Problem statement

The core of the project will be to implement the Hidden Markov Model (HMM) algorithm. We will apply the HMM to a large text and identify which of the symbols in the text are vowels and which are consonants.

The HMM is based on the Markov Chains that describe how a system can change state and the probabilities of moving from one state to another (or to the same state). The HMM tries to find hidden states based on a set of observations.

The project is based Mark Stamp's Article[1] from 17Oct 2018.

What we will use from the course:

Since the HMM can be divided into three main problems, we will make use of functional programming (**encapsulation - not sure here**) to solve these individually. Each problem can be divided into its own function and later used together.

We will apply the Python coding style learned in the course. Multiple data types will be used, depending on how we implement the algorithm.

What we will learn from this project:

- Introduction to Machine Learning using the HMM.
- Implementing advanced algorithms using python.
- Using the NumPy or Pandas library.

---

[1] A revealing introduction to Hidden Markov Models