

reducer.py

```
#!/usr/bin/env python
"""reducer.py"""
```

```
import sys
```

```
current_word = None
current_count = 0
word = None
```

```
for line in sys.stdin:
    line = line.strip()
    word, count = line.split('\t', 1)
```

```
    try:
        count = int(count)
    except ValueError:
        continue
```

```
    if current_word == word:
        current_count += count
    else:
        if current_word:
            print '%s\t%s' % (current_word, current_count)
            current_count = count
            current_word = word
```

```
    if current_word == word:
        print '%s\t%s' % (current_word, current_count)
```

mapper.py

```
#!/usr/bin/python
"""mapper.py"""
```

```
import sys
```

```
for line in sys.stdin:
    line = line.strip()
    words = line.split()
    for word in words:
        print '%s\t%s' % (word, 1)
```

steps:

1. `gedit mapper.py`
2. `gedit word.txt`
3. `cat word.txt | python mapper.py`
4. `gedit reducer.py`
5. `cat word.txt | python mapper.py | sort | python reducer.py`