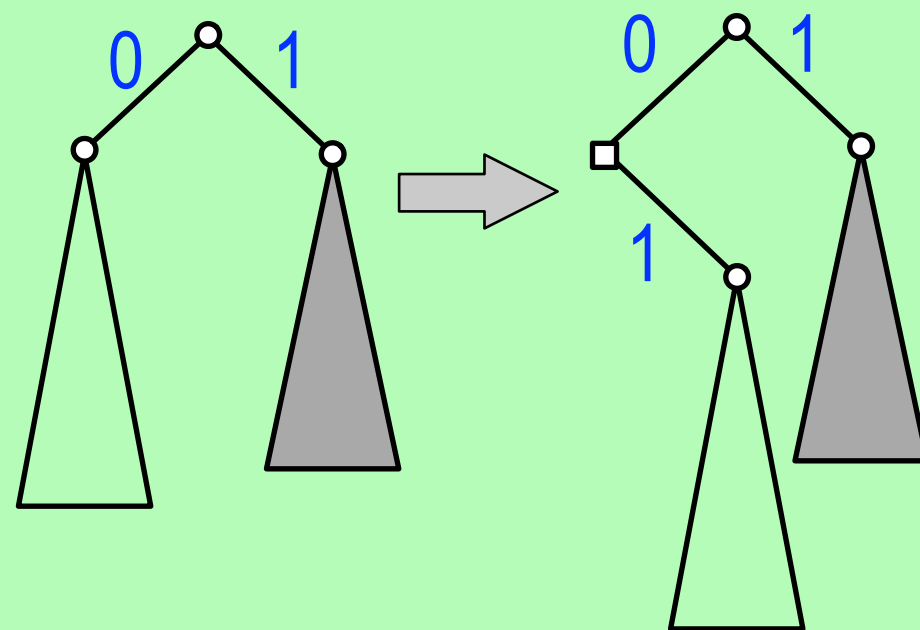
深さ $D \geq 2$ に対して,
 葉 a が $p_b < rp_a$ を満たすとき
 (r : 根の右の子の確率重み)

ハフマン符号木



$$\text{If } p_{\max} < \frac{-1 + \sqrt{5}}{2},$$

$$T_0 = T_{\text{base}} \quad T_1$$

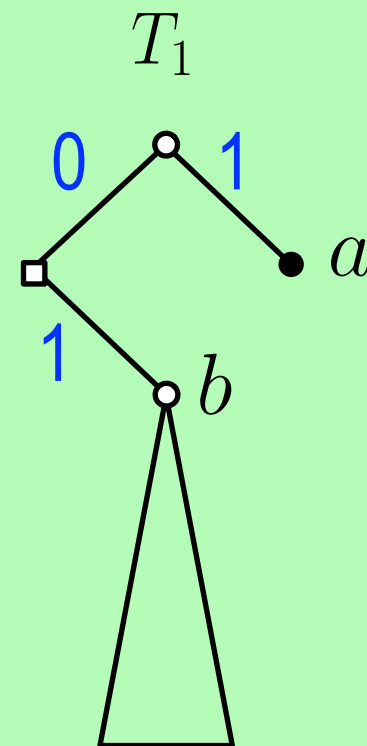
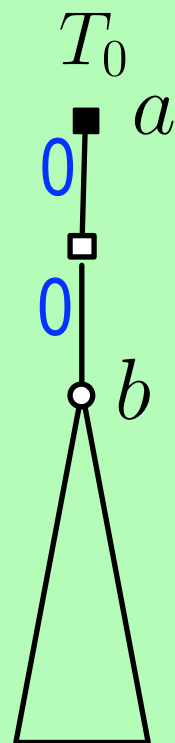
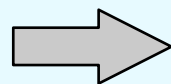
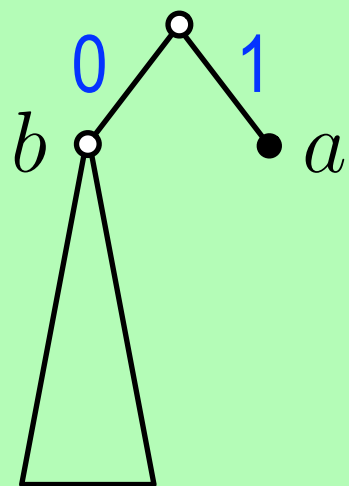


$$H(X) \leq L_{AIFV} < H(X) + 0.5$$

ハフマン符号木を用いた準瞬時FV符号木の簡易構成法 (Hu, Yamamoto, Honda, 2017) (平岡. 山本, 2017)

if $\frac{-1 + \sqrt{5}}{2} \leq p_{\max} < 1,$

T_{base}



$$H(X) \leq L_{AIFV} < H(X) + 0.5$$

動的ハフマン符号と動的準瞬時FV符号

静的ハフマン符号: 固定の符号木を使用

動的ハフマン符号:

情報源系列の1情報源シンボルを符号化すると共に、その時点の頻度分布に基づいて、符号木を更新する。

FGK algorithm (Faller, 1973, Gallager, 1978, Knuth, 1985)

Vitter algorithm (Vitter, 1989)

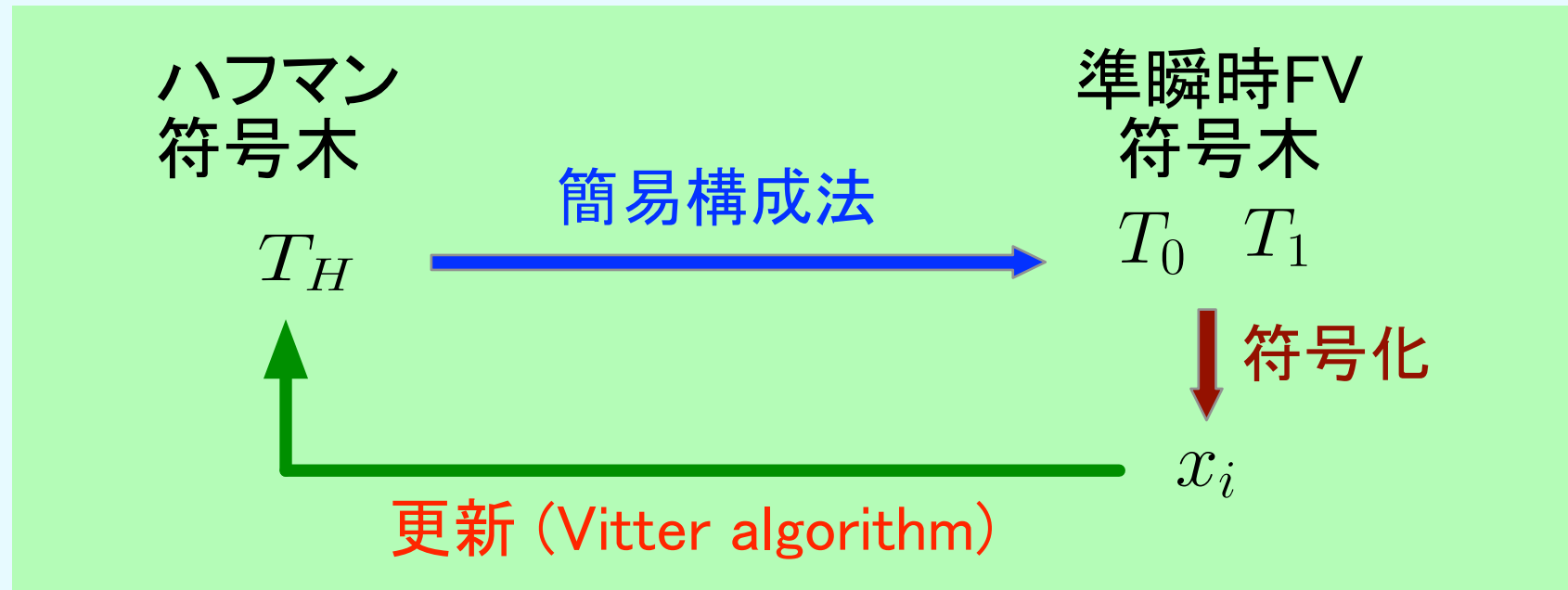
静的準瞬時FV符号: 最適な符号木の構成.
符号木の簡易構成法

動的準瞬時FV符号:

符号木の最適な更新法: Open problem
簡易構成法を利用した更新法

動的ハフマン木を利用した動的準瞬時FV符号化法

(平岡, 山本, 2017)



情報源系列: $x_1 x_2 \cdots x_i \cdots$

実際の x_i の符号化では, T_H の符号語の各兄弟節点の確率重みを比較して符号語を修正するだけで符号化可能.

➡ 符号化の複雑度は, 動的ハフマン符号とほぼ同じ



動的ハフマン木を利用した動的準瞬時FV符号化法

(平岡, 山本, 2017)

Canterbury Corpus

| File Name | Size (bytes) | Dynamic Huffman (Vitter) | Dynamic AIFV |
|--------------|--------------|--------------------------|--------------|
| alice29.txt | 152,089 | 4.618 | 4.618 |
| asyoulik.txt | 125,179 | 4.851 | 4.851 |
| cp.html | 24,603 | 5.303 | 5.303 |
| fields.c | 11,150 | 5.121 | 5.121 |
| grammar.lsp | 3,721 | 4.848 | 4.848 |
| kennedy.xls | 1,029,744 | 3.596 | 3.598 |
| lcet10.txt | 426,754 | 4.699 | 4.699 |
| plrabn12.txt | 481,861 | 4.578 | 4.578 |
| ptt5 | 513,216 | 1.664 | 1.347 |
| sum | 38,240 | 5.430 | 5.432 |
| xargs.1 | 4,227 | 5.087 | 5.087 |
| file1 | 1,372,962 | 2.112 | 1.903 |
| file2 | 1,268,406 | 2.736 | 2.570 |

(bits/byte)

アルファベティック符号とアルファベティック準瞬時FV符号

アルファベティック符号:

情報源系列の辞書順と符号語系列の辞書順が一致する
符号

単一符号木の場合の最適な符号の構成

Hu-Tucker algorithm (Hu, Tucker, 1971)

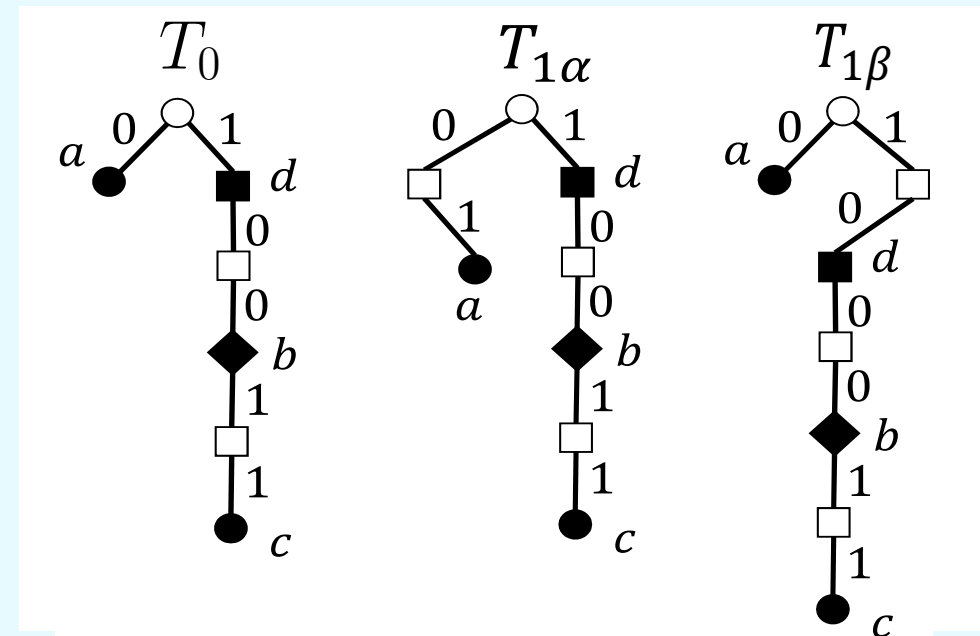
アルファベティック準瞬時符号:

符号木の最適な更新法:

Open problem

Hu-Tucker符号木を用いた簡易構成法

(平岡, 山本, SITA2017
11月29日4.2.3)



2元準瞬時符号の拡張

ハフマン符号:

符号木数: 1, 復号遅延: 0

最悪冗長度: 1 $H(X) \leq L_H < H(X) + 1$

準瞬時FV符号:

符号木数: 2, 復号遅延: 最大2ビット

最悪冗長度: 0.5 $H(X) \leq L_{AIFV} < H(X) + 0.5$

Question:

符号木数と復号遅延を, 2より増やすと, 最悪冗長度を0.5より減らせられるか?

2元準瞬時符号の拡張

ハフマン符号:

符号木数: 1, 復号遅延: 0

最悪冗長度: 1 $H(X) \leq L_H < H(X) + 1$

準瞬時FV符号:

符号木数: 2, 復号遅延: 最大2ビット

最悪冗長度: 0.5 $H(X) \leq L_{AIFV} < H(X) + 0.5$

準瞬時FV- m 符号(AIFV- m code): (Hu, Yamamoto, Honda, 2017)

符号木数: m , 復号遅延: 最大 m ビット

最悪冗長度: $1/m$ $H(X) \leq L_{AIFV-m} < H(X) + \frac{1}{m}$

($2 \leq m \leq 4$ に対して証明, $m \geq 5$ に対して予想)

準瞬時FV- m 符号の平均符号長

例.5 S1: $P_X = (0.900, 0.050, 0.049, 0.001)$
 S2: $P_X = (0.33, 0.3, 0.119, 0.1, 0.05, 0.04, 0.03, 0.02, 0.01, 0.001)$

平均符号長 (河井, 岩田, 山本, 2017)

| m | (Huffman) | 2 | 3 | 4 | 5 | $H(X)$ |
|-----|-----------|--------|--------|--------|--------|--------|
| S1 | 1.200 | 0.6252 | 0.4925 | 0.4395 | 0.4186 | 0.3373 |
| S2 | 2.336 | 2.2963 | 2.2865 | 2.2841 | 2.2839 | 2.2719 |

準瞬時FV- m 符号(AIFV- m code) : (Hu, Yamamoto, Honda, 2017)

符号木数: m , 復号遅延: 最大 m ビット

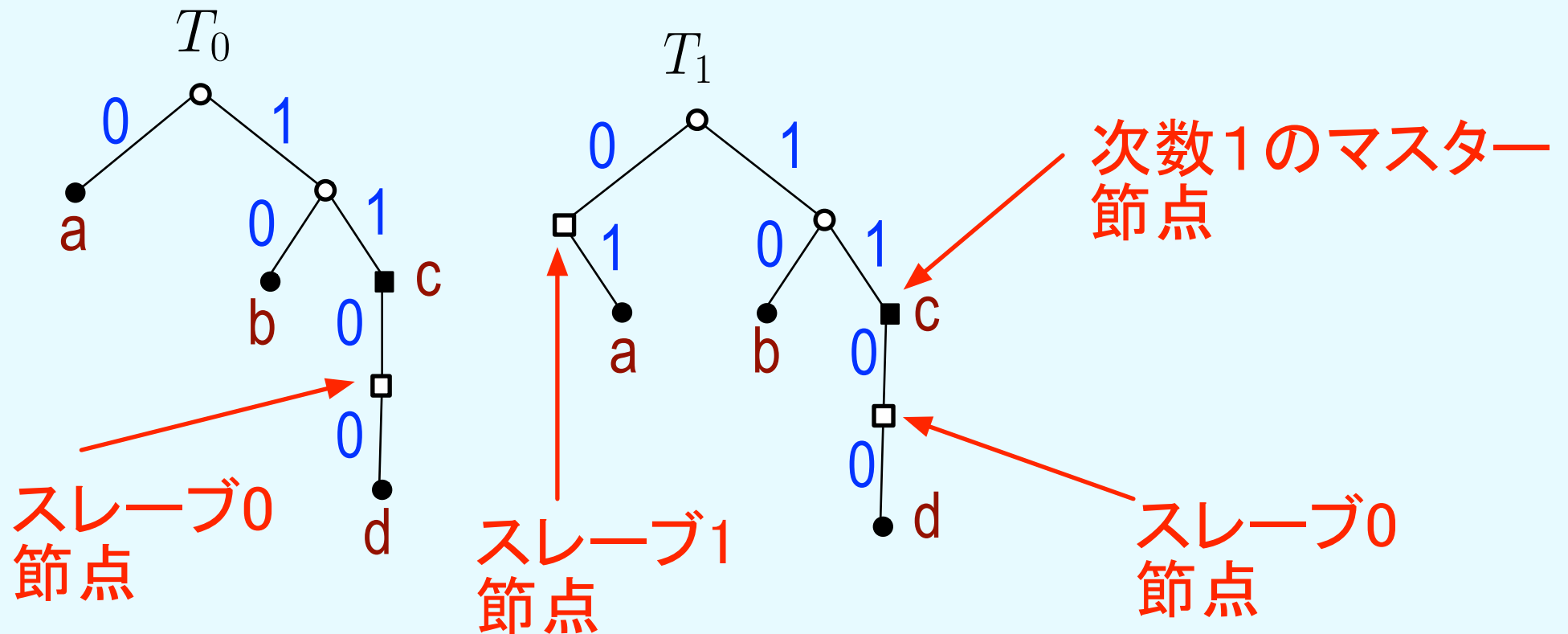
最悪冗長度: $1/m$

$$H(X) \leq L_{AIFV-m} < H(X) + \frac{1}{m}$$

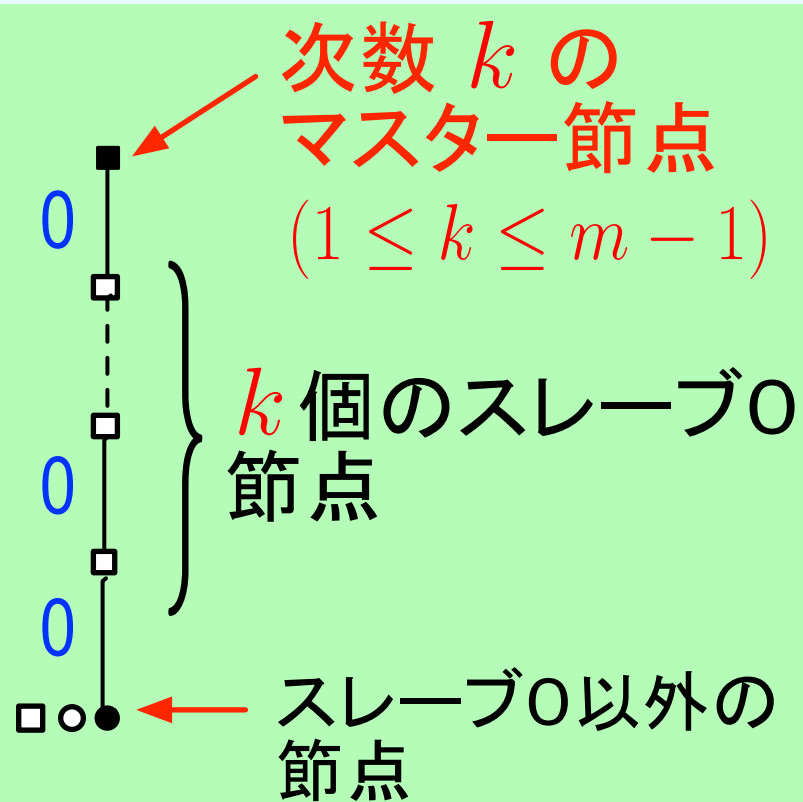
($2 \leq m \leq 4$ に対して証明, $m \geq 5$ に対して予想)

2元準瞬時FV- m 符号 (Hu, Yamamoto, Honda, 2017)

2元準瞬時FV- m 符号木



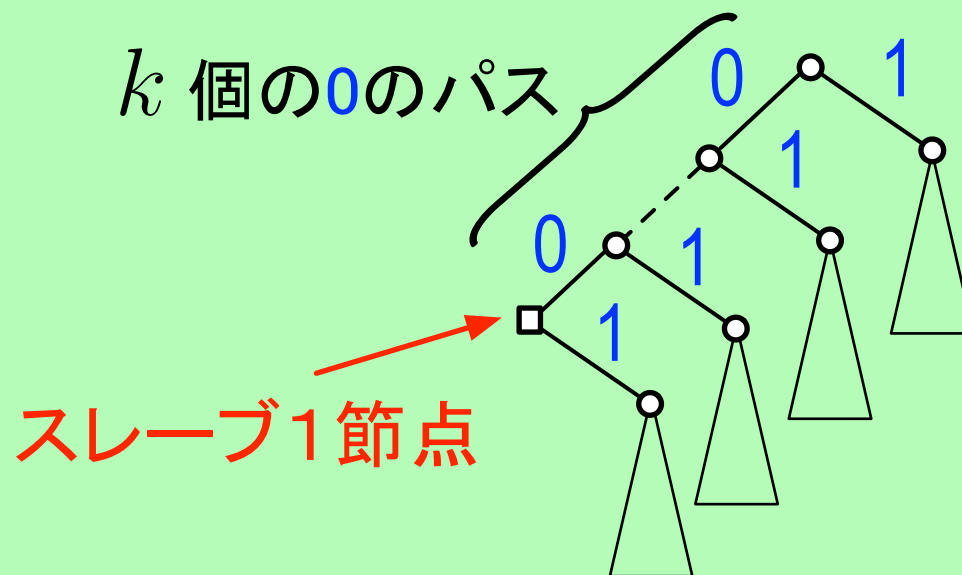
2元準瞬時FV- m 符号 (Hu, Yamamoto, Honda, 2017)



葉: 次数0のマスター
節点

符号木: T_0, T_1, \dots, T_{m-1}

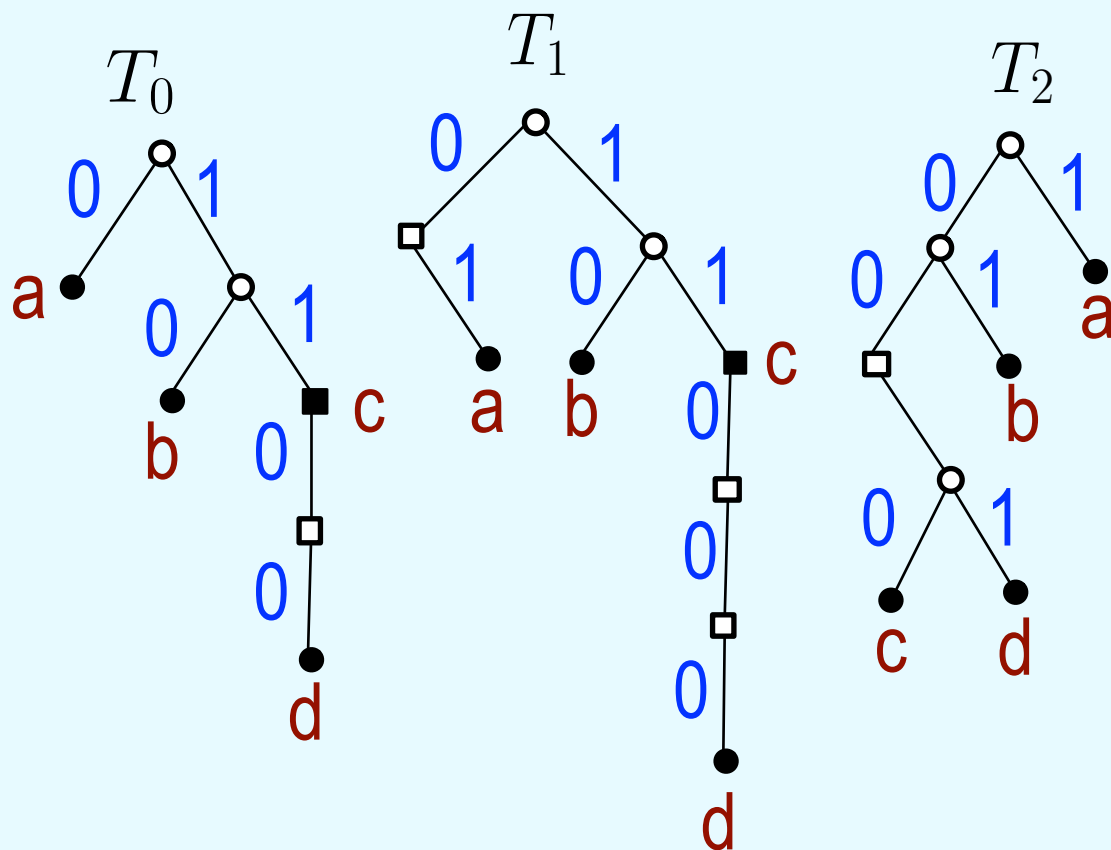
T_k ($1 \leq k \leq m - 1$)



情報源シンボルを次数0~ $m - 1$
のマスター節点に割り振る.

2元準瞬時FV- m 符号 (Hu, Yamamoto, Honda, 2017)

例.6 $\mathcal{X} = \{a, b, c, d\}, m = 3$



符号木の遷移則:

x_i が次数 k のマスター節点
で符号化される.

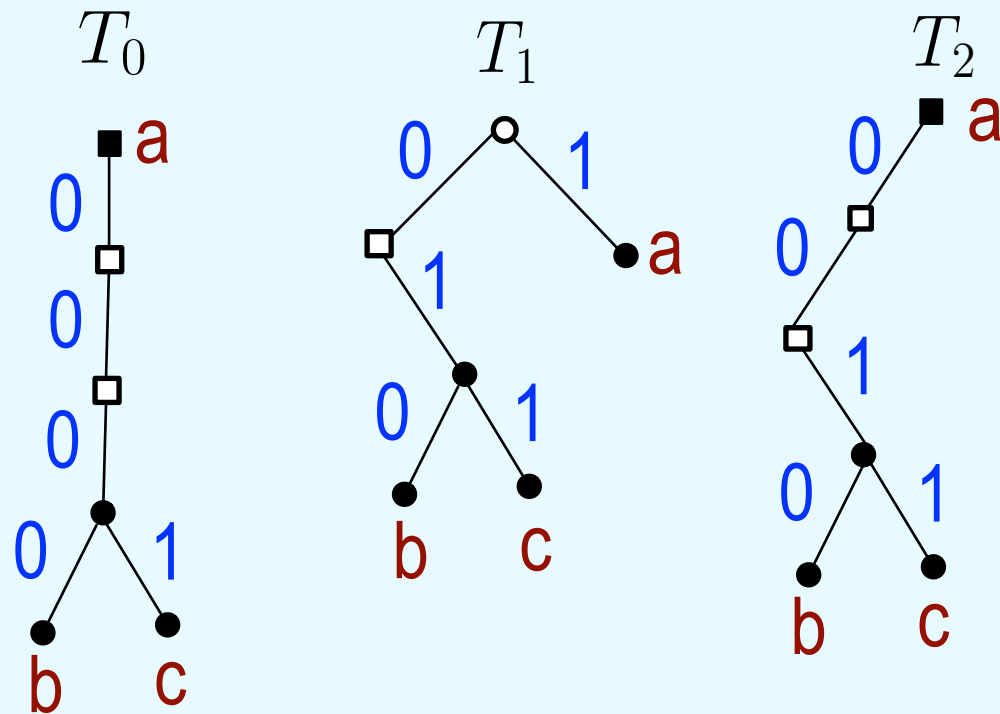


x_{i+1} を T_k で符号化

| | | | | |
|---|----|----|----|------|
| a | c | c | b | d |
| 0 | 11 | 11 | 01 | 1100 |

Binary AIFV- m codes (Hu, Yamamoto, Honda, 2017)

例.7 $\mathcal{X} = \{a, b, c\}, m = 3$



符号木の遷移即:

x_i が次数 k のマスター節点で符号化される.



x_{i+1} を T_k で符号化

a a a b a c
 λ λ 1 0000 λ 0011
 (λ : null)

100000011



2元準瞬時FV- m 符号 (Hu, Yamamoto, Honda, 2017)

$$p_{\max} = \max_{x \in \mathcal{X}} P(x)$$

準瞬時FV-3符号

準瞬時FV-4符号

IEEE Trans. Inform. Theory, vol.63, no.8, p.5084, Aug. 2017のFig.13を参照せよ.

IEEE Trans. Inform. Theory, vol.63, no.8, p.5085, Aug. 2017のFig.15を参照せよ.

$p_{\max} \leq 0.5$ のとき, 冗長度は0.25以下

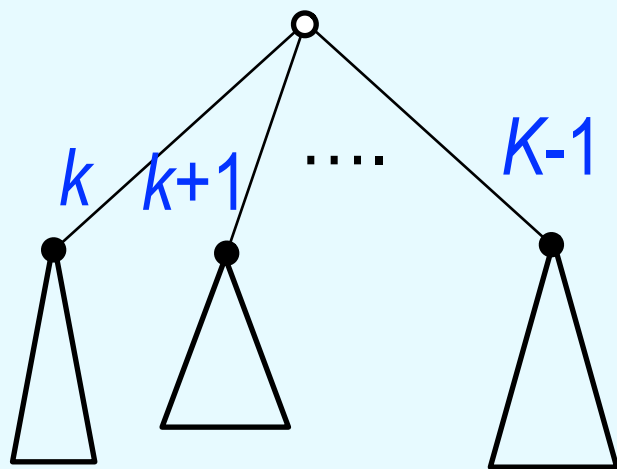
K 元準瞬時FV符号 ($K \geq 3$)

(Yamamoto, Wei, 2013)

(Yamamoto, Tsuchihashi, Honda, 2015)

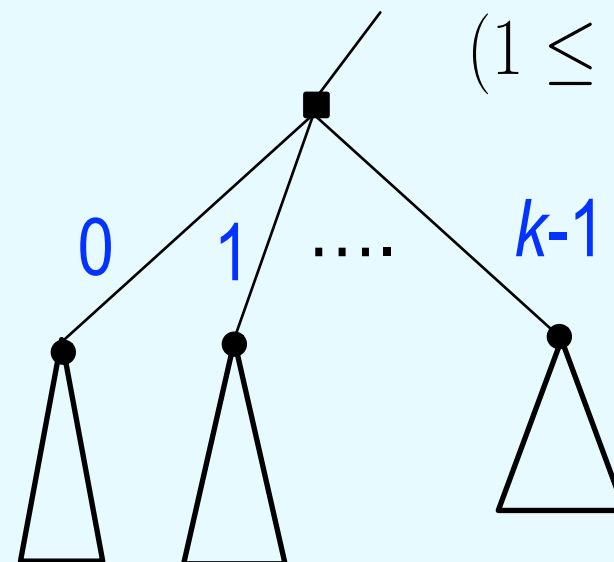
$$\mathcal{Y} = \{0, 1, 2, \dots, K-1\}$$

T_k の根 ($0 \leq k \leq K-2$)



次数 k のマスター節点

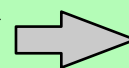
($1 \leq k \leq K-2$)



(葉: 次数0のマスター節点)

符号木の遷移即:

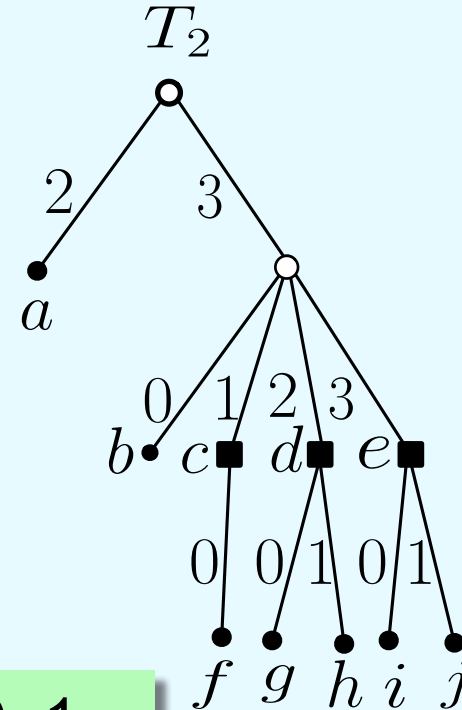
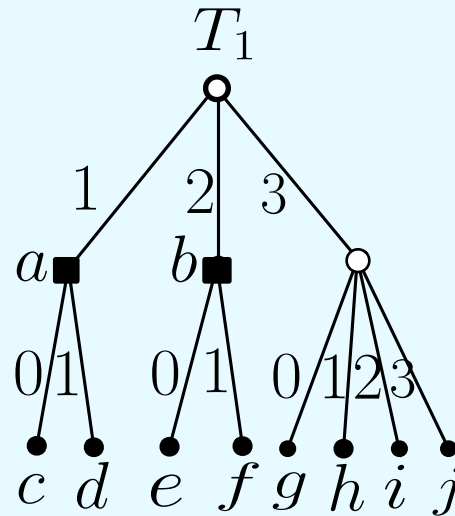
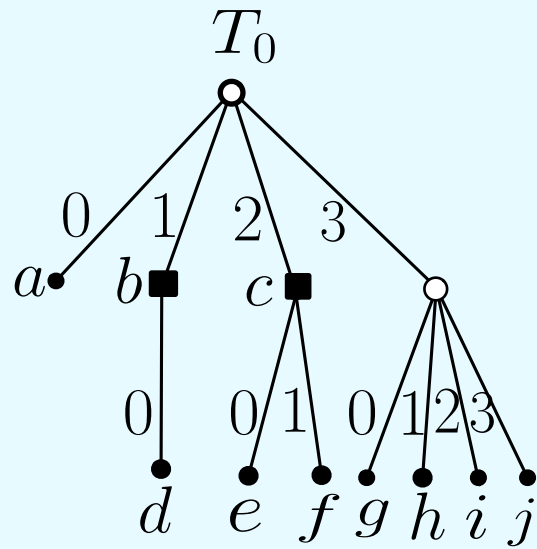
x_i が次数 k のマスター節点
で符号化される.



x_{i+1} の符号化に T_k を使用

K元準瞬時FV符号 ($K \geq 3$)

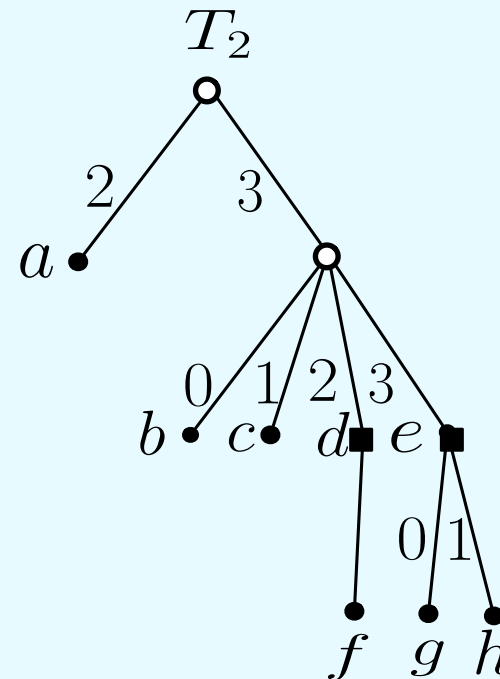
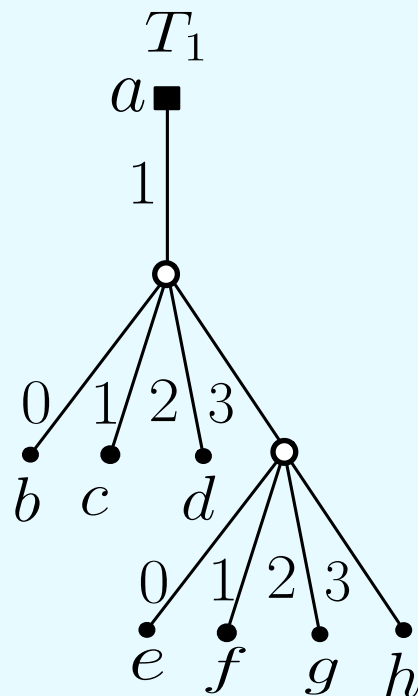
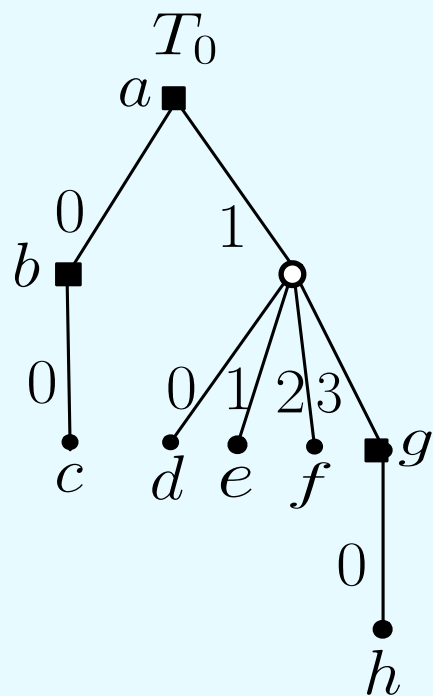
例.8 $K=4$ $\mathcal{Y} = \{0, 1, 2, 3\}$



| | | | | | | | | | | |
|---------|---|---|---|----|----|---|----|----|---|---|
| T_k : | 0 | 0 | 1 | 2 | 1 | 0 | 2 | 2 | 0 | 1 |
| 情報源系列: | a | b | a | c | g | c | e | b | b | d |
| 符号語系列: | 0 | 1 | 1 | 31 | 30 | 2 | 33 | 30 | 1 | 1 |

K元準瞬時FV符号 ($K \geq 3$)

例.9 $K=4$ $\mathcal{Y} = \{0, 1, 2, 3\}$



T_k の根が t 個の子

➡ $k + t$ 次の
マスター節点

| | | | | | | | | | | |
|---------|---|-----------|----|----|-----------|----|----|-----------|----|----|
| T_k : | 0 | 1 | 2 | 1 | 0 | 2 | 0 | 1 | 2 | 2 |
| 情報源系列: | b | a | d | b | a | c | g | a | e | c |
| 符号語系列: | 0 | λ | 32 | 10 | λ | 31 | 13 | λ | 33 | 31 |



数値実験例

$$\mathcal{X} = \{a_1, a_2, \dots, a_n\}$$

IEEE Trans. Inform. Theory, vol.61, no.12, p.6443, Dec. 2015のFig.15を参照せよ.

IEEE Trans. Inform. Theory, vol.61, no.12, p.6443, Dec. 2015のFig.16を参照せよ.

$$P_1(a_t) = \frac{t}{A_1}$$

$$P_2(a_t) = \frac{t^2}{A_2}$$

準瞬時符号

FV (fixed-to-variable length)符号

ハフマン符号

準瞬時FV符号(AIFV code)

VF (variable-to-fixed length)符号

タンストール符号
(Tunstall, 1967)

準瞬時VF符号(AIVF code)
(Yamamoto, Yokoo, 2001)

準瞬時FV符号の同期系列

同期系列

符号語のビット誤り ➡ 復号誤り伝播 ➡ 同期回復

(楮, 山本, SITA2017, 11月29日4.2.4)

まとめ

2元準瞬時FV- m 符号 m 個の符号木
最大復号遅延 m ビット

$$H(X) \leq L_{AIFV-m} < H(X) + \frac{1}{m} \quad \text{for } 2 \leq m \leq 4$$

最適な準瞬時FV- m 符号木の構成: 反復最適化 + IP, DP等

動的準瞬時FV符号:
動的ハフマン符号木 + 準瞬時FV符号木の簡易構成法

2元準瞬時FV符号木の総数: large Schröder number

K元準瞬時FV符号

アルファベティック準瞬時FV符号

準瞬時VF符号

準瞬時FV符号の同期系列

Open problems

2元準瞬時FV- m 符号

m 個の符号木
最大復号遅延 m ビット

$m \geq 5$ に対して $H(X) \leq L_{AIFV-m} < H(X) + \frac{1}{m}$ を証明せよ.

$m = 2$ の場合

2元準瞬時FV符号は, 符号木数2, 最大遅延2ビットの一位復号可能符号の中で最適か? (最適性証明 or 反例)

$m \geq 3$ の場合

準瞬時FV- m 符号は, 符号木数 m , 最大遅延 m ビットの一位復号可能符号の中で必ずしも最適でない.

➡ 最適な符号, より性能のよい符号

Open problems

2元準瞬時FV- m 符号

m 個の符号木
最大復号遅延 m ビット

最適な準瞬時FV- m 符号木の反復最適化を用いない構成法の導出.

動的準瞬時FV符号木の最適でシンプルな更新法

K元準瞬時FV符号の冗長度のタイトな上界の導出

⋮
⋮

Thank you for your attention.