

Probability Measure

Definition (Sample Space, Event, Probability Measure)

Let S be a *sample space*^a, and $A \subseteq S$ an *event*. $\mathcal{P}(S)$ is the set of all events of S . A *probability measure* is a function $\Pr: \mathcal{P}(S) \rightarrow [0, 1]$ mapping events to probabilities, satisfying $\Pr(S) = 1$ ^b, and for all disjoint events $A_i \subseteq S$ it holds that

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i).$$

^aAll possible outcomes of a random experiment.

^bThe *certain event*.

Lecture 1: Unique Decodability

Source

Preliminaries

Probability Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-Patterson

Epilogue

Summary

Next Session

Random Variable

Definition (Random Variable (Stochastic Variable))

A *random variable* is a function $X: S \rightarrow \mathbb{R}$, where S is the sample space of a random experiment. Denote $\mathcal{X} = \text{img}(X)$. X is *discrete* if \mathcal{X} is countable. X is *real* if $\mathcal{X} \subseteq \mathbb{R}$.

Lecture 1: Unique Decodability

Source

Preliminaries

Probability Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-Patterson

Epilogue

Summary

Next Session

Random Variable

Lecture 1: Unique Decodability

Definition (Random Variable (Stochastic Variable))

A *random variable* is a function $X: S \rightarrow \mathbb{R}$, where S is the sample space of a random experiment. Denote $\mathcal{X} = \text{img}(X)$. X is *discrete* if \mathcal{X} is countable. X is *real* if $\mathcal{X} \subseteq \mathbb{R}$.

Example

$S = \{\text{head}, \text{tails}\}$ as before. Then

$$X(s) = \begin{cases} 1 & \text{if } s = \text{head}, \\ 0 & \text{if } s = \text{tails}. \end{cases}$$

Quantifying the outcome of a surgeon's operation, proposal.

$S = \{\text{success}, \text{failure}, \text{handicap}, \text{death}\}$.

$$X(s) = \begin{cases} 10 & \text{if } s = \text{success}, \\ 5 & \text{if } s = \text{failure}, \\ 2 & \text{if } s = \text{handicap}, \\ 0 & \text{if } s = \text{death}. \end{cases}$$

Source

Preliminaries

**Probability
Theory**

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-

Patterson

Epilogue

Summary

Next Session

Definition (Probability Distribution)

The list $P = (p_1, \dots, p_n)$, where $p_i \in [0, 1]$, is a *probability distribution* if

$$\sum_{i=1}^n p_i = 1.$$

Example

$P = (\frac{1}{2}, \frac{1}{2})$ and $P' = (\frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{8})$ are probability distributions.
 $P'' = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6})$ is not.

Lecture 1: Unique Decodability

Source

Preliminaries
Probability Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique
Decodability

Definition

Block Codes

Good Codes

Sardinas-
Patterson

Epilogue

Summary

Next Session

Definition (Probability Distribution)

The list $P = (p_1, \dots, p_n)$, where $p_i \in [0, 1]$, is a *probability distribution* if

$$\sum_{i=1}^n p_i = 1.$$

Example

$P = (\frac{1}{2}, \frac{1}{2})$ and $P' = (\frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \frac{1}{8})$ are probability distributions.
 $P'' = (\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6})$ is not.

Definition (Probability Mass Function)

A *probability mass function* for a discrete random variable X is a function $p_X : \mathcal{X} \rightarrow [0, 1]$, where $p_X(x) = \Pr(X = x)$ and

$$\sum_{x \in \mathcal{X}} p_X(x) = 1$$

Example

$S = \{\text{head}, \text{tails}\}$. Then $p_X(X(\text{head})) = p_X(1) = \frac{1}{2}$.

Lecture 1: Unique Decodability

Source

Preliminaries

Probability Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-Patterson

Epilogue

Summary

Next Session

Source Definition

We now formalise the *source*.

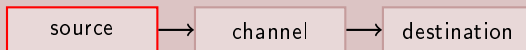


Figure: Communication System

Lecture 1: Unique Decodability

Source

Preliminaries

Probability

Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-

Patterson

Epilogue

Summary

Next Session

Source Definition

We now formalise the *source*.

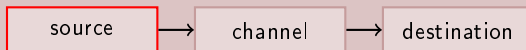


Figure: Communication System

Definition (Source)

A *source* \mathcal{S} is a pair (S, p) , where S is a finite set of symbols called the *source alphabet*, and $p : S \rightarrow [0, 1]$ is a probability mass function.

Note: It is tempting to consider \mathcal{S} a *random variable*. But \mathcal{S} never takes on a value.

Instead, we denote X_i as the *i*th *symbol* generated by \mathcal{S} . $\mathcal{X}_i = S$, and $p_{X_i} = p$.

Alternative Definition: S, p are lists, $|S| = |p|$, p is a probability distribution, and the likelihood of \mathcal{S} outputting $s_i \in S$ is $p_i \in p$.

Assumptions & More

Assumption (\mathcal{S} is stationary and memoryless)

- p does not change over time (stationary)
 - X_i does not depend on any X_j , where $j < i$ (memoryless)
-
- So, the X_i are *iid*¹. In more advanced treatment of the theory, these assumptions are lifted.
 - A finite sequence of symbols generated by \mathcal{S} : $\mathbf{s} = X_1 X_2 \dots X_n$ (Assume all such sequences are finite).
 - We denote the size of the source alphabet by q .

Lecture 1: Unique Decodability

Source

Preliminaries
Probability
Theory
Notation
Definition
Assumptions
Examples

Source Code

Motivation
Definition
Examples

Unique

Decodability
Definition
Block Codes
Good Codes
Sardinas-Patterson

Epilogue

Summary
Next Session

¹Independent and Identically Distributed

Example Sources

Example (S: Unbiased Die)

$S = \{1, \dots, 6\}$, $q = 6$, $s_i = i$ for $1 \leq i \leq 6$, $p_i = \frac{1}{6}$. X_j is the outcome of the j th throw.

Example (S: Weather observation at a particular place)

$S = \{1, 2, 3\}^a$, $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{2}$, $p_3 = \frac{1}{4}$. (ignore seasons: changing p_i).

^a1 for “good”, 2 for “moderate”, 3 for “bad”.

Note: We do not *care* about *what* $s \in S$ means.

Example (S: A Book)

S all symbols used in the book (English alphabet, space, punctuations, numerals, ...), X_i is the i th symbol in the book, p_i the frequency of i th symbol in S . (ignore preceding characters: dependence)

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples**

Source Code

- Motivation
- Definition
- Examples

Unique Decodability

- Definition
- Block Codes
- Good Codes
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Channel Radix $r < q$

Symbols are sent from the source through a channel.

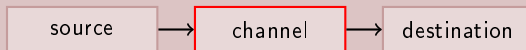


Figure: Communication System

- If the channel can carry q or more different types of values, sending a message s through the channel is trivial².

²Symbol-to-symbol mapping

Channel Radix $r < q$

Symbols are sent from the source through a channel.

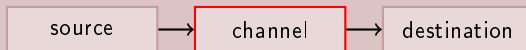


Figure: Communication System

- If the channel can carry q or more different types of values, sending a message s through the channel is trivial².
- If not, we need to *encode* s more cleverly to send it through the channel.

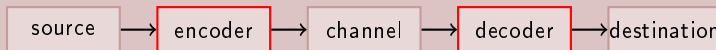


Figure: Communication System with Encoder & Decoder

We will thus assume $r < q$, and construct codes which utilise all r possible values. We call these r -ary codes. Usually, $r = 2$ (binary codes).

²Symbol-to-symbol mapping

Source Code Definition

Definition (Code Alphabet, Words)

Let $T = \{t_1, \dots, t_r\}$ be the finite set of r *code symbols* t_j , called the *code alphabet*.

$$T^* = \bigcup_{n \geq 0} T^n \quad \text{and} \quad T^+ = \bigcup_{n > 0} T^n.$$

$w \in T^*$ is a *word* in T . If $|w| = 0$, then w is the *empty word*, denoted by ϵ .

Definition (Source Code)

A *source code* \mathcal{C} is a function $\mathcal{C} : S \rightarrow T^+$. That is, $w_i = \mathcal{C}(s_i) \in T^+$.

Note: When we do not care about S , we simply say \mathcal{C} is a finite set of words over T . That is, $\mathcal{C} \subset T^+$.

Note: We can trivially extend \mathcal{C} to $\mathcal{C}^* : S^* \rightarrow T^*$, so \mathcal{C} encodes a sequence of code symbols to a sequence of code words.

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition**
- Examples

Unique

Decodability

- Definition
- Block Codes
- Good Codes
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Examples: Building Codes is Nontrivial

Example (\mathcal{S} : Unbiased Die, \mathcal{C} : Binary Code)

To send the result of a die throw over a binary channel, since $q > r = 2$, we need to construct a binary code.

1) Try $\mathcal{C}(s_i) = (s_i)_2$, the binary representation of s_i . For instance,

$$\mathcal{C}(s_1) = \mathcal{C}(1) = w_1 = 1, \mathcal{C}(s_2) = 10, \dots, \mathcal{C}(s_6) = 110.$$

So, if $s = 53214$, then s is encoded $t = 10111101100$. *Problem:*

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples**

Unique

Decodability

- Definition
- Block Codes
- Good Codes
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Examples: Building Codes is Nontrivial

Example (\mathcal{S} : Unbiased Die, \mathcal{C} : Binary Code)

To send the result of a die throw over a binary channel, since $q > r = 2$, we need to construct a binary code.

1) Try $\mathcal{C}(s_i) = (s_i)_2$, the binary representation of s_i . For instance,

$$\mathcal{C}(s_1) = \mathcal{C}(1) = w_1 = 1, \mathcal{C}(s_2) = 10, \dots, \mathcal{C}(s_6) = 110.$$

So, if $s = 53214$, then s is encoded $t = 10111101100$. *Problem:*

$$11 = 1.1$$

So, $t = 10111101100 = 101.11.10.1.100 = 101.1.1.10.1.100$ could be decoded as originally being $s = 53214$ or $s' = 511214$.

Definition of Unique Decodability

Definition (Unique Decodability)

A code \mathcal{C} is *uniquely decodable* iff $\mathcal{C}^* : S^* \rightarrow T^*$ is a injection.

We say that codes that are not uniquely decodable are *ambiguous*.

Example (\mathcal{S} : Unbiased Die, \mathcal{C} : Binary Code)

Since $\mathcal{C}^*(3) = \mathcal{C}^*(11) = 11$, \mathcal{C} is ambiguous.

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples

Unique Decodability

- Definition**
- Block Codes
- Good Codes
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Decoding Block-Codes is Trivial

Example (\mathcal{S} : Unbiased Die, \mathcal{C} : Binary Code)

2) We can remedy this by making \mathcal{C} a *block code*^a.

$$\mathcal{C}(1) = 000, \dots, \mathcal{C}(6) = 101.$$

^aAll codewords have the same length

Theorem

If \mathcal{C} is a block code, then \mathcal{C} is uniquely decodable.

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples

Unique

Decodability

- Definition
- Block Codes**
- Good Codes
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Decoding Block-Codes is Trivial

Example (\mathcal{S} : Unbiased Die, \mathcal{C} : Binary Code)

2) We can remedy this by making \mathcal{C} a *block code*^a.

$$\mathcal{C}(1) = 000, \dots, \mathcal{C}(6) = 101.$$

^aAll codewords have the same length

Theorem

If \mathcal{C} is a block code, then \mathcal{C} is uniquely decodable.

Proof.

By contradiction. Let l be the common word-length. Let $\mathbf{t} \in \mathcal{C}^*$ and assume an ambiguous factorisation $\mathbf{t} = \mathbf{u}_1 \cdots \mathbf{u}_m = \mathbf{v}_1 \cdots \mathbf{v}_n$ with $\mathbf{u}_i, \mathbf{v}_j \in \mathcal{C}$, and $m \neq n$.

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples

Unique

Decodability

- Definition
- Block Codes**
- Good Codes
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Decoding Block-Codes is Trivial

Example (\mathcal{S} : Unbiased Die, \mathcal{C} : Binary Code)

2) We can remedy this by making \mathcal{C} a *block code*^a.

$$\mathcal{C}(1) = 000, \dots, \mathcal{C}(6) = 101.$$

^aAll codewords have the same length

Theorem

If \mathcal{C} is a block code, then \mathcal{C} is uniquely decodable.

Proof.

By contradiction. Let l be the common word-length. Let $\mathbf{t} \in \mathcal{C}^*$ and assume an ambiguous factorisation $\mathbf{t} = \mathbf{u}_1 \cdots \mathbf{u}_m = \mathbf{v}_1 \cdots \mathbf{v}_n$ with $\mathbf{u}_i, \mathbf{v}_j \in \mathcal{C}$, and $m \neq n$. Since $lm = |\mathbf{t}| = ln$, then $m = n$ must be true. Since \mathcal{C} is a block code and $\mathbf{u}_i, \mathbf{v}_j \in \mathcal{C}$, then \mathbf{u}_1 and \mathbf{v}_1 must consist of the first l symbols of \mathbf{t} , so $\mathbf{u}_1 = \mathbf{v}_1$ must hold. Likewise for $\mathbf{u}_i = \mathbf{v}_i, \forall i$. A contradiction. As \mathbf{t} was arbitrary, any $\mathbf{t} \in \mathcal{C}^*$ must be uniquely factorisable. □

Lecture 1: Unique Decodability

Source

Preliminaries
Probability
Theory
Notation
Definition
Assumptions
Examples

Source Code

Motivation
Definition
Examples

Unique

Decodability

Definition
Block Codes
Good Codes
Sardinas-
Patterson

Epilogue

Summary
Next Session

What makes a code good?

Definition (Average Word-length)

For each $w_i \in \mathcal{C}$, let $l_i = |w_i|$. The *average word-length* of \mathcal{C} is

$$L(\mathcal{C}) = \sum_{i=1}^q p_i l_i,$$

with $p_i = p(s_i)$ where $C(s_i) = w_i$.^a

^a $s_i \in \mathcal{S}$, where \mathcal{C} is the source code of \mathcal{S} .

We are interested in codes which

- a) are easy and unambiguous to decode, and
- b) have a small/minimal $L(\mathcal{C})$.

We now look at (a).

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples

Unique

Decodability

- Definition
- Block Codes
- Good Codes**
- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Aren't Block Codes Good Enough?

Example (\mathcal{S} : Biased Die, \mathcal{C} : Binary Code)

Let us make our die biased. $p_1 = \dots = p_5 = \frac{1}{10}$ and $p_6 = \frac{1}{2}$. Our block code for \mathcal{S} has average word length 3.

$$L(\mathcal{C}) = \sum_{i=1}^6 3 * p_i = \frac{1}{2} + \frac{5}{10} = 3.$$

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples

Unique

Decodability

- Definition
- Block Codes
- Good Codes**

- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Aren't Block Codes Good Enough?

Example (\mathcal{S} : Biased Die, \mathcal{C} : Binary Code)

Let us make our die biased. $p_1 = \dots = p_5 = \frac{1}{10}$ and $p_6 = \frac{1}{2}$. Our block code for \mathcal{S} has average word length 3.

$$L(\mathcal{C}) = \sum_{i=1}^6 3 * p_i = \frac{1}{2} + \frac{5}{10} = 3.$$

3) We can do better. Define \mathcal{C} as the mapping

$$s_1 \mapsto w_1 = 0, s_2 \mapsto 01, \dots, s_6 \mapsto 01111, s_6 \mapsto 11111.$$

here,

$$L(\mathcal{C}) = \frac{1}{2} + \sum_{i=2}^5 \frac{i}{10} + \frac{5}{10} = 2 + \frac{2}{5}.$$

Extra: Compare the average word lengths for an unbiased die using these two codes.

Lecture 1: Unique Decodability

Source

- Preliminaries
- Probability Theory
- Notation
- Definition
- Assumptions
- Examples

Source Code

- Motivation
- Definition
- Examples

Unique

Decodability

- Definition
- Block Codes
- Good Codes**

- Sardinas-Patterson

Epilogue

- Summary
- Next Session

Back on Topic: Check Unique Decodability

Now we illustrate a technique to prove whether *any* code is *uniquely decodable*. First,

Definition (\mathcal{C}_n)

Let $\mathcal{C}_0 = \mathcal{C}_1$, and for $n \geq 1$,

$$\mathcal{C}_n = \left\{ w \in T^+ : \begin{array}{l} \exists u, v. \left(\begin{array}{l} u \in \mathcal{C}, v \in \mathcal{C}_{n-1} \text{ or} \\ v \in \mathcal{C}, u \in \mathcal{C}_{n-1} \end{array} \right) \\ uw = v \end{array} \text{ and } \right\}$$

Explain: You have

$$w \in T^*.$$

If you can pick $u \in \mathcal{C}$ s.t.

$$uw \in \mathcal{C}_{n-1},$$

or pick $u \in \mathcal{C}_{n-1}$ s.t.

$$uw \in \mathcal{C},$$

then

$$w \in \mathcal{C}_n.$$

Lecture 1: Unique Decodability

Source

Preliminaries
Probability
Theory
Notation
Definition
Assumptions
Examples

Source Code

Motivation
Definition
Examples

Unique

Decodability

Definition
Block Codes
Good Codes

Sardinas-Patterson

Epilogue

Summary
Next Session

In other words, if you can *postfix* a

- $u \in \mathcal{C}$ with $w \in T^+$ to form $uw \in \mathcal{C}_{n-1}$, or a
- $u \in \mathcal{C}_{n-1}$ with $w \in T^+$ to form $uw \in \mathcal{C}$,

then *the postfix*, w , is in \mathcal{C}_n . **Note:**

$$\mathcal{C}_1 = \{w \in T^+ : \exists u, v \in \mathcal{C} . uw = v\}.$$

$$\mathcal{C}_n = \emptyset \implies \mathcal{C}_{n'} = \emptyset, \forall n' > n.$$

$$\mathcal{C}_\infty = \bigcup_{n=1}^{\infty} \mathcal{C}_n.$$

What is the point?

Theorem (Sardinas-Patterson)

A code \mathcal{C} is uniquely decodable iff $\mathcal{C} \cap \mathcal{C}_\infty = \emptyset$.

Lecture 1: Unique Decodability

Source

Preliminaries

Probability

Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-Patterson

Epilogue

Summary

Next Session

Example (The better die code is uniquely decodable)

Recall

$$\mathcal{C} = \{0, 01, 011, 0111, 01111, 11111\}.$$

Then

$$\mathcal{C}_0 = \mathcal{C}.$$

$$\mathcal{C}_1 = \{1, 11, 111, 1111\}$$

$$\mathcal{C}_2 = \{1, 11, 111, 1111\}$$

...

Example (Exercise 1.2)

If $\mathcal{C} = \{02, 12, 120, 20, 21\}$, then

$$\mathcal{C}_1 = \{0\}$$

$$\mathcal{C}_2 = \{2\}$$

$$\mathcal{C}_3 = \{0, 1\}$$

$$\mathcal{C}_4 = \{2, 20\}$$

$$\mathcal{C}_5 = \{0, 1\}$$

...

Lecture 1: Unique Decodability

Source

Preliminaries

Probability

Theory

Notation

Definition

Assumptions

Examples

Source Code

Motivation

Definition

Examples

Unique

Decodability

Definition

Block Codes

Good Codes

Sardinas-**Patterson**

Epilogue

Summary

Next Session

Creating \mathcal{C}_∞ takes finitely many steps

- Why do I stop when I hit a period?
- Will this process always hit a period, or \emptyset ?

Exercise (1.1)

- Each \mathcal{C}_n is finite.
- $\mathcal{C}_0, \mathcal{C}_1, \dots$ is eventually periodic.

Lecture 1: Unique Decodability

Source

Preliminaries
Probability
Theory
Notation
Definition
Assumptions
Examples

Source Code

Motivation
Definition
Examples

Unique

Decodability

Definition
Block Codes
Good Codes

Sardinas-Patterson

Epilogue

Summary
Next Session

Creating \mathcal{C}_∞ takes finitely many steps

- Why do I stop when I hit a period?
- Will this process always hit a period, or \emptyset ?

Exercise (1.1)

- Each \mathcal{C}_n is finite.
 - $\mathcal{C}_0, \mathcal{C}_1, \dots$ is eventually periodic.
- a Assume \mathcal{C} has maximum code word length l . Let $w \in \mathcal{C}_n$ for some n . Then $|w| \leq l$ since w is always a postfix to a word in \mathcal{C}^a . There are finitely many possible words of length l ; these are the set $T_1^l = \bigcup_{i=1}^l T^i$. Since T_1^l is finite, then so is $\mathcal{P}(T_1^l)$. Any \mathcal{C}_n is an element of $\mathcal{P}(T_1^l)$, containing finite sets.

Lecture 1: Unique Decodability

Source

Preliminaries
Probability
Theory
Notation
Definition
Assumptions
Examples

Source Code

Motivation
Definition
Examples

Unique

Decodability
Definition
Block Codes
Good Codes
Sardinas-Patterson

Epilogue

Summary
Next Session

Creating \mathcal{C}_∞ takes finitely many steps

- Why do I stop when I hit a period?
- Will this process always hit a period, or \emptyset ?

Exercise (1.1)

- Each \mathcal{C}_n is finite.
 - $\mathcal{C}_0, \mathcal{C}_1, \dots$ is eventually periodic.
- Assume \mathcal{C} has maximum code word length l . Let $w \in \mathcal{C}_n$ for some n . Then $|w| \leq l$ since w is always a postfix to a word in \mathcal{C}^a . There are finitely many possible words of length l ; these are the set $T_1^l = \bigcup_{i=1}^l T^i$. Since T_1^l is finite, then so is $\mathcal{P}(T_1^l)$. Any \mathcal{C}_n is an element of $\mathcal{P}(T_1^l)$, containing finite sets.
 - There can at most be a sequence of length $|\mathcal{P}(T_1^l)|$ before \mathcal{C}_n repeats itself. Since each \mathcal{C}_{n+1} is uniquely determined by its predecessor, if ever $\mathcal{C}_n = \mathcal{C}_{n'}$ for $n \neq n'$, then $\mathcal{C}_{n+i} = \mathcal{C}_{n'+i}, \forall i$.

^aPerhaps a postfix to a postfix, or a postfix to a postfix to ...

Thus, at some point, inducing \mathcal{C}_n contributes nothing new to \mathcal{C}_∞ .

Our Good Codes

Definition (Unique Decodability (Simplified))

Assuming $w_i \neq w_j$ for $i \neq j$, a code \mathcal{C} is *uniquely decodable* iff any code word sequence \mathbf{w} is uniquely factorisable:

$$v_1 \cdots v_j = w_1 \cdots w_k \implies k = j \wedge v_i = w_i, \forall i.$$

Definition (Average Word-length)

For each $w_i \in \mathcal{C}$ w. $|\mathcal{C}| = q$, let $l_i = |w_i|$. The *average word-length* of \mathcal{C} is

$$L(\mathcal{C}) = \sum_{i=1}^q p_i l_i,$$

We are interested in codes which

- a) are easy and unambiguous to decode, and
- b) have a small/minimal $L(\mathcal{C})$.

Lecture 2: Optimal Codes

Recap

Definitions

Motivation

Instantaneous
Code

Definition

Instantaneous iff
Prefix

Construction

Inequalities

Kraft

McMillan

Sum Up

Optimal Codes

Do We Stop
Here?Definition and
Existence

Huffman

Huffman codes
are Instantaneous
(binary)

Epilogue

The Best Source
Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

1) Mapping to binary representation:

- $\mathcal{C} = (110, 101, 100, 11, 10, 1)$

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

1) Mapping to binary representation:

- $\mathcal{C} = (110, 101, 100, 11, 10, 1)$

ambiguous (ex: $21 = 5$).

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

1) Mapping to binary representation:

- $\mathcal{C} = (110, 101, 100, 11, 10, 1)$

ambiguous (ex: $21 = 5$).

2) Block code:

- $\mathcal{C} = (110, 101, 100, 011, 010, 001)$

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

1) Mapping to binary representation:

- $\mathcal{C} = (110, 101, 100, 11, 10, 1)$

ambiguous (ex: $21 = 5$).

2) Block code:

- $\mathcal{C} = (110, 101, 100, 011, 010, 001)$

sub-optimal. $L(\mathcal{C}) = 3$.

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

1) Mapping to binary representation:

- $\mathcal{C} = (110, 101, 100, 11, 10, 1)$

ambiguous (ex: $21 = 5$).

2) Block code:

- $\mathcal{C} = (110, 101, 100, 011, 010, 001)$

sub-optimal. $L(\mathcal{C}) = 3$.

3) Clever encoding; $L(\mathcal{C}) = 2 + \frac{2}{5}$:

- $\mathcal{C} = (0, 01, 011, 0111, 01111, 11111)$

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

1) Mapping to binary representation:

- $\mathcal{C} = (110, 101, 100, 11, 10, 1)$

ambiguous (ex: $21 = 5$).

2) Block code:

- $\mathcal{C} = (110, 101, 100, 011, 010, 001)$

sub-optimal. $L(\mathcal{C}) = 3$.

3) Clever encoding; $L(\mathcal{C}) = 2 + \frac{2}{5}$:

- $\mathcal{C} = (0, 01, 011, 0111, 01111, 11111)$

unbounded decoding delay.

Lecture 2: Optimal Codes

Recap
Definitions
Motivation

Instantaneous
Code

Definition
Instantaneous iff
Prefix
Construction

Inequalities
Kraft
McMillan
Sum Up

Optimal Codes
Do We Stop
Here?
Definition and
Existence
Huffman
Huffman codes
are Instantaneous
(binary)

Epilogue
The Best Source
Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

- 1) Mapping to binary representation: *ambiguous* (ex: 21 = 5).
 • $\mathcal{C} = (110, 101, 100, 11, 10, 1)$
- 2) Block code: *sub-optimal*. $L(\mathcal{C}) = 3$.
 • $\mathcal{C} = (110, 101, 100, 011, 010, 001)$
- 3) Clever encoding; $L(\mathcal{C}) = 2 + \frac{2}{5}$: *unbounded decoding delay*.
 • $\mathcal{C} = (0, 01, 011, 0111, 01111, 11111)$
- 4) The fix; $L(\mathcal{C}) = 2 + \frac{3}{5}$:
 • $\mathcal{C} = (0, 01, 011, 0111, 01111, \underline{0}11111)$

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

Examples: The Biased Die, binary code

Assume index-mapping of source symbols to probabilities & code words.

Example

Source could produce the symbols (6, 5, 4, 3, 2, 1) w. probability distribution $(\frac{1}{2}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10})$.

- 1) Mapping to binary representation: *ambiguous* (ex: 21 = 5).
 • $\mathcal{C} = (110, 101, 100, 11, 10, 1)$
- 2) Block code: *sub-optimal*. $L(\mathcal{C}) = 3$.
 • $\mathcal{C} = (110, 101, 100, 011, 010, 001)$
- 3) Clever encoding; $L(\mathcal{C}) = 2 + \frac{2}{5}$: *unbounded decoding delay*.
 • $\mathcal{C} = (0, 01, 011, 0111, 01111, 11111)$
- 4) The fix; $L(\mathcal{C}) = 2 + \frac{3}{5}$: *bounded decoding delay: 1*.
 • $\mathcal{C} = (0, 01, 011, 0111, 01111, \underline{0}11111)$

Lecture 2: Optimal Codes

Recap
Definitions
Motivation
Instantaneous Code
Definition
Instantaneous iff Prefix
Construction
Inequalities
Kraft
McMillan
Sum Up
Optimal Codes
Do We Stop Here?
Definition and Existence
Huffman
Huffman codes are Instantaneous (binary)
Epilogue
The Best Source Codes

What we have seen

- Source coding is not trivial.



Figure: Classes of Source Codes.

Lecture 2: Optimal Codes

Recap

Definitions

Motivation

Instantaneous
Code

Definition

Instantaneous iff
Prefix

Construction

Inequalities

Kraft

McMillan

Sum Up

Optimal Codes

Do We Stop
Here?

Definition and
Existence

Huffman

Huffman codes
are Instantaneous
(binary)

Epilogue

The Best Source
Codes

What we have seen

- Source coding is not trivial.

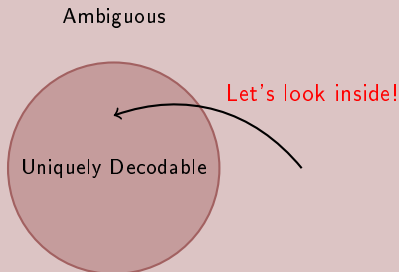


Figure: Classes of Source Codes.

Lecture 2: Optimal Codes

Recap

Definitions

Motivation

Instantaneous
Code

Definition

Instantaneous iff
Prefix

Construction

Inequalities

Kraft

McMillan

Sum Up

Optimal Codes

Do We Stop
Here?

Definition and
Existence

Huffman

Huffman codes
are Instantaneous
(binary)

Epilogue

The Best Source
Codes