Question 1.

What is the amount of information when transmitting a character in ASCII code?

Answer:

The amount of information when transmitting a character y_i in ASCII code is
$I(y_i) = -\log_2 p(y_i)$

In the case of standard ASCII code, the probability that y_i is chosen is $p(y_i) = \frac{1}{128}$

$I(y_i) = -\log_2 \frac{1}{128} = 7$

In the case of extended ASCII code, the probability that y_i is chosen is $p(y_i) = \frac{1}{256}$

$I(y_i) = -\log_2 \frac{1}{256} = 8$

Question 2

What is the maximal average amount of information (entropy) that can be transmitted with a Poisson spike train with an average firing rate of 100 Hz?

Answer

The entropy of observing N spikes in the time interval T with a Poisson spike train is

$$S(N;T) \approx \frac{1}{2}\log_2 N + \frac{1}{2}\log_2 2\pi \ , \qquad N = rT$$

And the average firing rate is 100HZ, we can get $N = 100T$.

Therefore, the maximal average amount of information that can be transmitted with a Poisson spike train with an average firing rate of 100HZ is

$$S(N;T) \approx \frac{1}{2}\log_2 100T + \frac{1}{2}\log_2 2\pi \ , \qquad T \text{ is the time interval.}$$

It is maximum in the sense that we consider only the noiseless case.

Question 3

What is the maximal firing rate of an integrate-and-fire neuron with an absolute refractory time of 2ms?

Answer :

The firing rate of an integrate-and-fire neuron is

$$\bar{r} = \frac{1}{t^{ref} - T_m \ln \frac{v - RI}{u_{res} - RI}}$$

also $t^f = -T_m \ln \frac{v - RI}{u_{res} - RI} \geq 0$, therefore, if $t^f = 0$, $\bar{r}$ will have maximal value, that is

$$\bar{r} = \frac{1}{t^{ref}} = \frac{1}{2 \times \frac{1}{1000}} = 500HZ$$

Question 4

How many hidden layers are necessary to implement the Boolean XOR function with a feedforward neural network? Can the activation function of the hidden nodes be linear? (Explain briefly)

Answer:

1) We don't have to use hidden layers to implement the Boolean XOR function. Because if we employ a nonlinear activation function, such as a rotation combined with a non-monotonic function, we can represent the Boolean XOR function with a single sigma node.

2) We need to use a non linear activation function for the hidden nodes, since units of hidden layers using linear functions will not be more powerful than a well chosen single layer. Suppose we use linear activation functions $g(x) = ax + b$ to all the nodes in a two hidden layer network.

$$r^{out} = g(w^{out} g(w^{h2} g(w^{h1} r^{in}))) = g(w^{out} g(aw^{h2} w^{h1} r^{in} + bw^{h2}))$$
$$= g(a^2 w^{out} w^{h2} w^{h1} r^{in} + abw^{out} w^{h2} + bw^{out}) = g(Ar^{in} + B)$$

In the last step, I use A to stand for the multiplication of a series of weights matrices and constant, use B to stand for the other part.

We can see that each layer would perform as a purely linear operation on its inputs and the multiple layers network would still perform as a single-layer network.

Question 5

A Boolean function can be defined with a truth table. A specific Boolean function is given by the following truth table

Answer:

x_1	x_2	y_1	y_2
true	true	true	true
true	false	false	true
false	true	true	false
False	false	true	true