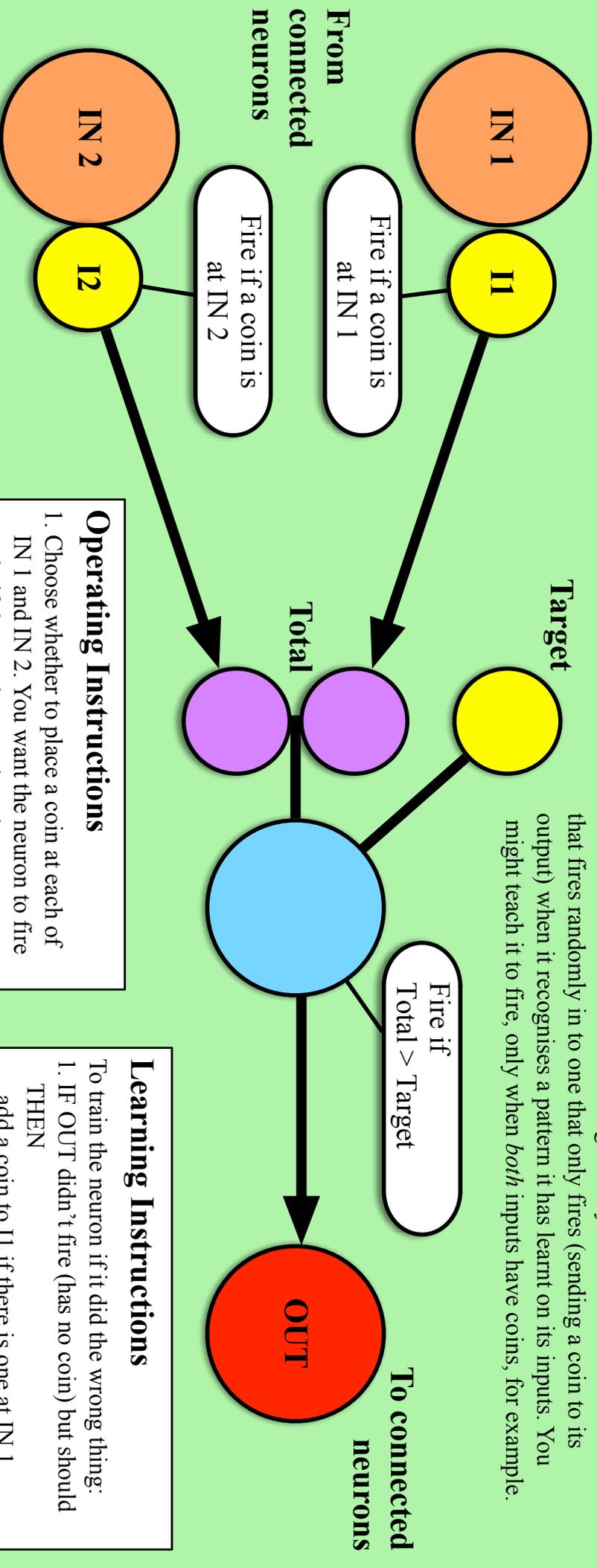


# Training an Artificial Neuron

[www.cs4fn.org/machinelearning/](http://www.cs4fn.org/machinelearning/)

These instructions show how to gradually turn an artificial neuron that fires randomly in to one that only fires (sending a coin to its output) when it recognises a pattern it has learnt on its inputs. You might teach it to fire, only when *both* inputs have coins, for example.



## Set Up Instructions

1. Place 0, 1, 2 or 3 small coins on each of the three yellow circles at random.
2. Place a large coin on the blue circle.

## Operating Instructions

1. Choose whether to place a coin at each of IN 1 and IN 2. You want the neuron to fire only if these coins match a given pattern.
2. If the linked rules (in white boxes) say to fire then move the piles of coins down the arrows from yellow to purple circles.
3. Fire the coin from blue to red OUT circle if there are more coins at Total than Target.
4. After each round move all coins back to where they came from to reset.

## Learning Instructions

To train the neuron if it did the wrong thing:

1. IF OUT didn't fire (has no coin) but should THEN add a coin to I1 if there is one at IN 1 add a coin to I2 if there is one at IN 2
2. IF OUT fired (has a coin) but shouldn't have THEN remove a coin from I1 if there is one at IN 1 remove a coin from I2 if there is one at IN 2

Note: if a pile runs out of coins you can put IOUs (coins marked -1) there.

## How does a neuron learn?

Everything you do, think and feel involves neurons in your brain sending messages to each other. One of the most powerful things about our brains is the way neurons allow us to learn. Computer scientists have been inspired by the way they work to make machines that learn too, programming networks of artificial neurons called neural networks.

### Our brains

You have billions of neurons. Each is a very simple computational unit. It has a central *nucleus*, lots of inputs called *dendrites*, and a single output called an *axon*. The outputs link to the inputs of other neurons, passing chemical messages, called *neurotransmitters*, to each other.

Each neuron follows a simple rule about when to *fire*: when to send a message to the neurons it is connected to. A neuron fires when the number of messages arriving reaches a *threshold*. It's a bit like blowing up a balloon until eventually it can take no more and pops. For a neuron, 'popping' involves sending neurotransmitter messages down its output axon to all the neurons it is connected to. That leads to those neurons edging towards their own threshold and so eventually firing themselves.

### Learning

So what is going on in your brain when you learn? What are the neurons doing? It actually boils down to individual neurons just adjusting their threshold, adjusting when they fire. They do this based on feedback about whether they did the right thing in the past or not. Some neurons may need more signal to fire at the right time, some less. By making repeated gradual adjustments they eventually fire only when they should. For example, a small child does not have fine control over their muscles and so struggle to draw smooth lines with a pencil. The more they draw, the more their neurons adjust when to fire, until they have the control to draw smooth lines.

### Playing the game

The simple board game overleaf demonstrates how an individual neuron gradually learns the right threshold for a given task. Coins are used to represent the neurotransmitters. Initially the artificial neuron fires randomly as it has a random number of internal coins, but gradually it turns in to a neuron that only fires (sending a coin to its output) when it recognises a pattern it has learnt on its inputs. In the game you follow the learning rules, adjusting the weights of the signals from its inputs and so its threshold by adding or taking away coins. Real neurons follow the rules themselves of course. You might teach it to fire only when *both* inputs have coins, for example, or alternatively to fire when only one does.

### Machine Learning

Neural networks, that form the basis of machine learning programs, learn in a similar way, adjusting the thresholds of their artificial neurons based on training data. They are shown lots of examples of the pattern the network is to learn: pictures of cats perhaps. The neurons fire according to their settings, and the network as a whole makes a judgement as to whether the pattern is a cat. If wrong, the neurons adjust. They do this repeatedly over lots of pictures until the whole network fires only when the picture is of a cat.