	00	000	0000	
		000	0001	
		001	0010	ß
0		001	0011	nd
0		0.1.0	0100	D
	01	010	0101	pde
		011	0110	C C
		011	0111	
	10	100	1000	m
		100	1001	Sy
	10	101	1010	tal
1		101	1011	to
	11	110	1100	he
		110	1101	E E
	•••	111	1110	
			1111	

Figure 5.1. The symbol coding budget. The 'cost' 2^{-l} of each codeword (with length l) is indicated by the size of the box it is written in. The total budget available when making a uniquely decodeable code is 1. You can think of this diagram as showing a codeword supermarket, with the codewords arranged in aisles by their length, and the cost of each codeword indicated by the size of its box on the shelf. If the cost of the codewords that you take exceeds the budget then your code will not be uniquely decodeable.

Book by David MacKay

Krait inequality and code budgets

	a_i	$c(a_i)$	l_i				C_3 :					C	\overline{C}				C_6 :		
		1000	4		a_i	$c(a_i)$	p_i	$h(p_i)$	l_i	_		C_4	C_5		a_i	$c(a_i)$	p_i	$h(p_i)$	l_i
C_0 :	h	0100	Λ		a	0	1/2	1.0	1		a	00	0		a	0	$1/_{2}$	1.0	1
0.	D	0100			b	10	$1/_{4}$	2.0	2		b	01	1		h	01	$1/_{A}$	$\frac{1.0}{2.0}$	2
	С	0010	4		С	110	$1/_{8}$	3.0	3		С	10	00		0	011	1/0	2.0	2 9
	d	0001	4		d	111	$1/_{8}$	3.0	3		4	- • 1 1	11		C	110	-/8 1/0	3.0	ე ე
						111	70	0.0	0			ΤΤ				111	1/8	3.0	3
			C_0					C_3					\mathcal{K}				C	7 6	
				000	0			000) –	0000			000	0000				000	0000
		00		001 001	0		00	001	l –	0010		00	001	0010			00	001 -	0010
		0		010 010	0	0	01	010) –	0100 0101	0	01	010	0100 0101		0	01	010	0100 0101
		01		011 011 011	0			011	L	0110 0111			011	0110 0111			.01	011	0110 0111
		10		100 100	0		10	100) -	1000 1001		10	100	1000 1001			10	100 -	1000 1001
]	1		101 101	0	1	10	101		1010 1011			101	1010 1011		1	10	101 -	1010 1011
		1		110 110	0		11	110		1100 1101		11	110	1100 1101		1	11	110 -	1100 1101
				111 111	0			111	-	1110 1111			111	1110 1111				111 -	1110 1111

Book by David MacKay

ECE 534 by Natasha Devroye

a_i	p_i	$\log_2 \frac{1}{p_i}$	l_i	$c(a_i)$
a	0.0575	4.1	4	0000
b	0.0128	6.3	6	001000
С	0.0263	5.2	5	00101
d	0.0285	5.1	5	10000
е	0.0913	3.5	4	1100
f	0.0173	5.9	6	111000
g	0.0133	6.2	6	001001
h	0.0313	5.0	5	10001
i	0.0599	4.1	4	1001
j	0.0006	10.7	10	1101000000
k	0.0084	6.9	7	1010000
1	0.0335	4.9	5	11101
m	0.0235	5.4	6	110101
n	0.0596	4.1	4	0001
0	0.0689	3.9	4	1011
р	0.0192	5.7	6	111001
q	0.0008	10.3	9	110100001
r	0.0508	4.3	5	11011
S	0.0567	4.1	4	0011
t	0.0706	3.8	4	1111
u	0.0334	4.9	5	10101
v	0.0069	7.2	8	11010001
W	0.0119	6.4	7	1101001
х	0.0073	7.1	7	1010001
у	0.0164	5.9	6	101001
Z	0.0007	10.4	10	1101000001
_	0.1928	2.4	2	01



Figure 5.6. Huffman code for the English language ensemble (monogram statistics).

Book by David MacKay

ure 5.6. This code has an expected length of 4.15 bits; the entropy of the ensemble is 4.11 bits. Observe the disparities between the assigned codelengths and the ideal codelengths $\log_2 1/p_i$.

			- 00000	Context (sequence thus far)	e thus far) Probability of next symbol				
			= 00001 0000		P(a) = 0.425	P(b) = 0.425	$P(\Box) = 0.15$		
			= 00010 = 00011 0001	b	P(a b) = 0.28	$P(b \mid b) = 0.57$	$P(\Box \mid \mathbf{b}) = 0.15$		
			= 000011 00 = 00100 0010 00	bb	$P(\mathbf{a} \mid \mathbf{bb}) = 0.21$	$P(\mathbf{b} \mid \mathbf{b}\mathbf{b}) = 0.64$	$P(\Box \texttt{bb}) {=} 0.15$		
			= 00101 0010	bbb	$P(\mathbf{a} \mid \mathbf{bbb}) = 0.17$	$P(b bbb) {=} 0.68$	$P(\Box \texttt{bbb}) {=} 0.15$		
а			$\frac{-00110}{-00111}$ 0011	bbba	$P(\mathbf{a} \mathbf{bbba}) \!=\! 0.28$	$P(\mathbf{b} \mathbf{bbba}) {=} 0.57$	$P(\Box \texttt{bbba}) {=} 0.15$		
			$ \begin{array}{r} - 01000 \\ - 01001 \\ - 01001 \\ - 01010 \\ - 01011 \\ - 01011 \\ - 01100 \\ - 01101 \\ - 01100 \\ - 01101 \\ - 01100 \\ - 0100 \\ - 000 \\ -$	Figure 6.4 shows th 0.425 of $[0, 1)$. The i	e corresponding internation internation of the middle state of the	vals. The interval b le 0.567 of b, and so	is the middle forth.		
	ba		$ \begin{array}{c} - 01101 \\ - 01110 \\ - 01111 \\ - 01111 \\ - 10000 \\ - 10001 \\ \end{array} 011 $	hhhaa	<u>- 10010111</u> - <u>- 100110</u> 00 - 10011001				
- b	bb	bba <u>bbba</u>	$= \frac{10001}{1001} 100 / \\= \frac{10010}{1001} 1001 / \\= \frac{10100}{1010} 1010 / \\= 10110 101 / \\= 10110 101 / \\= 10110 / \\= 10110 / \\= 10110 / \\= 10110 / \\= 10110 / \\= 10110 / \\= 101 / \\= 10110 / \\= 101 / \\$	bbbaa bbbab bbbaD	$ \begin{array}{r} -\frac{-10011010}{-10011010} \\ -10011011 \\ -10011100 \\ -10011101 \\ -10011110 \\ -10011110 \\ \end{array} $	11			
		bbb bb	$\begin{array}{r} - & 10110 \\ - & 10111 \\ - & 10111 \\ - & 11000 \\ - & 11001 \\ \end{array} \begin{array}{r} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 $		- <u>10011111</u> - <u>10100000</u> 100111101				
b	b□		$-\frac{11001}{-11010} 110$ $-\frac{11001}{-11011} 1101$ $-\frac{11100}{-11100} 1110$						
			$\frac{-11101}{-11110} 1111$		Figure 6 arithme sequence	5.4. Illustration etic coding process $coding$ is trace	n of the cess as the .nsmitted.		

Book by David MacKay