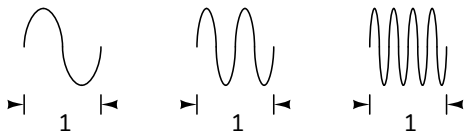
Solutions to Quiz 1

Question 1

A data source can generate three different waveforms:



The duration of each waveform is 1 µs (or 1 ms). The probabilities of the waveforms are 0.125, 0.5 and 0.375.

- What is the symbol rate?
- What is the entropy of the source?
- What is the information rate of the source?

Solution

- The symbol rate is the inverse of the symbol period which is given as 1 µs (or 1 ms). The symbol rate is thus $1/1\times10^{-6} = 1$ MHz (or $1/1\times10^{-3} = 1$ kHz).
- The entropy of a source is defined as:

$$H = \sum_i (-\log_2(P_i) \times P_i) \quad \text{bits/message}$$

  In this question the message probabilities, P_i, are 0.125, 0.5 and 0.375 so the entropy of this source is:

$$H = -\log_2(0.125) \times 0.125 - \log_2(0.5) \times 0.5 - \log_2(0.375) \times 0.375 \approx 1.4 \quad \text{bits/message}$$

- The information rate in bits/second is given as the information rate in bits per message divided by the message duration in seconds per message: $1.4\,\text{bits/message}/1\,\mu\text{s/message} = 1.4$ Mb/s (or $1.4\,\text{bits/message}/1\,\text{ms/message} = 1.4$ kb/s).

Question 2

The Japanese character for Table is 卓 and has a Unicode code point of U+5353 (or Electricity: 電, U+96FB). What is the UTF-8 encoding? Give your answer as a sequence of bytes in hexadecimal.

Solution

For the character 卓 the binary representation of **`0x5353`** is **`0101 0011 01 01 0011`**. According to the UTF-8 encoding table this requires a 3-byte encoding.

The values of the z, y and x fields are **`0101`** (**`0x05`**), **`001101`** (**`0x0D`**) and **`010011`** (**`0x13`**). When these are added to the prefixes (**`E0`**, **`80`**, **`80`**) the result is the bytes **`0xE5`**, **`0x8D`** and **`0x93`**.

For the character 電 the binary representation of **`0x96FB`** is **`0101 0011 01 01 0011`**. According to the UTF-8 encoding table this requires a 3-byte encoding.

The values of the z, y and x fields are **`1001`** (**`0x09`**), **`011011`** (**`0x1B`**) and **`111011`** (**`0x3B`**). When these are added to the prefixes (**`E0`**, **`80`**, **`80`**) the result is the bytes **`0xE9`**, **`0x9B`** and **`0xBB`**.

Question 3

A serial interface uses a DB-9 connector. You measure the voltage on pin 4 (or 6) as 9 V.

- Is this connector wired as a DTE or DCE? Briefly explain how you arrived at your answer.
- Could you connect this port directly to a typical PC serial port which is set up as a DTE? If not, how could such a connection be achieved?

Solution

- The table in the lecture notes shows that pin 4 on a "DB-9" connector is DTR. A voltage of 9 V indicates this pin is an output. DTR is an output

on a DTE so this connector must be wired as a DTE.

The table in the lecture notes shows that pin 6 on a "DB-9" connector is DSR. A voltage of 9 V indicates this pin is an output. DSR is an output on a DCE so this connector must be wired as a DCE.

- A port wired as a DTE can only communicate with a port wired as a DCE and vice-versa. If pin 4 (DTR) is an output the connector is a DTE and it would need to be connected to a DCE, not a DTE such as typical PC. A null-modem adapter would be required.

  If pin 6 (DSR) is an output then the connector is wired as a DCE and it could be connected to a DTE. No adaptor would be required (assuming the connectors are physically compatible).

Question 4

The bits corresponding to the ASCII (or UTF-8) encodings of the characters **`AT`** are transmitted lsb-first. Write out the bits in the order they would be transmitted. Do not include start, stop or parity bits – only the data bits in lsb-first order.

Solution

The ASCII encoding for **`A`** is **`0x41`** (**`0100 0001`** in binary) [1] and the encoding for **`T`** is **`0x54`** (**`0101 0100`** in binary). If these characters were transmitted lsb-first the data bits transmitted would be:

```
1000 0010 0010 1010
```

[1] ASCII encodings are the same as UTF-8.