

프로그래밍 과제 W02

1. 입력으로 제공된 텍스트 파일 `pride_and_prejudice.txt`를 읽어서 이 파일에 등장하는 길이가 7이상인 모든 단어의 목록과 각 단어의 등장 빈도를 구하여 화면으로 출력하는 프로그램을 작성하라.

- 프로그램 내에서 단어들은 강의 슬라이드 16페이지와 같은 자료구조로 저장되어야 한다.
- 대소문자의 구분은 없다. 출력에서 모든 단어들은 소문자로 구성되어야 한다.
- 단어들은 사전식 순서로 정렬되어 출력되어야 한다.
- 영문 알파벳이 하나도 포함되지 않은 문자열은 무시되어야 한다. 또한 단어의 앞과 뒤에 붙은 알파벳이 아닌 기호들은 제거되어야 한다. 예를 들어 쉼표, 마침표, 물음표 등은 제거되어야 한다.
- 단, 단어의 중간에 삽입된 따옴표(')와 하이픈(-)은 단어의 일부로 간주한다. 예를 들어 `o'clock`이나 `brother's` 혹은 `breakfast-parlour` 등은 그 자체로 하나의 단어로 취급한다. 하지만 단어의 중간에 따옴표(')와 하이픈(-)을 제외한 다른 기호나 혹은 따옴표(')와 하이픈(-)이라고 하더라도 두 개 이상 연속될 경우에는 그것을 기준으로 문자열을 분리하여 2개의 별개의 단어로 취급한다. 예를 들어 `enough--one`은 `enough`와 `one`이라는 2개의 단어로 취급한다. 가령 `half-an-hour--was`는 `half-an-hour` 와 `was`라는 2개의 단어로 취급한다.
- 출력은 화면으로 한다. 한 줄에 하나의 단어와 그 단어의 등장 빈도를 출력하라. 단어와 등장빈도 사이에 한 칸을 띄워라. 모든 단어를 출력하지 말고 10개씩 건너뛰어서 출력한다. 즉, 정렬된 순서에서 1번째 단어, 11번째 단어, 21번째 단어,... 이런 식으로 출력한다.

출력 예:

```
absolutely 2
acquaint 1
adjusting 1
against 2
amusement 3
apothecary 1
arrived 1
astonished 1
authorized 1
bennet's 5
...
```