

Getting close to $H_r(\mathcal{S})$

Recall

$$H_r(\mathcal{S}) = \sum_{i=1}^q p_i \log_r \frac{1}{p_i}.$$

$H(\mathcal{S})$ is “sandwiched”:

$$0 \leq H_r(\mathcal{S}) \leq \log_r q$$

Theorem

If \mathcal{C} is a uniquely decodable r -ary code for source \mathcal{S} , then $H_r(\mathcal{S}) \leq L(\mathcal{C})$.

With luck, $L(\mathcal{C}) = H_r(\mathcal{S})$

Corollary (Equality when $p_i = r^{e_i}$; $e_i \in \mathbb{Z}$; $e_i \leq 0$)

Let \mathcal{S} be a source w. probabilities p_i . There exists a uniquely decodable r -ary code \mathcal{C} for \mathcal{S} with $L(\mathcal{C}) = H_r(\mathcal{S})$ iff $\log_r p_i \in \mathbb{Z}$, $\forall i$.^a

^aThat is, $p_i = r^{e_i}$, for some integer $e_i \leq 0$

Shannon-Fano Coding

What do we do then, to get as close to $H_r(\mathcal{S})$ as possible¹?

- Any $p_i \in P$ can, for any r and P , be written as $\frac{1}{r^{x_i}}$ for $x_i \in \mathbb{R}$. We want code-word lengths $l_i = x_i$, but l_i must be *integers*.

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source Coding Theorem

$H_r(\mathcal{S})$ and Source Extension

Theorem

Code Generation Story

Epilogue

The Big Picture Next Week

¹ *without* applying *source extension*.

² You can construct these with the tree-building technique for instantaneous codes.

Shannon-Fano Coding

What do we do then, to get as close to $H_r(\mathcal{S})$ as possible¹?

- Any $p_i \in P$ can, for any r and P , be written as $\frac{1}{r^{x_i}}$ for $x_i \in \mathbb{R}$. We want code-word lengths $l_i = x_i$, but l_i must be *integers*.
- Overapproximate*; let $l_i = \lceil x_i \rceil = \left\lceil \log_r \frac{1}{p_i} \right\rceil$. So

$$\log_r \frac{1}{p_i} \leq l_i < 1 + \log_r \frac{1}{p_i}.$$

and $p_i = \frac{1}{r^{x_i}} \geq \frac{1}{r^{l_i}} = r^{-l_i}$. Thus

$$K = \sum_{i=1}^q r^{-l_i} \leq \sum_{i=1}^q p_i = 1.$$

We have shown, by “Kraft”, the *existence* of an instantaneous r -ary code \mathcal{C} for \mathcal{S} with word-lengths l_i . And $L(\mathcal{C})$ is nearly $H_r(\mathcal{S})$. \mathcal{C} built from these l_i is a *Shannon-Fano code*².

¹without applying source extension.

²You can construct these with the tree-building technique for instantaneous codes.

What is Really interesting...

Theorem

Every r -ary Shannon-Fano code \mathcal{C} for a source \mathcal{S} satisfies

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S}).$$

Proof.

Recall $l_i < 1 + \log_r \frac{1}{p_i}$.^a Multiply by p_i , and sum for each i :

$$\sum_{i=1}^q p_i l_i < \sum_{i=1}^q p_i (1 + \log_r \frac{1}{p_i}) = 1 + \sum_{i=1}^q p_i \log_r \frac{1}{p_i}.$$

Thus $L(\mathcal{C}) < 1 + H_r(\mathcal{S})$. Extendible to $L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$. □

^a l_i are the lengths of the Shannon-Fano code \mathcal{S} , and p_i are the probabilities of source s_i which \mathcal{C} attaches w_i ; $|w_i| = l_i$ to.

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Why is that interesting?

Corollary

For any optimal r -ary code \mathcal{D} ,

$$H_r(S) \leq L(\mathcal{D}) \leq 1 + H_r(S)$$

Proof.

Since any Shannon-Fano code \mathcal{C} is uniquely decodable, we have

$$H_r(S) \leq L(\mathcal{C}) \leq 1 + H_r(S).$$

Since $H_r(S) \leq L(\mathcal{D}) \leq L(\mathcal{C})$ (\mathcal{D} optimal), we get

$$H_r(S) \leq L(\mathcal{D}) \leq 1 + H_r(S).$$



The Hunt For $H_r(\mathcal{S}) = L(\mathcal{D})$

So, $L(\mathcal{D})$ for an optimal \mathcal{D} is “sandwiched”.

$$H_r(\mathcal{S}) \leq L(\mathcal{D}) \leq 1 + H_r(\mathcal{S}).$$

Our optimal codes are, well, fairly optimal.

- With immense luck, we find a \mathcal{D} s.t. $H_r(\mathcal{S}) = L(\mathcal{D})$.
- Most likely, $L(\mathcal{D})$ is between $H_r(\mathcal{S})$ and $1 + H_r(\mathcal{S})$ (not biased towards one or the other).
- We want to get $H_r(\mathcal{S}) = L(\mathcal{D})$. At least *arbitrarily* close.

Spoiler: You can do that with *source extension*.

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Can we do Better?

Example (Biased coin, further improvement?)

Recall

$$\mathcal{C} = \{0, 1\},$$

the Huffman code for $\mathcal{S} = (S, P)$, with $S = \{\text{head}, \text{tails}\}$, and $P = (\frac{2}{3}, \frac{1}{3})$.

We have $L(\mathcal{C}) = \frac{2}{3} + \frac{1}{3}$

- We have proven Huffman codes are optimal for P .

Q: Have we at last found the best encoding for \mathcal{S} ? Can we do better?

Can we do Better?

Example (Biased coin, further improvement?)

Recall

$$\mathcal{C} = \{0, 1\},$$

the Huffman code for $S = (S, P)$, with $S = \{\text{head}, \text{tails}\}$, and $P = (\frac{2}{3}, \frac{1}{3})$.

We have $L(\mathcal{C}) = \frac{2}{3} + \frac{1}{3}$

- We have proven Huffman codes are optimal for P .

Q: Have we at last found the best encoding for S ? Can we do better?

A: Yes! We can!

Optimality of
Huffman Codes $L(\mathcal{C})$ of Huffman
CodesHuffman Codes
are Optimal

Source Extension

Information

 $I(s)$ measure $I(s)$ is uniquely
defined

Entropy

 $H(S)$ measureProperties of $H(S)$

Epilogue

Can we do Better?

Example (Biased coin, further improvement?)

Recall

$$\mathcal{C} = \{0, 1\},$$

the Huffman code for $\mathcal{S} = (S, P)$, with $S = \{\text{head}, \text{tails}\}$, and $P = (\frac{2}{3}, \frac{1}{3})$.

We have $L(\mathcal{C}) = \frac{2}{3} + \frac{1}{3}$

- We have proven Huffman codes are optimal for P .

Q: Have we at last found the best encoding for \mathcal{S} ? Can we do better?

A: Yes! We can!

Let $\mathcal{S}^n = (S^n, P^n)$ be the n th *extension* of \mathcal{S} . Let's try for $n = 2$.

Can we do Better?

Lecture 3: A Measure of Information

Example (Biased coin, further improvement?)

Recall

$$\mathcal{C} = \{0, 1\},$$

the Huffman code for $\mathcal{S} = (S, P)$, with $S = \{\text{head}, \text{tails}\}$, and $P = (\frac{2}{3}, \frac{1}{3})$.

We have $L(\mathcal{C}) = \frac{2}{3} + \frac{1}{3}$

- We have proven Huffman codes are optimal for P .

Q: Have we at last found the best encoding for \mathcal{S} ? Can we do better?

A: Yes! We can!

Let $\mathcal{S}^n = (S^n, P^n)$ be the n th *extension* of \mathcal{S} . Let's try for $n = 2$. Here, $S^2 = \{\text{hh}, \text{ht}, \text{th}, \text{tt}\}$, and $P^2 = (\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9})$. Huffman code \mathcal{C}^2 of \mathcal{S}^2 :

$$\mathcal{C}^2 = \{1, 01, 001, 000\}.$$

Now $L_2 = L(\mathcal{C}^2) = 1 * \frac{4}{9} + 2 * \frac{2}{9} + 3 * \frac{2}{9} + 3 * \frac{1}{9} = \frac{17}{9}$. How does that help?

Optimality of Huffman Codes
$L(\mathcal{C})$ of Huffman Codes
Huffman Codes are Optimal
Source Extension
Information
$I(s)$ measure
$I(s)$ is uniquely defined
Entropy
$H(S)$ measure
Properties of $H(S)$
Epilogue

Can we do Better?

Lecture 3: A
Measure of
Information

Example (Biased coin, further improvement?)

Recall

$$\mathcal{C} = \{0, 1\},$$

the Huffman code for $\mathcal{S} = (S, P)$, with $S = \{\text{head}, \text{tails}\}$, and $P = (\frac{2}{3}, \frac{1}{3})$.
We have $L(\mathcal{C}) = \frac{2}{3} + \frac{1}{3}$

- We have proven Huffman codes are optimal for P .

Q: Have we at last found the best encoding for \mathcal{S} ? Can we do better?

A: Yes! We can!

Let $\mathcal{S}^n = (S^n, P^n)$ be the n th *extension* of \mathcal{S} . Let's try for $n = 2$. Here, $\mathcal{S}^2 = \{\text{hh}, \text{ht}, \text{th}, \text{tt}\}$, and $P^2 = (\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9})$. Huffman code \mathcal{C}^2 of \mathcal{S}^2 :

$$\mathcal{C}^2 = \{1, 01, 001, 000\}.$$

Now $L_2 = L(\mathcal{C}^2) = 1 * \frac{4}{9} + 2 * \frac{2}{9} + 3 * \frac{2}{9} + 3 * \frac{1}{9} = \frac{17}{9}$. How does that help?
Since each $s^2 \in \mathcal{S}^2$ represents some *two* symbols $s_i, s_j \in \mathcal{S}$ *in sequence*, as an *encoding of \mathcal{S}* , \mathcal{C}^2 has average word length

$$\frac{L_2}{2} = \frac{17}{18} = 0.944 \dots < L(\mathcal{C}) = 1.$$

Optimality of
Huffman Codes
 $L(\mathcal{C})$ of Huffman
Codes
Huffman Codes
are Optimal
Source Extension
Information
 $I(s)$ measure
 $I(s)$ is uniquely
defined
Entropy
 $H(S)$ measure
Properties of $H(S)$
Epilogue

Amazing!

- We obtained an *encoding* which reduced the average word length *even further!*
- This is allowed since $\sum_{i,j} p_{i,j} = (p_1 + \dots + p_q)^n = 1^n = 1$, that is, P^n is a probability distribution.
- This works since the “gap” between $\min(P^n)$ and $\max(P^n)$ increases (if P was not uniform, that is).

Advantage: We obtain a better code for S ! Map $w \in \mathcal{C}^n$ to $s^n \in S$ to $s \in S^*$. Uniquely decodable!

Disadvantage: Not instantaneous; *bounded* decoding delay.

Lecture 3: A Measure of Information

Optimality of Huffman Codes
$L(\mathcal{C})$ of Huffman Codes
Huffman Codes are Optimal
Source Extension
Information
$I(s)$ measure
$I(s)$ is uniquely defined
Entropy
$H(S)$ measure
Properties of $H(S)$
Epilogue

Spoiler

We could continue:

$$\frac{L_3}{3} = \frac{76}{81} = 0.938\ldots < \frac{L_2}{2}$$

Lecture 3: A Measure of Information

Optimality of Huffman Codes

$L(C)$ of Huffman Codes

Huffman Codes are Optimal

Source Extension

Information

$I(s)$ measure

$I(s)$ is uniquely defined

Entropy

$H(S)$ measure

Properties of $H(S)$

Epilogue

Spoiler

We could continue:

$$\frac{L_3}{3} = \frac{76}{81} = 0.938\dots < \frac{L_2}{2}$$

Now, the ultimate question is:

What happens to $\frac{L_n}{n}$ as $n \rightarrow \infty$?

Lecture 3: A Measure of Information

Optimality of Huffman Codes
$L(\mathcal{C})$ of Huffman Codes
Huffman Codes are Optimal
Source Extension
Information
$I(s)$ measure
$I(s)$ is uniquely defined
Entropy
$H(S)$ measure
Properties of $H(S)$
Epilogue

Does it Ever End?

We have seen that by applying *source extension*, we obtain better codes, at the cost of getting a decoding delay.



Figure: Communication System with Source Extension

We have seen a measure of *information*.

Example (Biased coin)

For $\mathcal{S} = ((h, t), (\frac{2}{3}, \frac{1}{3}))$, we have $H_r(\mathcal{S}) \approx 0.9183$. Recall $L_1 = 1$, $\frac{L_2}{2} = 0.944$, and $\frac{L_3}{3} = 0.938$.

We have now laid the groundwork for Shannon's Theorem!

Theorem (Shannon's Source Coding Theorem)

By encoding \mathcal{S}^n with sufficiently large n , one can find uniquely decodable r -ary encodings of \mathcal{S} with average word-lengths arbitrarily close to $H_r(\mathcal{S})$.

So, $\frac{L_n}{n} \rightarrow H_r(\mathcal{S})$ as $n \rightarrow \infty$. Amazing. We will see its **proof** tomorrow.

Entropy of Extended Source

Let \mathcal{S}, \mathcal{T} be any two sources. p_i, q_j probability distributions for \mathcal{S} and \mathcal{T} .

- $\mathcal{S} \times \mathcal{T}$ produce symbol pairs, $(s_i, t_j) = s_i t_j$ w. $\Pr(s_i \text{ and } t_j)$.
- $\mathcal{S} \times \mathcal{T}$ independent if $\Pr(s_i \text{ and } t_j) = p_i q_j$.

Consider $\mathcal{S} \times \mathcal{T}$ as 1 source. So you sample from \mathcal{S}, \mathcal{T} simultaneously.

Theorem (\mathcal{S}^n provides n times the information of \mathcal{S})

If \mathcal{S} is any source, then $H_r(\mathcal{S}^n) = nH_r(\mathcal{S})$

Proof Idea.

Prove that

$$H_r(\mathcal{S} \times \mathcal{T}) = H_r(\mathcal{S}) + H_r(\mathcal{T}).$$

Then by induction,

$$H_r(\mathcal{S}^n) = \underbrace{H_r(\mathcal{S} \times \cdots \times \mathcal{S})}_{n \text{ times}} = \underbrace{H_r(\mathcal{S}) + \cdots + H_r(\mathcal{S})}_{n \text{ times}} = nH_r(\mathcal{S}).$$



Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano
Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source
Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension

Theorem

Code Generation
Story

Epilogue

The Big Picture
Next Week

Lemma

If \mathcal{S} and \mathcal{T} are independent sources,

$$H_r(\mathcal{S} \times \mathcal{T}) = H_r(\mathcal{S}) + H_r(\mathcal{T}).$$

Proof.

Independence gives $\Pr(s_i t_j) = p_i q_j$. So,

$$\begin{aligned} H_r(\mathcal{S} \times \mathcal{T}) &= - \sum_{i,j} p_i q_j \log_r p_i q_j = - \sum_i \sum_j p_i q_j (\log_r p_i + \log_r q_j) \\ &= - \sum_i \sum_j (p_i q_j \log_r p_i + p_i q_j \log_r q_j) \\ &= - \sum_i \sum_j p_i q_j \log_r p_i - \sum_i \sum_j p_i q_j \log_r q_j \\ &= \left(\sum_j q_j \right) \left(- \sum_i p_i \log_r p_i \right) + \left(\sum_i p_i \right) \left(- \sum_j q_j \log_r q_j \right) \\ &= H_r(\mathcal{S}) + H_r(\mathcal{T}), \end{aligned}$$

as $\sum_i p_i = \sum_j q_j = 1$. □

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source Coding Theorem

$H_r(\mathcal{S})$ and Source Extension

Theorem

Code Generation Story

Epilogue

The Big Picture Next Week

Theorem

Theorem (Shannon's Source Coding Theorem)

By encoding \mathcal{S}^n with sufficiently large n , one can find uniquely decodable r -ary encodings of a source \mathcal{S} with average word-lengths arbitrarily close to $H_r(\mathcal{C})$.

Proof.

Let $L_n = L(\mathcal{C}^n)$, the average word-length of an optimal code for \mathcal{S}^n .

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source Coding Theorem

$H_r(\mathcal{S})$ and Source Extension

Theorem

Code Generation Story

Epilogue

The Big Picture Next Week

Theorem

Theorem (Shannon's Source Coding Theorem)

By encoding \mathcal{S}^n with sufficiently large n , one can find uniquely decodable r -ary encodings of a source \mathcal{S} with average word-lengths arbitrarily close to $H_r(\mathcal{S})$.

Proof.

Let $L_n = L(\mathcal{C}^n)$, the average word-length of an optimal code for \mathcal{S}^n .
From optimality:

$$H_r(\mathcal{S}^n) \leq L_n \leq 1 + H_r(\mathcal{S}^n).$$

By above-proven theorem concerning entropy of source products,

$$nH_r(\mathcal{S}) \leq L_n \leq 1 + nH_r(\mathcal{S}).$$

Divide by n ,

$$H_r(\mathcal{S}) \leq \frac{L_n}{n} \leq \frac{1}{n} + H_r(\mathcal{S}).$$

As $\frac{L_n}{n}$ is the average word-length of \mathcal{C}^n regarded as an encoding of \mathcal{S} , the result follows:

$$\lim_{n \rightarrow \infty} \frac{L_n}{n} = H_r(\mathcal{S}).$$



$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Theorem

A Note

While “Shannon’s Source Coding Theorem” says we can get $\frac{L_n}{n}$ as close to $H_r(S)$ as we want, getting near enough may be hard, or yield an unacceptable code.

- Constructing codes: time-consuming.
- Decoding delay increases for big n .

Practical rule: Choose a maximal decoding delay d you deem acceptable. Generate codes, and source extensions up to d , and pick the best code from that.

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(S) \leq L(C)$$

$$L(C) \leq 1 + H_r(S)$$

Shannon's Source Coding Theorem

$H_r(S)$ and Source Extension

Theorem

Code Generation Story

Epilogue

The Big Picture Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (\mathcal{S} := (h, t), P := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous*.

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano
Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source
Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension
Theorem

Code Generation
Story

Epilogue

The Big Picture
Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (\mathcal{S} := (h, t), P := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous*.

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem:

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano
Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source
Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension
Theorem

Code Generation
Story

Epilogue

The Big Picture
Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (\mathcal{S} := (\mathbf{h}, \mathbf{t}), \mathbf{P} := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous*.

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem: *not instantaneous*.

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano
Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source
Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension
Theorem

Code Generation
Story

Epilogue

The Big Picture
Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (S := (h, t), P := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous*.

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem: *not instantaneous*.

Improvement: Turn \mathcal{C} into a blockcode: Set $\mathcal{C} = (10, 01)$.

Problem:

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source Coding Theorem

$H_r(\mathcal{S})$ and Source Extension Theorem

Code Generation Story

Epilogue

The Big Picture Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (\mathcal{S} := (\mathbf{h}, \mathbf{t}), \mathbf{P} := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous.*

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem: *not instantaneous.*

Improvement: Turn \mathcal{C} into a blockcode: Set $\mathcal{C} = (10, 01)$.

Problem: *not optimal.*

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano
Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source
Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension
Theorem

Code Generation
Story

Epilogue

The Big Picture
Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (S := (h, t), P := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous.*

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem: *not instantaneous.*

Improvement: Turn \mathcal{C} into a blockcode: Set $\mathcal{C} = (10, 01)$.

Problem: *not optimal.*

Improvement: Apply Huffman. $\mathcal{C} = (0, 1)$.

Problem:

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension
Theorem

Code Generation Story

Epilogue

The Big Picture
Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (S := (h, t), P := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous*.

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem: *not instantaneous*.

Improvement: Turn \mathcal{C} into a blockcode: Set $\mathcal{C} = (10, 01)$.

Problem: *not optimal*.

Improvement: Apply Huffman. $\mathcal{C} = (0, 1)$.

Problem: Does not match $H_r(\mathcal{S}) \approx 0.918$ with equality.

Lecture 4: Shannon's Source Coding Theorem

Shannon-Fano Coding

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Shannon's Source Coding Theorem

$H_r(\mathcal{S})$ and Source
Extension
Theorem

Code Generation Story

Epilogue

The Big Picture
Next Week

From Worst to Best

Let $r = 2$ and $\mathcal{S} = (\mathcal{S} := (h, t), P := (\frac{2}{3}, \frac{1}{3}))$, where $q = 2$.

Naïve approach: Set $\mathcal{C} = (1, 11)$.

Problem: *ambiguous.*

Improvement: Turn \mathcal{C} into a *uniquely decodable* code: Set $\mathcal{C} = (0, 01)$.

Problem: *not instantaneous.*

Improvement: Turn \mathcal{C} into a blockcode: Set $\mathcal{C} = (10, 01)$.

Problem: *not optimal.*

Improvement: Apply Huffman. $\mathcal{C} = (0, 1)$.

Problem: Does not match $H_r(\mathcal{S}) \approx 0.918$ with equality.

n	1	2	3	4	5
L	1.333	1.333	1	1.083	0.933
η	0.689	0.689	0.918	0.848	0.984

Table: Shannon-Fano Coding for \mathcal{S}^n

n	1	2	3	4	5
L	1	0.944	0.938	0.938	0.923
η	0.918	0.972	0.979	0.979	0.995

Table: Huffman Coding for \mathcal{S}^n

$$H_r(\mathcal{S}) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(\mathcal{S})$$

Full Overview of Source Coding

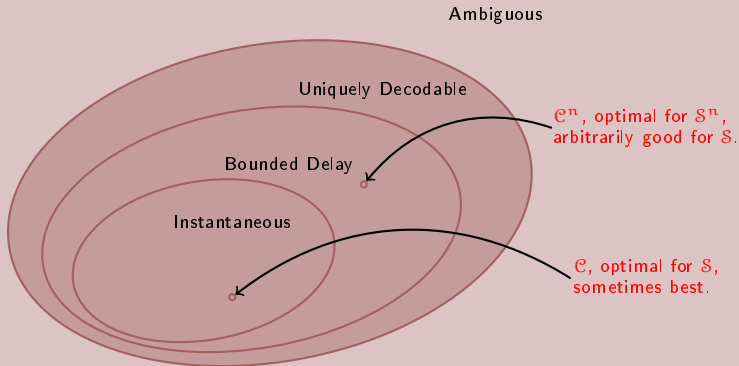


Figure: Classes of Source Codes.

Lecture 4:
Shannon's Source
Coding TheoremShannon-Fano
Coding

$$H_r(S) \leq L(\mathcal{C})$$

$$L(\mathcal{C}) \leq 1 + H_r(S)$$

Shannon's Source
Coding Theorem

$H_r(S)$ and Source
Extension
Theorem

Code Generation
Story

Epilogue

The Big Picture
Next Week