Chapter 1

Where does information come from?

1.1 Introduction

This book is designed to provide you with an explanation of the basic concepts of information collecting and processing systems. To do this we will examine examples ranging from secret codes to compact disc players. Using these practical examples you should be able to see how the mathematics of *Information Theory* can be applied in practical situations to make *Instruments* which perform useful tasks. This first chapter is intended to a be a general outline. Most of the concepts introduced here will be looked at more carefully later.

Scientists and engineers devote considerable attention to the processing and storage of information, yet questions relating to how information is produced generally attract less consideration. To some extent, this blind spot seems to stem from a belief that any interest in this area smells strongly of philosophy, not engineering. In general, practically minded scientists don't want to 'waste their time' with philosophy — although there are many notable exceptions to this rule.

This book is not about philosophy. No time will be devoted to questions like:
'What is the meaning of meaning?'
'How do we know what we know?'
etc.
Despite this, when trying to understand information based systems it's vital to have some idea of how information is created or captured.

1.2 What ***is*** *information?*

For our purposes, we can say that information initially comes from some form of sensor or transducer. This generates some form of response which can then be measured. It is this measurement or detection which 'creates' information. (In fact, the sensor is reacting to the arrival of some input pattern of energy or power. It would be fairer to say it 'picks up' the information, but we'll ignore this fact.) Once we adopt this starting point it becomes clear that the topics of instrumentation and measurement form the basis of all practical information systems.

This viewpoint provides us with a double advantage over someone who is studying information theory purely as a branch of mathematics. Firstly, it gives us a way to understand information processing systems in terms of the physical properties of the real world. Secondly, it helps us sort out questions related to the 'value' or 'meaning' of information without the risk of being dragged into metaphysics. Instead we can simply ask, 'How was this information produced?'

What is an 'instrument'? At first glance, it can appear to science and engineering students that the subject called *Instrumentation* is obsessed with describing how voltmeters and oscilloscopes work. Yet the subject covers a much wider and more important area. A colour TV is an instrument. A digital computer is an instrument. Each senses some form of input and responds by producing an appropriate output. The TV responds to an electromagnetic wave from a distant transmitter to produce a corresponding picture on a screen and sound from a loudspeaker. The computer can be affected by various sorts of input, from a keyboard, a mouse, or by reading a magnetic disc. It can respond by altering the electronic pattern held in its memory, by altering its monitor display, or recording something on a disc.

Most of the examples we'll look at in this book will be electronic or optical. This is because optical and electronic methods are powerful and widely used. Despite this, it's important to realise that the basic points made in this book aren't only true in these areas. To emphasise this, we can start by considering a simple mechanical measurement system — a kitchen balance — to make some fundamental points which apply to all measurement (information gathering) systems.

The balance has a pan or plate supported by a spring. When we place something on the pan the added weight presses down on the spring, compressing it. The pan moves downwards until the compression force from the squeezed spring balances the force of the increased weight. Most balances have a rotary dial with a pointer attached to the pan. The movement caused by the weight rotates the pointer to give us a 'reading' of the weight.

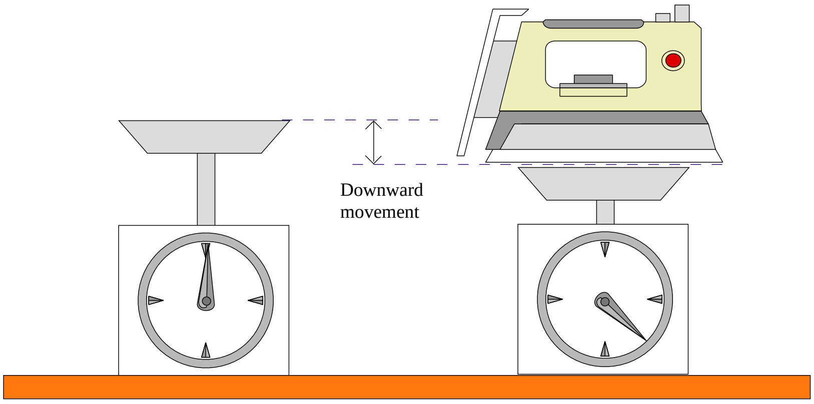

Figure 1.1 Kitchen technology measurement system.

The first point to note is that, like most measurement systems, this one is *indirect*. What we actually observe is a movement (rotation) of the pointer. We don't actually see the magnitude of the weight. If, for example, we put an iron on the pan we might see the pointer move around through 120 degrees. If we liked, we could also use a ruler to find that the pan moved down 2 cm. However, we don't usually quote weights in degrees or centimetres! In order to make sense of these observed values we have to *calibrate* the balance. To do this we can place two or three different known weights on the scales and make a note of how far the pointer goes around (or the pan falls) each time. We can then use these results to make a series of calibration marks on the face of the dial. Now, when we put something — e.g an iron — on the scales we can read off its weight from the dial. This calibration process means that the balance provides us with a means to compare the weight of the iron with a set of other 'standard' weights. In general, all measurements are *Comparisons* with some defined standard.

Usually, we buy a kitchen balance which should already be calibrated (i.e. its dial is marked in kg, lb, etc, not degrees) and we don't bother to calibrate the weighing instrument for ourselves. However, when we consider the need for a calibration process an awkward question springs to mind — where did the 'known' weights come from that were used to calibrate the readings? If all measurements are comparisons, how were the values of those weights known? They, too, would need to have been weighed on some weight measurement system. If so, how was that system calibrated?

Any measurement we make is the last link in a chain of similar measurements. Each one calibrates a system or a ‘standard’ (e.g. a known weight) which can be used for the next step. Right back at the beginning of this chain (at the National Physical Laboratory in the UK and other standards labs around the world) there will a *Primary Reference* system or standard which is used to define what we mean by ‘1 kg’, or ‘1 second’ or whatever. In effect, when we plonk something on the pan of a kitchen balance we're indirectly comparing it with the standard kg weight kept under a glass cover at the NPL.

When we place an iron on the pan we have to wait a second or two to let the system settle down and allow the pointer to stop moving. Similarly, when we remove the iron the system takes a short time to recover. The second point we can make about the measurement system is, therefore, that it has a finite *Response Time* — i.e. we have to wait for a specific time after any change in the weight before we can make a reliable reading. This limits our ability to measure any changes which take place too quickly for the system.

The third point to note is fairly obvious from our choice of an iron. If we put too large a weight on the pan the pointer will go right around and move ‘off scale’. (If the iron is very heavy we may even smash the scales!) No matter how well we search the shops, we can't find scales which can accurately measure any weight, no matter how big. Every real instrument is limited to operate over some finite *Range.* Beyond this range it won't work properly and *Overloads* or *Saturates* to give a meaningless response.

The fourth and final basic point is something we won't usually notice using ordinary scales since the effect is relatively small. All of the atoms in the scales, including those in its spring, will be at room temperature. (In a kitchen this probably means at or above 20 Celsius or 293 Kelvin.) As a result, they'll be moving around with random thermal motions. Compared to the effect of placing an iron on the scales these movements are quite small. However, if we looked at the pointer very carefully with a powerful microscope we'd see its angle fluctuating randomly up and down a little bit because of the motions of the atoms in the spring. As a result, if we wanted to measure the weight very accurately this thermal jittering would limit the precision of our reading. As a result, no matter how good the scales, our ability to make extremely accurate measurements is limited by thermal random effects or thermal *noise.*

1.3 Accuracy and resolution

It is important to realise that the amount of information we can collect is always finite. The example of kitchen scales has introduced us to the limiting effects of clipping, noise, and response time. It doesn't matter how clever we are, these problems occur in all physical systems since they are consequences of the way the real world works. To see some of the other problems which arise when we're collecting information, consider the system in figure 1.2. This diagram represents a diffraction grating being used to measure the power/frequency spectrum produced by a light source.

The system is intended to provide us with information about how bright the light source is at various light wavelengths. It relies upon the reflection properties of a surface made with a series of parallel ridges called a *Reflection Grating*. For an ordinary plane mirror, the angle of reflection equals the angle of incidence. For a grating, the angle of reflection also depends upon the wavelength of the light and the details of the grooved surface pattern. Hence the arrangement shown acts as a sort of adjustable filter. Only those light wavelengths which reflect at the appropriate angle will make their way through the output slit onto the detector.

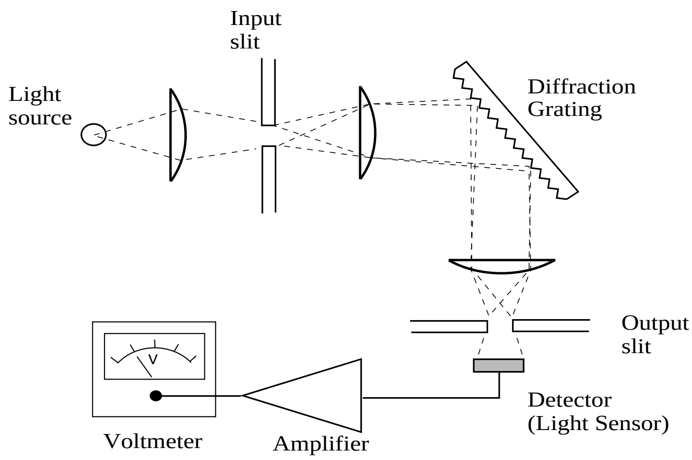

Figure 1.2 Simple diffraction grating spectrometer.

As with the kitchen scales, the system provides an indirect way to measure the light's spectrum. We use the angle of the diffraction grating to tell us the wavelength being observed. The voltage displayed on the meter indicates the light power falling on the detector. To discover the light's spectrum we slowly rotate the grating (or move the output lens/slit/detector) and note how the voltmeter reading varies with the grating angle. To convert these angles and voltages into wavelengths and light powers we then need to know the *Sensitivity* of the detector/amplifier system and the angles at which various wavelengths would be reflected by the grating – i.e. the system must be calibrated.

In most cases the instrument will be supplied with appropriate display scales. The voltmeter will have a dial marked in units of light power, not volts. The grating angle display will be marked in wavelengths, not degrees. These scales will have been produced by a calibration process. If the measurements we're making are important it will probably be sensible to check the calibration by making some measurements of our own on a 'known' light source.

As with the kitchen balance, our ability to measure small changes in the light level will be limited by random noise — in this case random movements of the electrons in the measurement system and fluctuations in the rate at which photons strike the detector. The accuracy of the power measurement will depend upon the ratio of the light power level hitting the detector to the random noise. We could increase the light level and improve the precision of the power measurement by widening the slits and allowing more light through. However, this would have the disadvantage of allowing light reflected over a wider *range* of angles to reach the detector. Since the angle of reflection depends upon the light wavelength this means we are allowing through a wider range of wavelengths.

In fact, looking at the system we can see that it always allows through a range of wavelengths. Unless the slits are narrowed down to nothing (cutting off all the light!) it will always allow light reflected over some range of angles, $\Delta\Theta$ (and hence having a range of wavelengths, $\Delta\lambda$) to get through. As a result there is an unavoidable 'trade off' between the instrument's power sensitivity and its frequency *Resolution* or ability to distinguish variations in power confined to a narrow frequency interval. This kind of trade off is very common in information collection systems. It stems from basic properties of the physical world and means that the amount of information we can collect is always finite — i.e. we can never make perfect measurements with absolute accuracy or precision or certainty.

Summary

You should now know that information is collected by *Instruments* which perform some kind of *Measurement.* That measurement systems usually give an *Indirect* indication of the measured quantity and that all measurements are *Comparisons* which have to be *Calibrated* in some way. The amount of information we can collect is always finite, limited by the effects of *Noise, Saturation* (or *Overload*), and *Response Time.* That many information gathering techniques involve a *Trade-Off* between various quantities — for example, between the *Resolution* of a wavelength measurement and the *Sensitivity* of a related power measurement. That these limitations arise from the properties of the physical world, not poor instrument design.