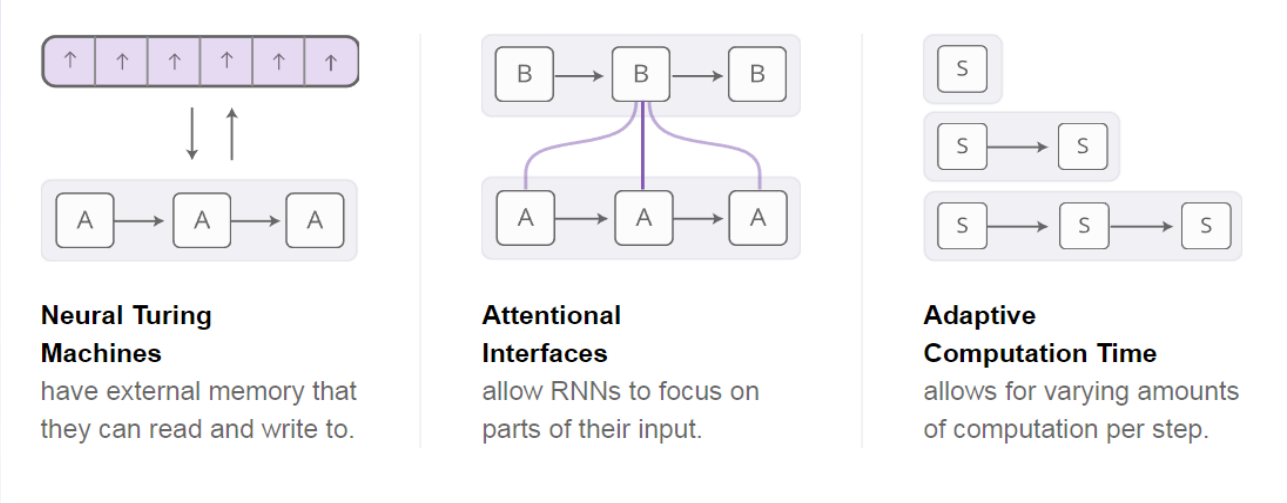
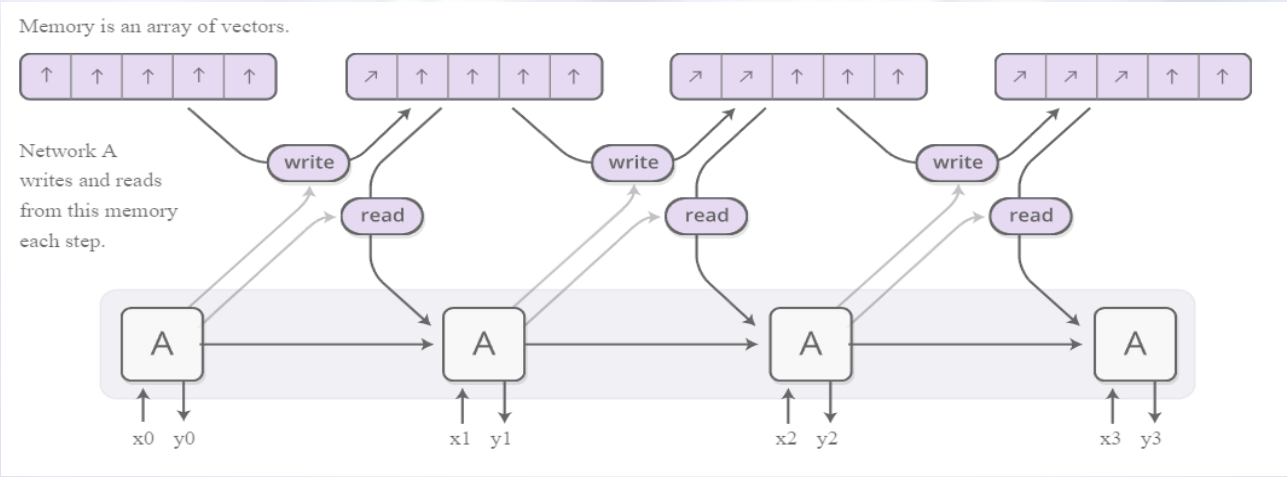
(E) attention units:

Ongoing research on LSTMs with attention units produced many different attempts at improving these units. We will focus on the Neural Turing machine, but will also review two additional types of attention unit to demonstrate their shared attributes:

1. Neural Turing machine
2. Attentional Interfaces
3. Adaptive Computation



Neural Turing machine



The **Neural Turing Machine (NTM)** adds an external memory to the basic LSTM network, which it can write to and read from, allowing the network to better understand what it is doing (e.g. mimic a lookup table with key-value, learn to store a long sequence in memory and so on). The external memory is a vector array, the question arises where in the vector array should the network write to and read from?