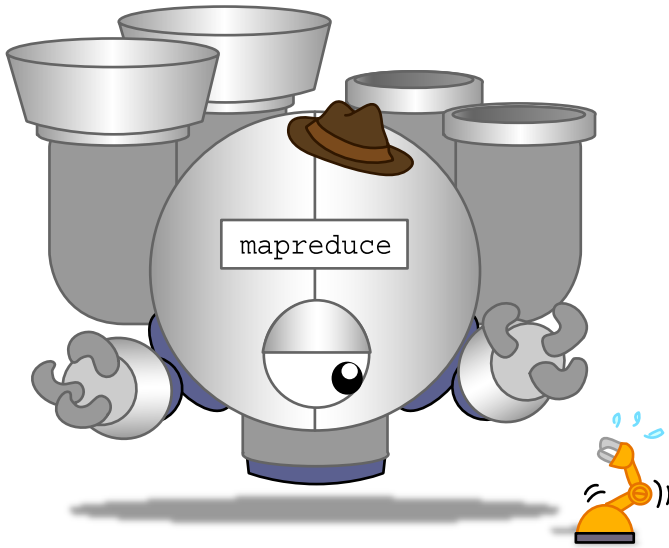




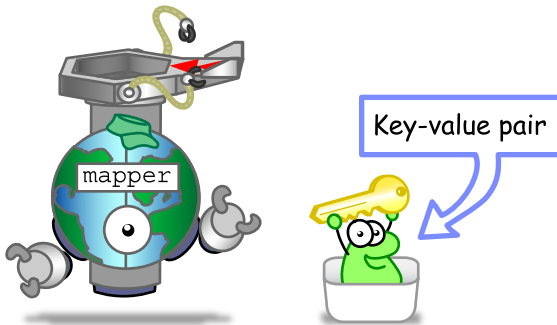
# MapReduce

Parallelism and Functional Programming



MapReduce is a system that makes writing parallel code easier for programmers.

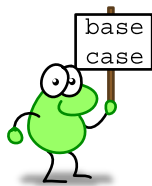
It takes four arguments:



The first argument is the mapper, a function that takes one key-value pair and outputs a list of intermediate key-value pairs. (A mapper is the kind of function used as an argument to `map`.)



Next is the reducer, which collects and combines values. It plays a vital role in turning intermediate key-value pairs into output. (A reducer is the kind of function used as an argument to `accumulate`, also known as `reduce`.)



Then comes the base case, which is usually the "identity" of the reducer function given to `accumulate`.



Finally, there's the vast body of data to be processed, encoded as a stream of key-value pairs.