NLP

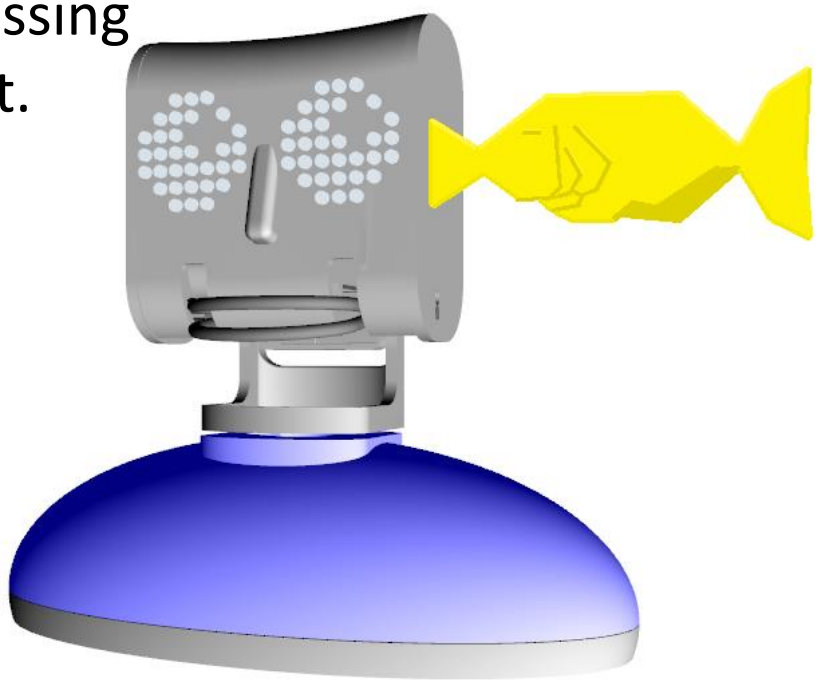
Part 1

Understand how a computer attempts to process human language

Humans learn language skills through experience. By interacting and communicating with other people we pick up on context clues and patterns in speech. The more we practise, the better we get at communication.

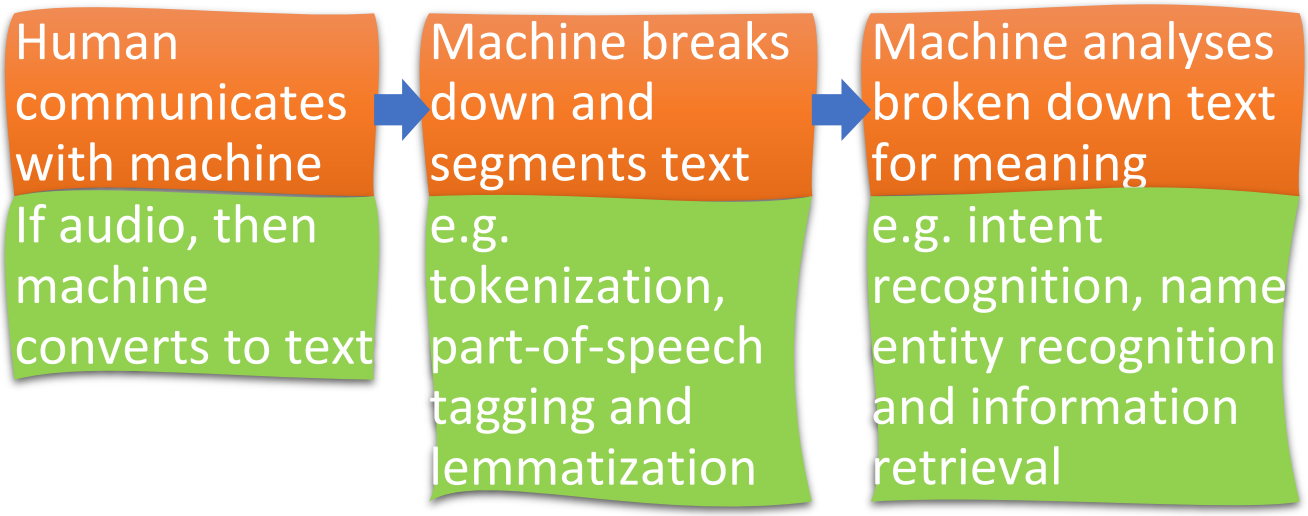
Unfortunately computers don’t learn like that and even if they did, they don’t get out very much. So people learn programming languages in order to communicate with computers.

If we want computers to understand human language then we have to insert a Babel fish into the USB port. Just kidding, we use natural language processing or NLP for short.



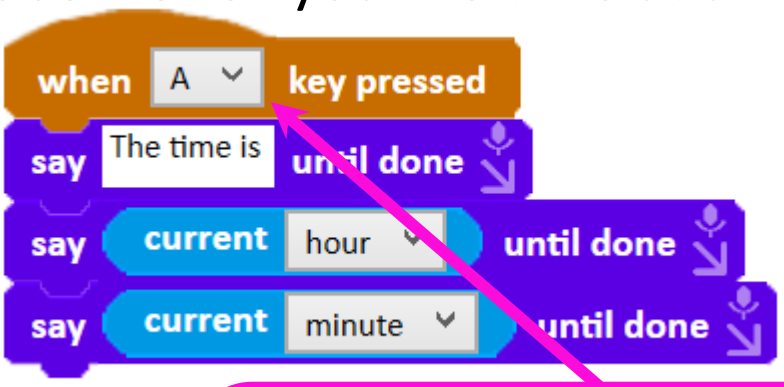
NLP is the method by which computers derive meaning from raw data.

In a very basic way it looks something like this…



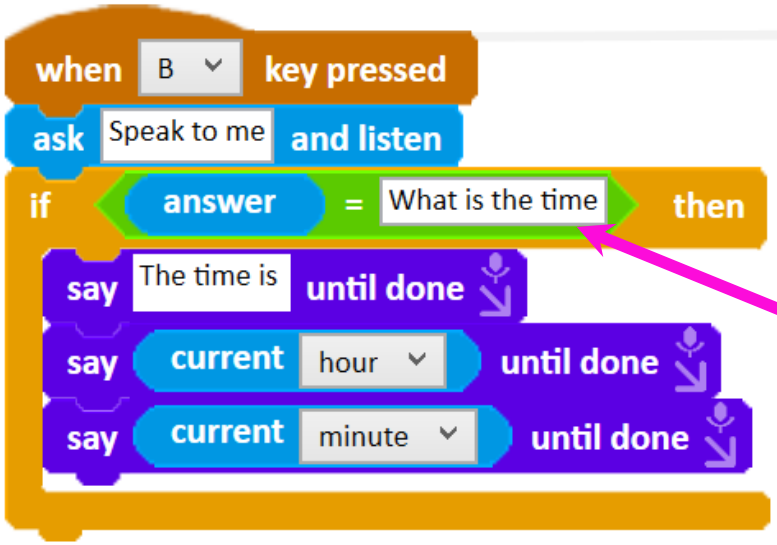
Let’s use asking for the time as an example and try to get our robot to understand us using NLP.

To start off, let's look at how you can get the robot to tell you the time at all.



When the "A" key is pressed, the robot will say "The time is" followed by the current hour and minute on your computer's clock

Now let's introduce some human language into the program.

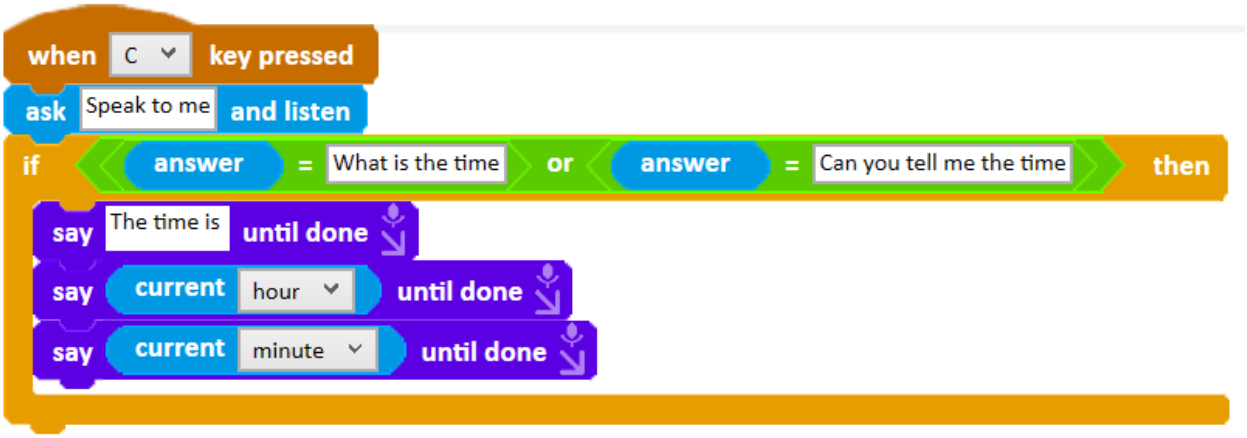


Press "B" and the robot will say "Speak to me", if you reply "What is the time" then it will say the time

Can you see the limitations of this program?

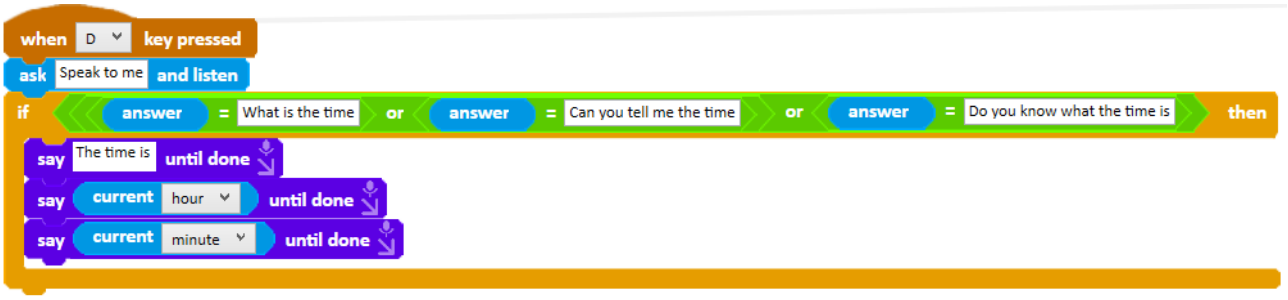
In that last example, the robot will only reply if it hears the exact phrase, “what is the time?”

There are many different ways of asking for the time. So what is the solution?



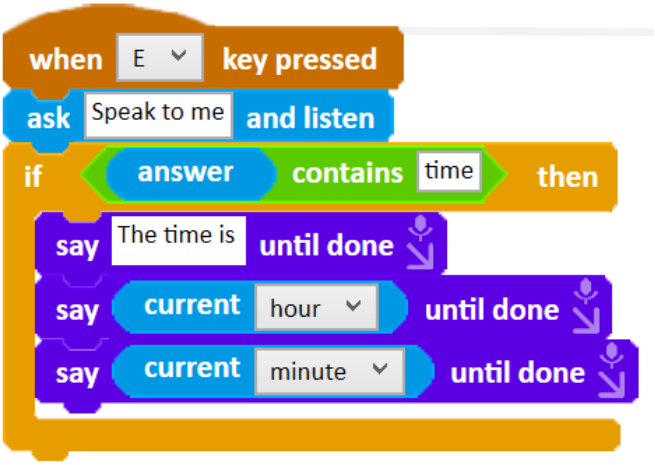
Can you see how this example is better?

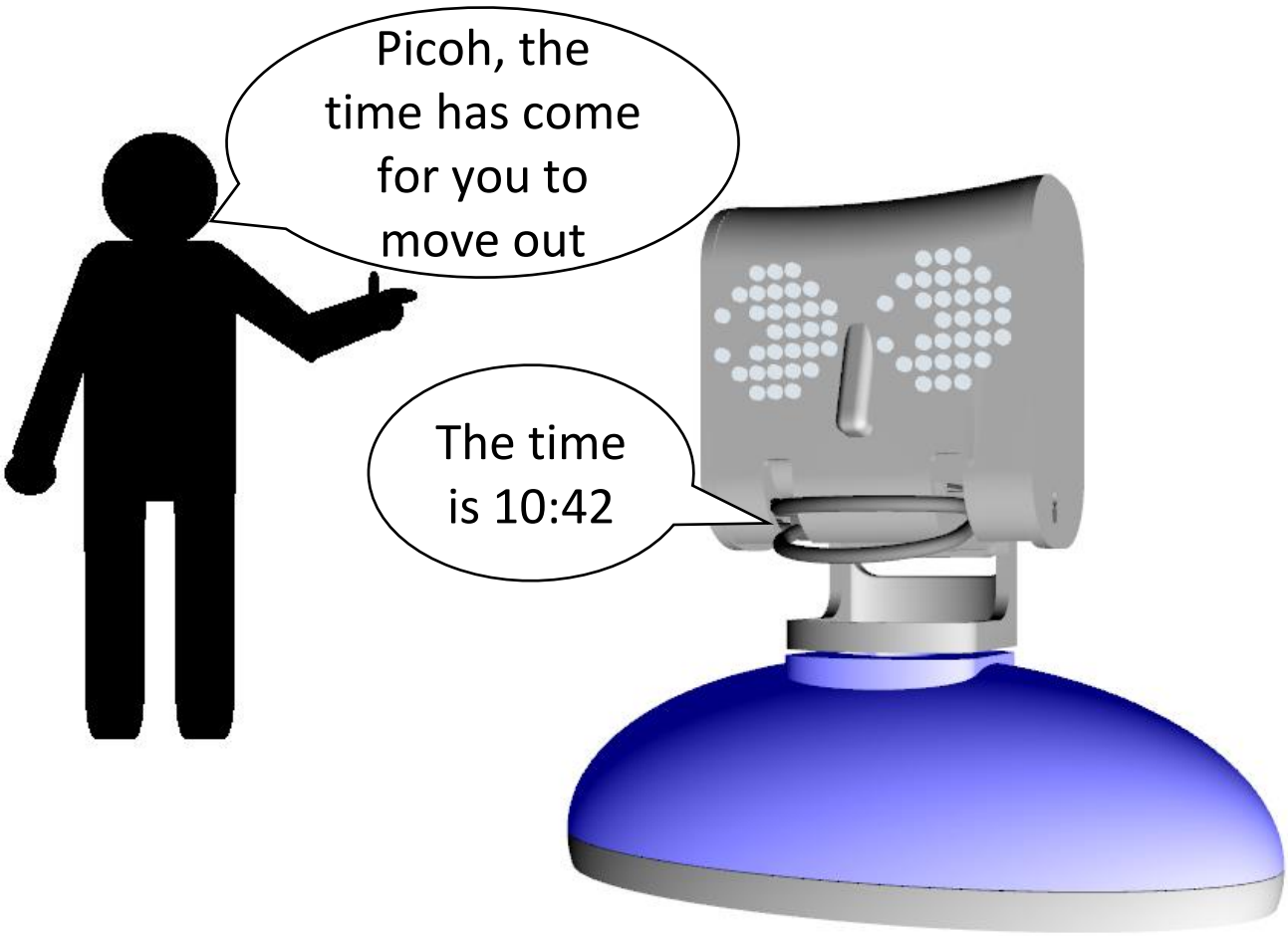
I wonder how many different ways there are of asking for the time.



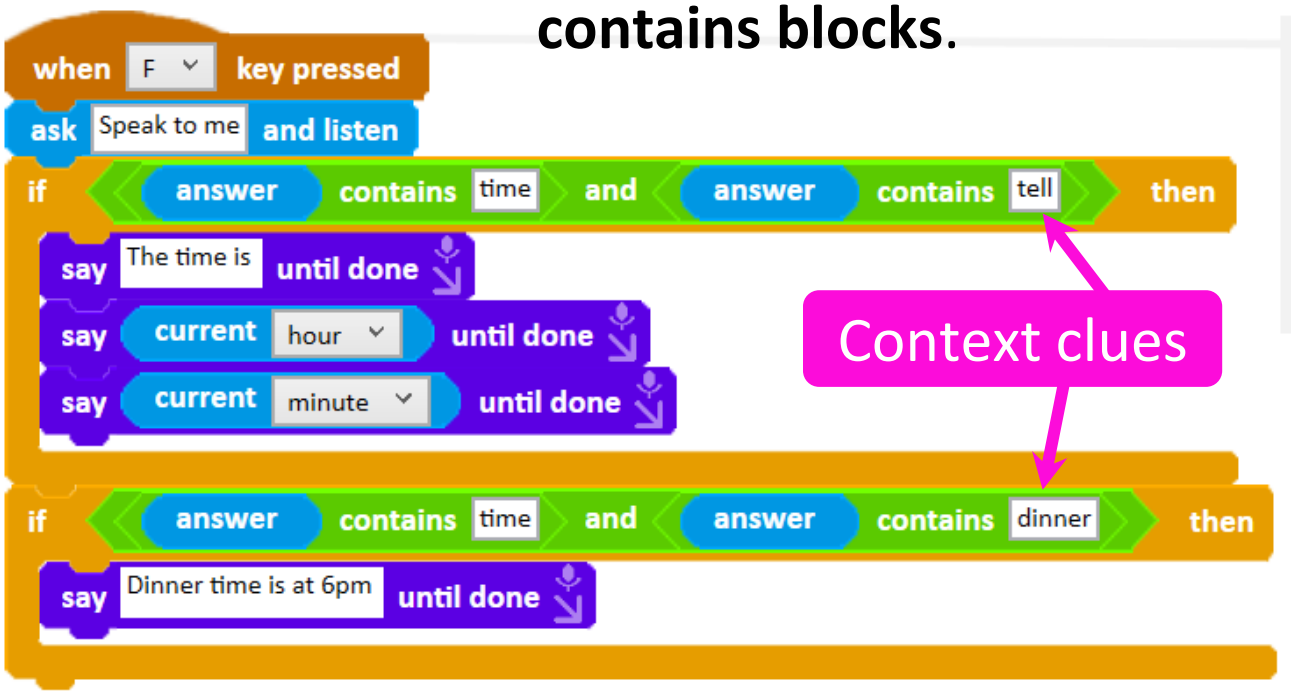
We could include them all but this is getting seriously cumbersome. Can you see the pros and cons to this approach?

Here's a way based on inference. The code here 'guesses' that you want to know the time when a phrase you say contains the word time. Can you see advantages and disadvantages to this approach?



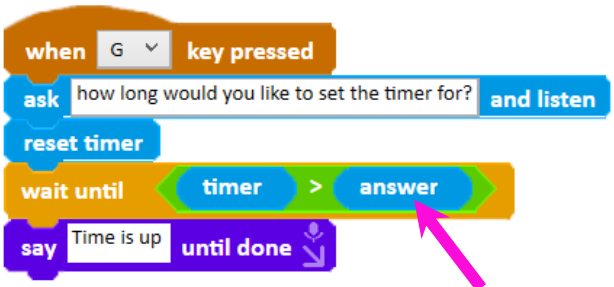


The problem being that assuming the word “time” is only ever used to ask for the time is wrong. For example, we use phrases like “dinner time”. If we ask the robot, “is it dinner time?” we don’t want to know what the current time is. So we have to give the robot some context by including other words in the **contains blocks**.

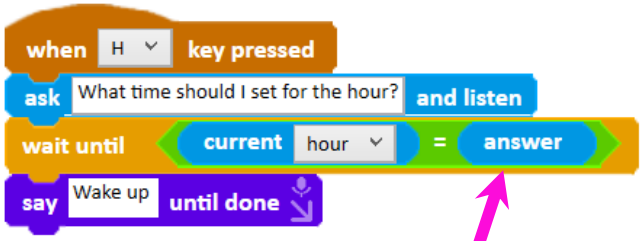


We’re still using inference rather than exact sentence matching but this way the robot has a better chance of giving a useful answer.

Apart from the different ways we use the word time, there are different functions we want a clock to perform. Like setting an alarm or a timer.



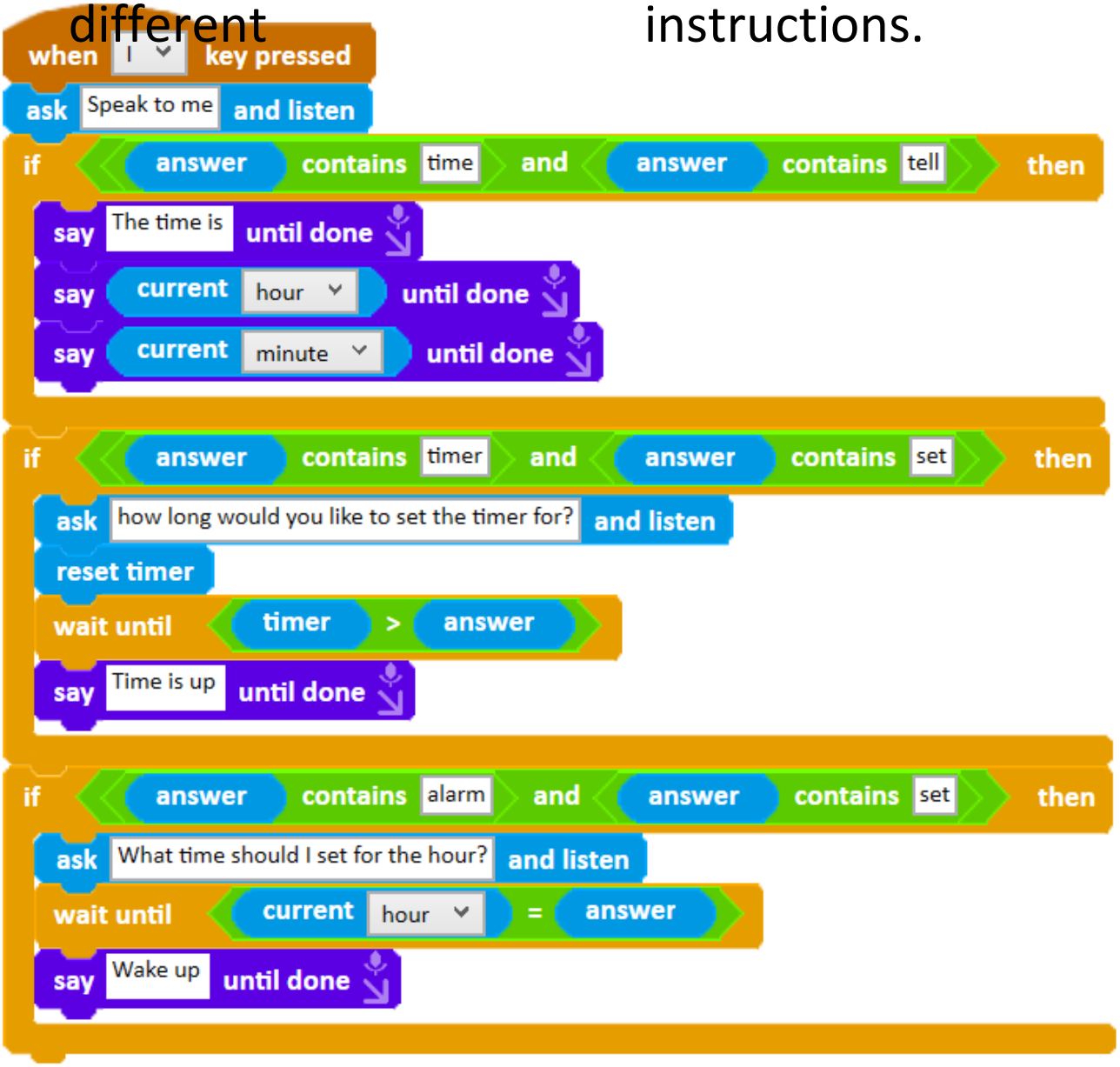
Answer with a number and the robot will count to that number of seconds



Answer with an hour of the day and the robot will say “wake up” when it is that time

Using these programs and one for asking the time, can you think how you would combine them all? And what words would you choose for the **contains blocks**?

By selecting the right pair of words for the **contains blocks** the robot is able to listen to our speech and distinguish between different instructions.



For further explanation of NLP and more powerful examples, take a look at the next NLP resource.