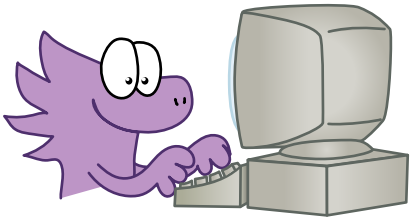
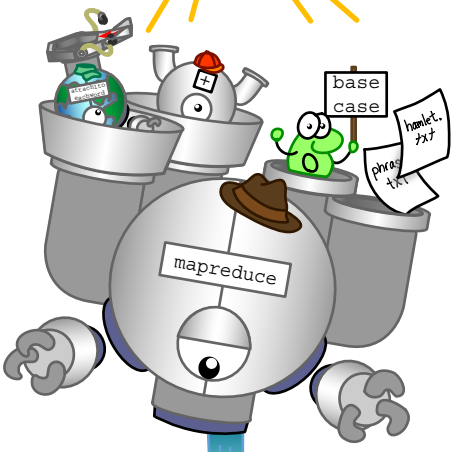
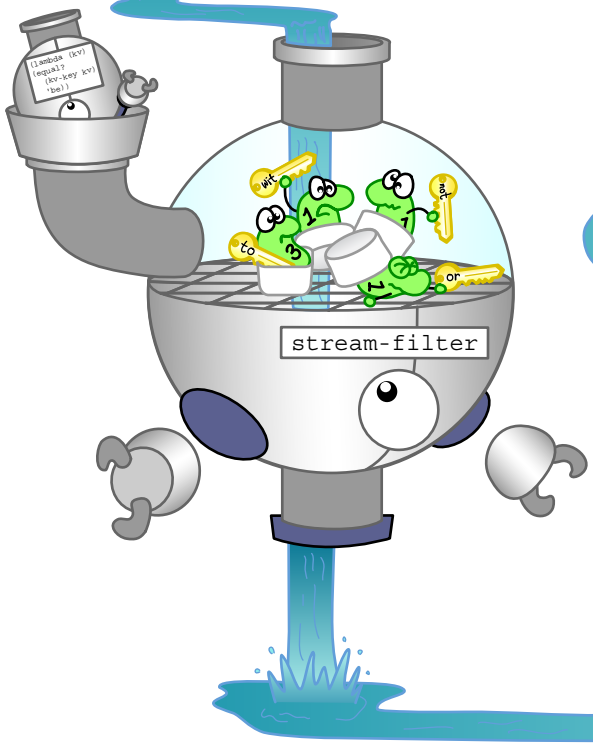
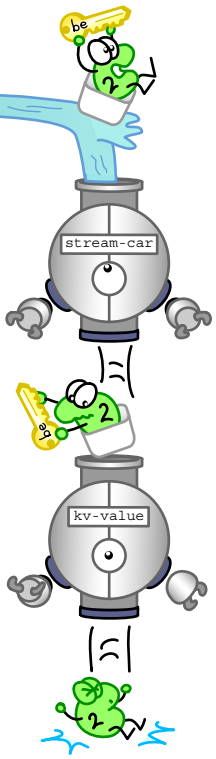
MapReduce

The Wordcount in Code



```
unix% ls /docs
hamlet.txt phrases.txt

unix% stk
;; Define the mapper
;; Typical kv-pair:
;; ("phrases.txt" . (to wit))
;; Output:
;; ((to . 1) (wit . 1))
> (define (attach1toeachword document-line-kv-pair)
    (map (lambda (wd-in-line)
           (make-kv-pair wd-in-line 1))
         (kv-value document-line-kv-pair)))
attach1toeachword

;; Invoke mapreduce as a distributed word count
> (define wordcounts
    (mapreduce attach1toeachword + 0 "/docs"))
wordcounts

;; Display the elements of the output stream
> (show-stream wordcounts 5)
((or . 1) (be . 2) (not . 1) (wit . 1) (to . 3))

;; Query the stream for the count of the word "be"
> (kv-value (stream-car (stream-filter
    (lambda (kv) (equal? (kv-key kv) 'be)) wordcounts)))
2
```