8 CODES IN EVERYDAY USE

After studying this chapter you should

* appreciate the role of codes in a highly technological society;
* understand how and why check digits are used;
* understand why particular designs are used for particular codes.

8.0   Introduction

Although you might not have appreciated it, many aspects of life today depend on the effective use of codes. Examples include

> **Satellite transmission**
>
> **Bar codes**
>
> **Postcodes**
>
> **Catalogue codes**
>
> **Bank codes**
>
> **Computer codes**

- the list could go on and on. Whilst the average member of the public does **not** need to know how these codes are designed or how they work, it has become a very important subject for mathematicians to study. In this chapter you will look at a number of codes used in practical situations.

8.1   Historical perspective

Although codes have now become indispensible to modern life, they are not a new invention, and our study will start with two codes which have been around for some time.

Braille

Braille is a method of writing that can be used by blind people. It was invented in 1829 by the Frenchman, *Louis Braille* (1809-52). When he was three years old he lost the sight of one eye while playing with one of his father's knives (his father was a harness maker), and soon lost his sight completely.

Online Classes : Megalecture@gmail.com
www.youtube.com/megalecture
www.megalecture.com

Chapter 8 Codes in Everyday Use

An earlier system for soldiers passing messages in the dark had been developed by another Frenchman, *Charles Barbier*; this used up to **twelve** embossed dots, 6 vertical in 2 rows, as shown opposite. Each letter is made up of a pattern of raised dots which the reader can feel with his fingers. Of course, it is just as important to be able to tell when a dot is missing.

Braille revised the pattern by using a base of **six** positions, 3 vertical in 2 rows, as shown opposite.

How many different patterns exist using this system?

Activity 1

Investigate how many different patterns exist using just

(a)   1 dot   (b) 2 dots   (c)   3 dots

(d)   4 dots   (e) 5 dots   (f)   6 dots.

Check the final answer with your answer to the earlier discussion point.

The chart in Appendix 1 gives the list for the alphabet, number and punctuation. Study the chart carefully and then proceed to the next activity.

Activity 2

(a)   What patterns have not been used in the given Braille chart?

(b)   Can you suggest what these other patterns can be used for?

(c)   Consider systems that use four or five dots as the basis rather than six. What are the advantages or disadvantages of such systems?

Morse code

This was designed in America by *Samuel Morse*, 1791 -1872, and was first used in 1844 for the telegraph line between Baltimore and Washington. Although modern technology has largely superseded the need for morse code as a form of communication, the 'SOS' code is still universally used for shipping in distress.

Online Classes : Megalecture@gmail.com
www.youtube.com/megalecture
www.megalecture.com

Activity 3

Find out the actual codes used in Morse Code. Analyse why it takes its particular form, and suggest improvements.

There are many other important historical codes, including secret codes used in the World Wars. It is the view of some that the eventual cracking of the ENIGMA code, used by the Germans in the Second World War, by the team at Bletchley, was one of the most significant factors in helping the Allies to defeat Germany. This chapter, however, will deal with codes in everyday use.

8.2   Check digits

Many codes have been designed for use with new technology. These include bar codes, ISBN numbers, ASCII codes, post codes, bank account numbers; many of these modern codes employ a checking device, often referred to as a **check digit**. An example of this is that of ISBN numbers, now used universally on all new books. Each ISBN has **ten** digits made up from components, as illustrated opposite.

The check digit is designed so that any **one** error in the previous nine digits is spotted. It is calculated in the following way.

Group identifier
(up to 4 digits long) which gives information about contry of publication (UK uses '0' or '1')

Check digit

↓   ↓
0   8 5 0 2 0   0 1 4   8
↑   ↑

Publisher's prefix
(one to seven digits long) uniquely identifies the publisher

Title number
identifies book number in publisher's list

ISBN number

> Multiply the first nine numbers by 10, 9, 8, ..., 2 respectively and find the sum of the resulting numbers. The check number is the smallest number that needs to be added to this total so that it is exactly divisible by 11.
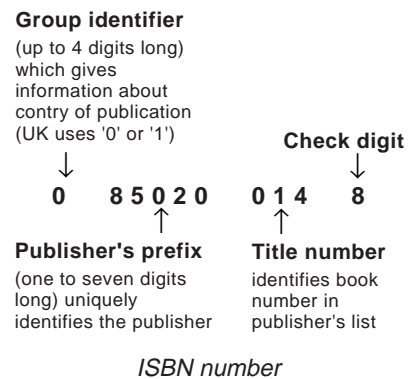
For the example above, we have

$$0\times10+\left(8\times9+5\times8+0\times7+2\times6+0\times5\right)+\left(0\times4+1\times3+4\times2\right)$$

$$=135$$

so the check digit must be 8, since 143 is divisible by 11. Note that if the number 10 is needed for the check digit, the symbol X is used.

Example

Determine the check digit, *a*, for the following ISBN numbers:

(a)   1 869931 00 *a*   (b)   1 7135 2272 *a*

Online Classes : Megalecture@gmail.com
www.youtube.com/megalecture
www.megalecture.com

Solution

(a) The number

$$1 \times 10 + (8 \times 9 + 6 \times 8 + 9 \times 7 + 9 \times 6 + 3 \times 5 + 1 \times 4) + (0 \times 3 + 0 \times 2 + 1 \times a)$$

must be divisible by 11; i.e. $266 + a$ must be divisible by 11 and $0 \le a \le 10$. Hence $a = 9$.

(b) Again, the number

$$1 \times 10 + (7 \times 9 + 1 \times 8 + 3 \times 7 + 5 \times 6) + (2 \times 5 + 2 \times 4 + 7 \times 3 + 2 \times 2 + 1 \times a)$$

must be divisible by 11; i.e. 11 divides $(175 + a)$, giving $a = 1$.

Why do ISBN numbers use a check digit of this particular form?

Activity 4   Error detection and correction

There is one error in each of these ISBN numbers. Can you correct them?

(a)   1 869932 23 8   (b)   0 7458 1078 5

Activity 5

A publisher is given a set of ISBN numbers of the form

$$1 \quad 834721 \quad m \; n \; x$$

for $0 \le m \le 9, 0 \le n \le 9$, and x is the check digit. Design a ready reckoner or algorithm to determine the check digit for all appropriate values of m and n.

8.3   Bar code design

Bar codes are nearly universal today, being used in just about every industry. They were first suggested for automation in grocery stores in 1932 in the thesis of a Harvard Business School student, but it was not until the 1950s that the idea of a scanner installed at check-outs was conceived. It took another two decades for a combination of technology advancement and economic pressure to bring about the commercial use of bar codes and optical readers in retail trading. In 1973 the **UPC** (Universal Product Code) was adopted as a standard.   In 1976 a variation known as the **EAN** (European Article Numbers) was also standardised.

Online Classes : Megalecture@gmail.com
www.youtube.com/megalecture
www.megalecture.com

Chapter 8 Codes in Everyday Use

Other types of bar codes, for example, **Code 3 of 9** and
Interleaved Two of Five (ITF), have also been developed.

In the UK, the **Article Number Association** was formed to
administer and promote the use of article numbering, and the
association provides information packs and educational material.

8-Digit EAN

Three examples of 8-digit EAN symbols are shown opposite.
These are used by large stores for their own brands. Each bar code
consists of

> **left hand guard**
>
> **left hand four numbers**
>
> **centre guard**
>
> **right hand four numbers**
>
> **right hand guard.**

Looking at the code for each number, you will notice that the
representation of a number is dependent on whether it is on the left
or right hand side. In fact, each representation is designed using a
seven module system. For example, a left hand side 5 is shown
magnified opposite (the dashes are shown here to emphasise the
seven module design - they are not actually shown on the code).

Each number has two white and two black strips of varying
thickness but following the rules that

(a)   the first module must be white;

(b)   the last module must be black;

(c)   there are in total either 3 or 5 black modules.
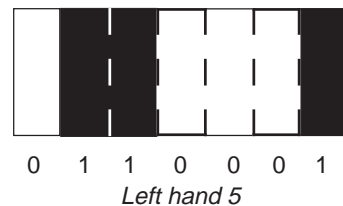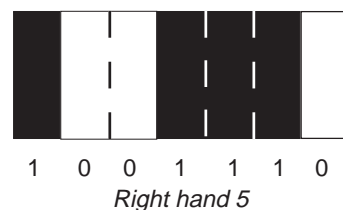
A convenient way of representing each number is given by using 0
(white) 1 (black) giving 0 1 1 0 0 0 1 for 5, as shown.

0 1 1 0 0 0 1
Left hand 5

Activity 6   Left hand codes

With the rules listed above, write down all the possible codes for
left hand numbers.

Appendix 2 gives the complete set of codes for left hand numbers -
called **Number Set A**. The codes for the right hand side are
determined by interchanging 0 s and 1 s (i.e. white and black
interchanged) - called **Number Set C**.

Why is a different code needed for right hand numbers?

1 0 0 1 1 1 0
Right hand 5

Online Classes : Megalecture@gmail.com
www.youtube.com/megalecture
www.megalecture.com

Chapter 8 Codes in Everyday Use

As with ISBN numbers, these bar codes incorporate a check digit, again the last one. It is chosen so that,

$$3 \times (1\text{st} + 3\text{rd} + 5\text{th} + 7\text{th number}) + (2\text{nd} + 4\text{th} + 6\text{th} + 8\text{th number})$$

is exactly divisible by 10. For example, for

0 0 3 3   7 7 9 3

it means that

$$3 \times (0 + 3 + 7 + 9) + (0 + 3 + 7 + 3) = 3 \times 19 + 13 = 70$$

is exactly divisible by 10.

Example

Find the check digit, *a*, for the 8-digit EAN code

5 0 2 1   4 2 1 *a*

Solution

The number

$$3 \times (5 + 2 + 4 + 1) + (0 + 1 + 2 + a) = 36 + 3 + a = 39 + a$$

must be exactly divisible by 10, so $a = 1$.

Activity 7   Errors

The 8-digit EAN code

5 0 2 6   8 0 2 0

has one error. Can you identify it?

What are the advantages of this method of determining the check digit?

13-Digit EAN

Examples of this code are found on many grocery products. Three such codes are shown opposite.

The first digit, which as you can see is not represented directly in the code, together with the second digit, indicates the country in which the article number was allocated; e.g. 50 represents the UK, 31 represents France, etc. The next five digits are issued to a particular manufacturer, and the next five identify the product. The final number is again the check digit.

9 770951 862057

7 610400 013130

5 010031 661038

All six right hand numbers are coded using **Number Set C** but the six left hand numbers are coded using a combination of **Number Sets A** and **B** (see Appendix 3) according to the first digit. For example, if the first digit is 5, then the next six digits are coded according to the Number Sets **A B B A A B**..

Using the tables in Appendix 3, can you see how Number Set B is obtained?

Activity 8

Using three As and three Bs, how many different possible combinations exist for the coding of the six left hand numbers in the code?

In fact, the first digit 0 uses the code **A A A A A A**, whereas all other first digits are coded using 3 As and 3 Bs as indicated in Appendix 4.

13-digit EAN codes use the same method as 8-digit EAN codes for determing the check digit, except that all 13 numbers are included, so that the number

$$3 \times (2\text{nd} + 4\text{th} + \ldots + 12\text{th number}) + (1\text{st} + 3\text{rd} + \ldots + 13\text{th number})$$

must be divisible by 10.

Activity 9

Check that the three 13-digit EAN codes shown earlier have correct check digits.

There are many other types of bar codes in use, some having a completely different design (e.g. library cards).

Exercise 8A

1. Design a method of coding for alphanumeric (i.e. number and letter) characters used for display on calculators.

2. Find out the code used for semaphore. Is it an efficient method of coding?

3. Marks and Spencer, who only stock their own label brands, use a special 7-digit bar code. Find out what method is used for the check digit.

8.4   Postcodes

Much of the mail in the UK is now sorted automatically.   This has been made possible by the introduction of POSTCODES, which were started in 1966 and are now used throughout the UK.

Post boxes — Sorting Office Town A — Transport — Sorting Office Town B — Postal rounds
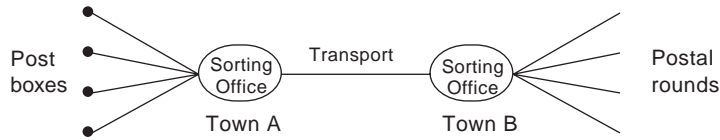
After collection, letters are sorted at the local Sorting Office into **areas** and **districts**. They are then forwarded to the appropriate Sorting Office where they are sorted again into **sectors** and **units.**

The postcode shown opposite illustrates these aspects.

Why is a mixture of numbers and letters used?

Area	District	Sector	Unit
E X	1 3	1	P F
↑	↑	↑	↑
any letters	any one or two digits	any digit 1 to 9	any letters

Postcode

Activity 10

Keeping in mind the restrictions indicated in the postcode diagram, estimate the maximum number of units which can be defined.

In fact,there are

   12 areas, 2900 districts, 9000 sectors and 2 000 000 units.

Since there are about 24 million household and business addresses in the UK, the average number of addresses per unit is given by

$$\frac{24 \times 10^6}{2 \times 10^6} \;=\; 12.$$

Why do you think the Post Office does not identify each address with a unique postcode?

Finally, it should also be noted that each postcode has to be coded (with a series of small blue dots) on the envelope to enable the automatic sorting to take place. So yet another code is used in order to make use of the first code!

Activity 11

Design a coding system, which can be put on envelopes to represent postcodes using a series of dots, to facilitate automatic sorting.

8.5   Telephone numbers

Until 1995, most UK telephone numbers took the form of 10 digits as shown opposite.

The first digit was always 0, and the first digit of both the **area code** and the **local number** did not use 0 or 1.

0	**392**	**217113**
↑	↑	↑
fixed	area code	local number

Telephone number

Activity 12

With the restrictions given above, how many unique telephone numbers existed?

In 1994 there were about 25 million numbers in use in the UK but British Telecom was in fact running out of usable numbers.

Can you suggest why?

In April 1995, to solve the problem of lack of codes, BT adopted a new system of area codes.

Most local numbers did **not** change, but all area codes had a '1' inserted after the initial '0'. For example:

0392	became	01392
0742	became	01742
071	became	0171
081	became	0181
etc.		

What advantage did this new system have?

Activity 13

List the possible disadvantages of the new system implemented in 1995. Consider other solutions to the problem, giving the advantages and disadvantages.

There are numerous other codes used extensively; for example

> **Vehicle registration numbers**
>
> **Home shopping catalogue numbers**
>
> **ASCII codes in computing**
>
> **Mariner 9 code**
>
> **Cyphers**

all of which have been designed to solve particular problems.

8.6   Computing codes

There are many codes used in computing, but the most commonly used code is **ASCII** (American Standard Code for Information Interchange). The code is summarised opposite. It is in ascending binary order in each section.

Activity 14

How many possible codewords are there, using the ASCII system?

This code is not particularly efficient and for computers with limited memory space (e.g. hand-held calculators) often different codes are used.

One particular code in which the use of particular letters or numbers is very varied is called a **Huffman Code**.

As an example, consider a code needed for just five letters, say,

E   A   M   N   T

in which they are listed in order of decreasing frequency; that is, E is used more than A, A more than M, etc.

A possible Huffman code for five letters is shown below. The code for each letter is found by using a '1' for a left hand branch, and '0' for a right hand branch. So E is coded as '1', A as '0 0', etc., as shown.

Letter	Code
E	1
A	0 0
M	0 1 0
N	0 1 1 0
T	0 1 1 1

Why is this an efficient way of coding for this problem?

Note that there is no need to put gaps between codes for different letters as there can be no confusion, as you will see in the next example.
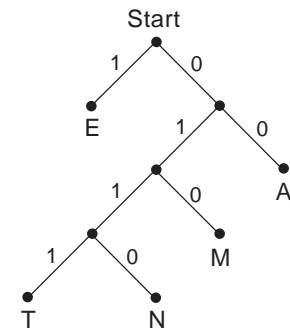
Example

Decode   0 1 1 0 0 0 0 1 0 1 0 1 0 0 0 0 1 1 0.

Character	Code	
Space	0 1 0	0 0 0 0
0	0 1 1	0 0 0 0
1	0 1 1	0 0 0 1
2	0 1 1	0 0 1 0
3	0 1 1	0 0 1 1
...	...	...
...	...	...
9	0 1 1	1 0 0 1
+	0 1 0	1 0 1 1
−	0 1 0	1 1 0 1
=	0 1 1	1 1 0 1
A	1 0 0	0 0 0 1
B	1 0 0	0 0 1 0
...	...	...
O	1 0 0	1 1 1 1
P	1 0 1	0 0 0 0
Q	1 0 1	0 0 0 1
...	...	...
...	...	...
Z	1 0 1	1 0 1 0

ASCII code

134

Online Classes : Megalecture@gmail.com
www.youtube.com/megalecture
www.megalecture.com

Chapter 8 Codes in Everyday Use

Solution

Using the diagram (or the table) you can follow through the code, stopping when any letter is reached:

$$0\ 1\ 1\ 0 \quad 0\ 0 \quad 0\ 1\ 0 \quad 1 \quad 0\ 1\ 0 \quad 0\ 0 \quad 0\ 1\ 1\ 0$$

$$\textbf{N} \quad \textbf{A} \quad \textbf{M} \quad \textbf{E} \quad \textbf{M} \quad \textbf{A} \quad \textbf{N}$$

Huffman code for five letters

Exercise 8B

1. Decode the following, using the Huffman code above

   (a) 0 1 0 0 0 0 1 1 0 0 0 0 1 1 1 1 0 1 1 1 1 0 0

   (b) 0 1 0 1 1 0 1 1 1 0 1 1 1 1 0 0 0 1 0 0 0 0 1 1
       1 0 1 1 1 1 0 1 1 0

2. Design a Huffman code if the only codewords used are as shown below and all words are used equally frequently.

BUS	CUPS	MUSH	PUSS
SIP	PUSH	CUSS	HIP
PUP	PUPS	HIPS	

8.7 Miscellaneous Exercises

1. The list below shows the International Morse Code (in 'dots and dashes') for some letters of the alphabet.

   J ●−−−   K −●−   L ●−●●   M −−   N −●

   Using these letters only,

   (a) give an example to show that in Morse Code even a single error can go undetected;

   (b) give an example of 7 dots and dashes to show that, unless a pause is left between letters, a message received in Morse Code may be decoded in more than one way.

2. A new furniture mail order company is designing a coding system for its variety of products. Information required to be coded includes:

   type of product
   size
   colour
   price
   catalogue number.

   Design a bar code system for identifying products in this company. Explain the rationale behind your design.

3. Research into one of the commonly used codes not covered earlier, and write a report outlining

   (a) the design of the code used

   (b) how it works in practice

   (c) advantages and disadvantages of the code.

4. Design a new coding system to solve a particular practical problem.

136