

2024 자료구조

[과제 4 정렬]



과목명 : 자료구조

학과 : 소프트웨어학과

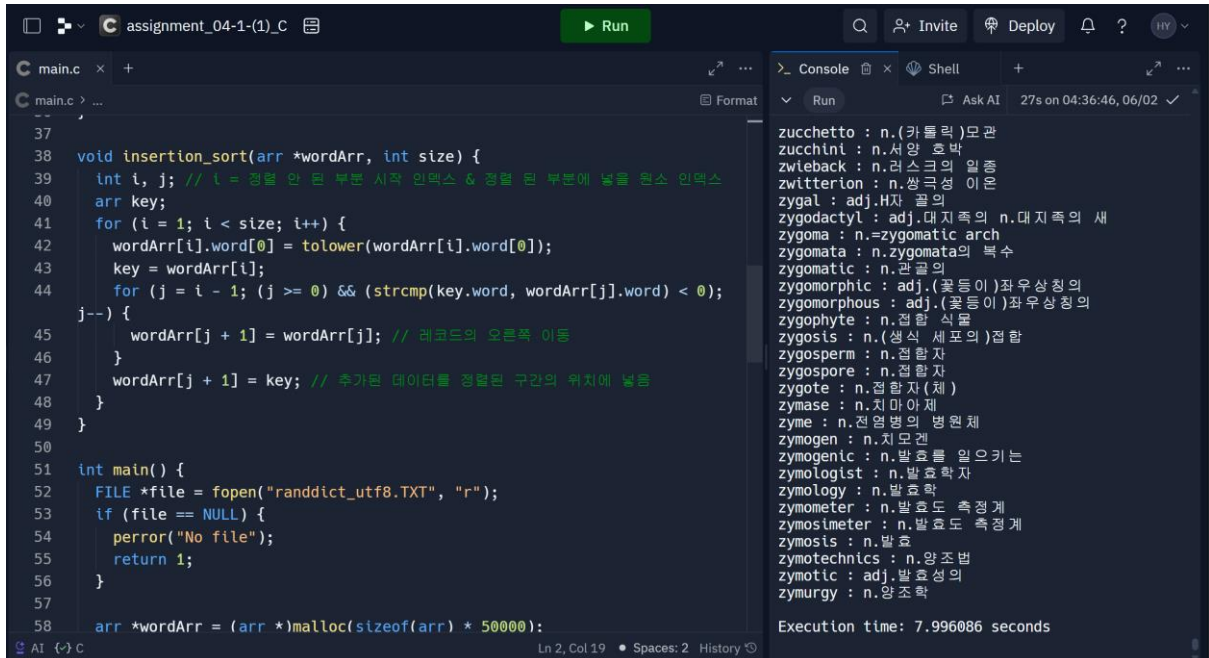
학번 : 20231371

이름 : 김혜린

I 사전 파일 정렬(1)

: 사전 데이터가 저장된 (순서 없음) 배열을 단어순으로 정렬

(1)_C 삽입 정렬 : 실행시간 = 7.996086s



The screenshot shows a C program in a code editor. The program reads a file named 'randdict_utf8.TXT' and sorts its contents using an insertion sort algorithm. The console output displays a list of words and their definitions, sorted alphabetically. The execution time is 7.996086 seconds.

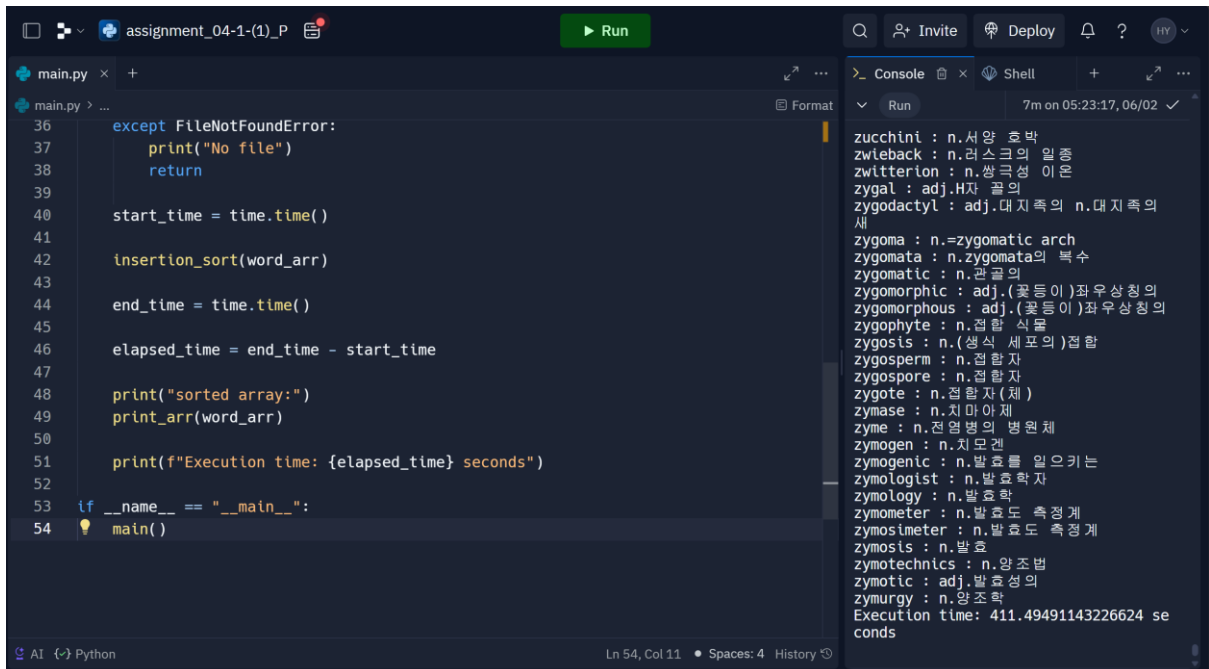
```
37
38 void insertion_sort(arr *wordArr, int size) {
39     int i, j; // i = 정렬 안 된 부분 시작 인덱스 & 정렬 된 부분에 넣을 원소 인덱스
40     arr key;
41     for (i = 1; i < size; i++) {
42         wordArr[i].word[0] = tolower(wordArr[i].word[0]);
43         key = wordArr[i];
44         for (j = i - 1; (j >= 0) && (strcmp(key.word, wordArr[j].word) < 0); j--) {
45             wordArr[j + 1] = wordArr[j]; // 레코드의 오른쪽 이동
46         }
47         wordArr[j + 1] = key; // 추가된 데이터를 정렬된 구간의 위치에 넣음
48     }
49 }
50
51 int main() {
52     FILE *file = fopen("randdict_utf8.TXT", "r");
53     if (file == NULL) {
54         perror("No file");
55         return 1;
56     }
57
58     arr *wordArr = (arr *)malloc(sizeof(arr) * 50000);
```

Console output:

```
zucchetto : n.(카톨릭)모관
zucchini : n.서양 호박
zwieback : n.러스크의 일종
zwitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
```

Execution time: 7.996086 seconds

(1)_P 삽입 정렬 : 실행시간 = 411.494911 s



The screenshot shows a Python program in a code editor. The program reads a file named 'randdict_utf8.TXT' and sorts its contents using an insertion sort algorithm. The console output displays a list of words and their definitions, sorted alphabetically. The execution time is 411.494911 seconds.

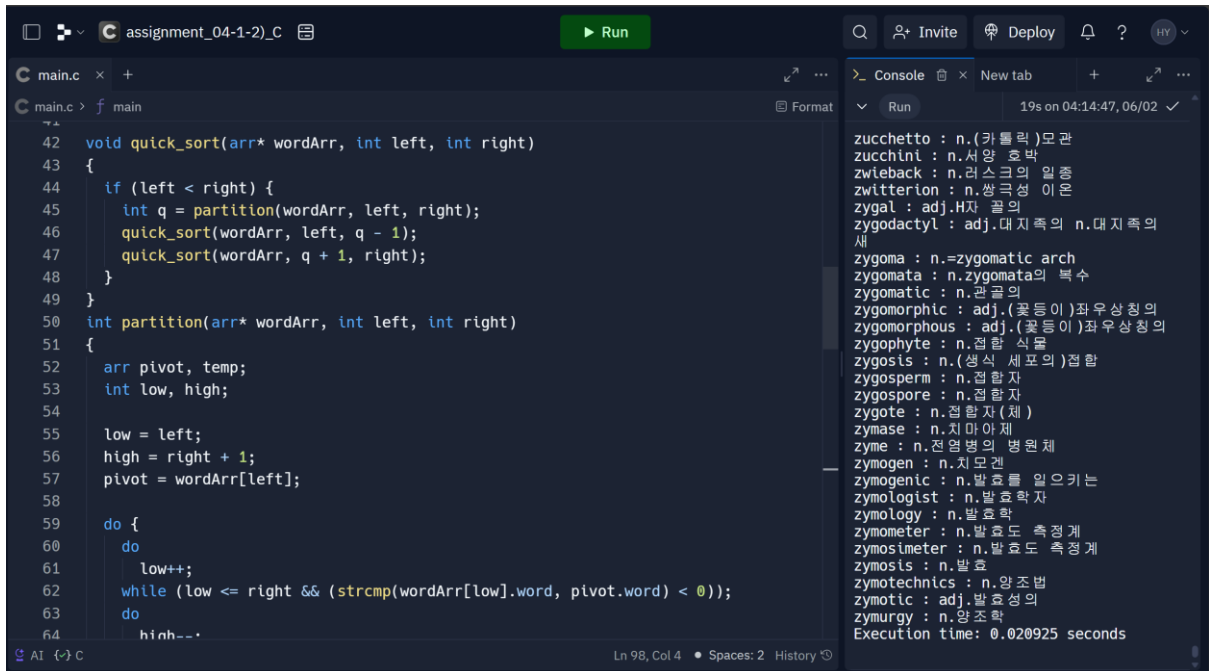
```
36 except FileNotFoundError:
37     print("No file")
38     return
39
40 start_time = time.time()
41
42 insertion_sort(word_arr)
43
44 end_time = time.time()
45
46 elapsed_time = end_time - start_time
47
48 print("sorted array:")
49 print_arr(word_arr)
50
51 print(f"Execution time: {elapsed_time} seconds")
52
53 if __name__ == "__main__":
54     main()
```

Console output:

```
zucchini : n.서양 호박
zwieback : n.러스크의 일종
zwitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
```

Execution time: 411.49491143226624 seconds

(2)_C 퀵 정렬 : 실행시간 = 0.020925 s

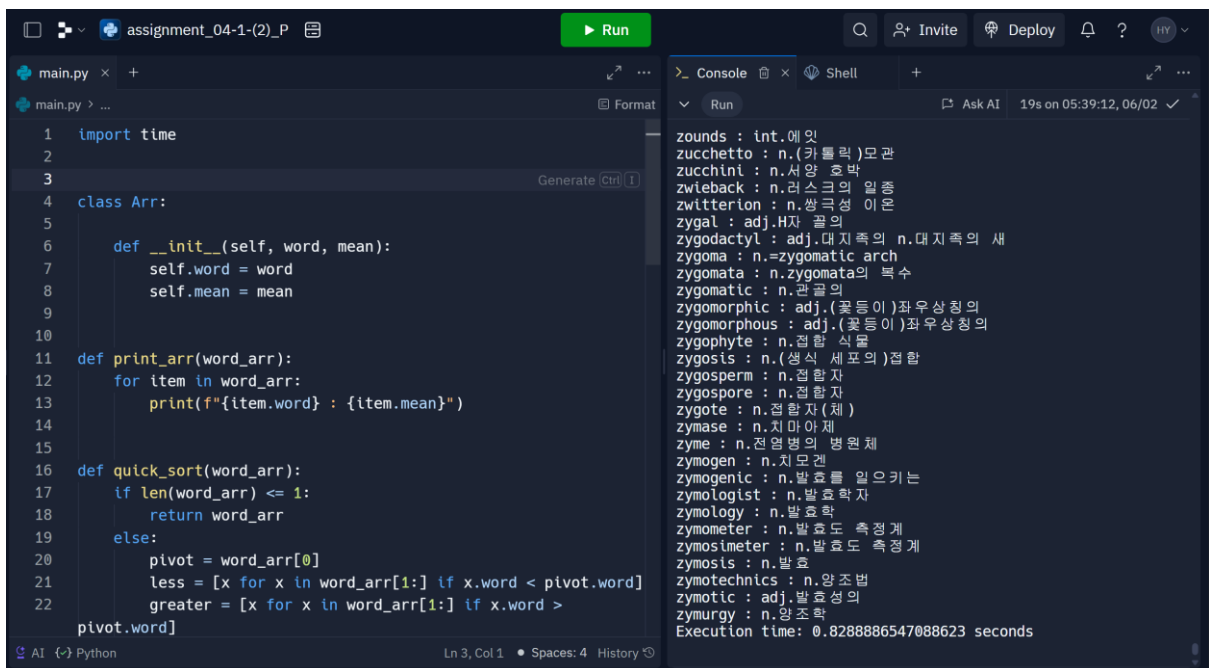


```
main.c x +
main.c > f main
42 void quick_sort(arr* wordArr, int left, int right)
43 {
44     if (left < right) {
45         int q = partition(wordArr, left, right);
46         quick_sort(wordArr, left, q - 1);
47         quick_sort(wordArr, q + 1, right);
48     }
49 }
50 int partition(arr* wordArr, int left, int right)
51 {
52     arr pivot, temp;
53     int low, high;
54
55     low = left;
56     high = right + 1;
57     pivot = wordArr[left];
58
59     do {
60         do
61             low++;
62         while (low <= right && (strcmp(wordArr[low].word, pivot.word) < 0));
63     }
64     high--;
```

Console

zucchetto : n.(카톨릭)모관
zucchini : n.서양 호박
zwieback : n.라스크의 일종
twitterton : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
Execution time: 0.020925 seconds

(2)_P 퀵 정렬 : 실행시간 = 0.828888 s

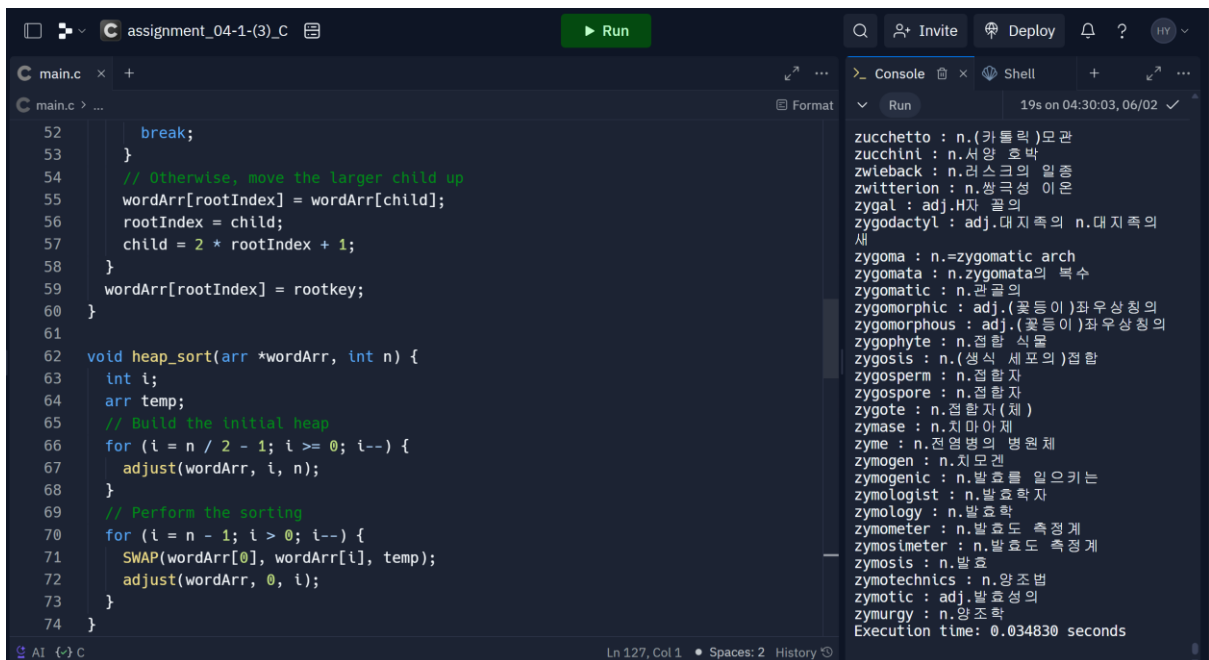


```
main.py x +
main.py > ...
1 import time
2
3
4 class Arr:
5
6     def __init__(self, word, mean):
7         self.word = word
8         self.mean = mean
9
10
11 def print_arr(word_arr):
12     for item in word_arr:
13         print(f"{item.word} : {item.mean}")
14
15
16 def quick_sort(word_arr):
17     if len(word_arr) <= 1:
18         return word_arr
19     else:
20         pivot = word_arr[0]
21         less = [x for x in word_arr[1:] if x.word < pivot.word]
22         greater = [x for x in word_arr[1:] if x.word >
                pivot.word]
```

Console

zounds : int.에잇
zucchetto : n.(카톨릭)모관
zucchini : n.서양 호박
zwieback : n.라스크의 일종
twitterton : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
Execution time: 0.828886547088623 seconds

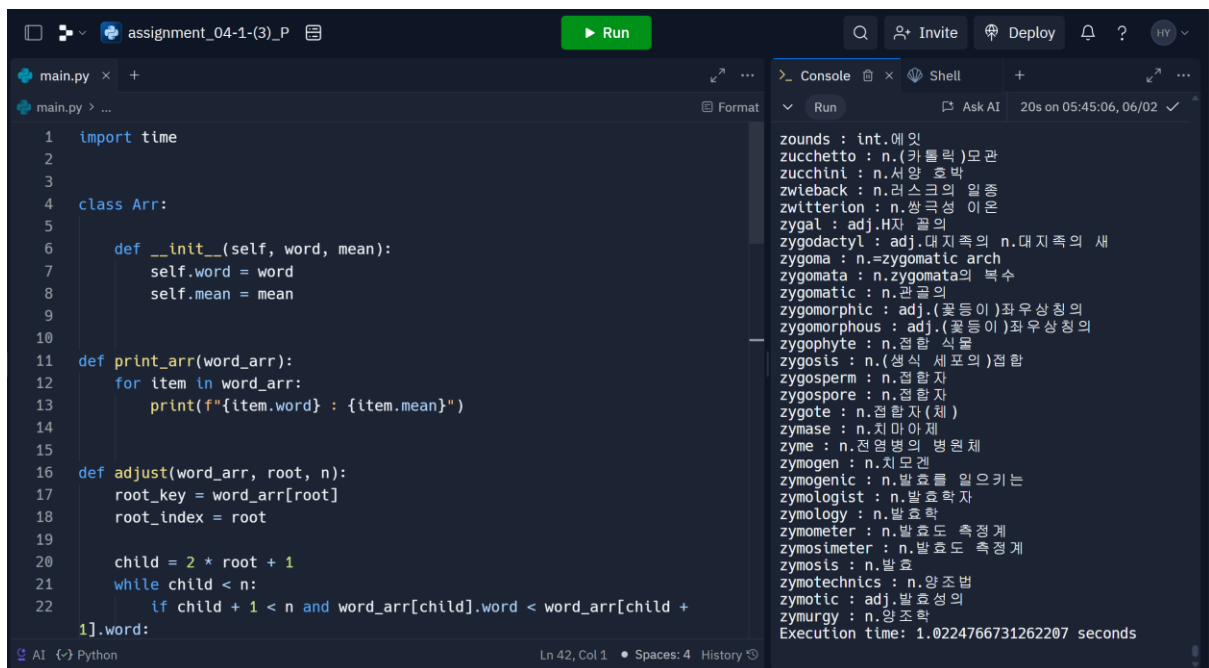
(3)_C 힙 정렬 : 실행시간 = 0.034830 s



```
52     break;
53 }
54 // Otherwise, move the larger child up
55 wordArr[rootIndex] = wordArr[child];
56 rootIndex = child;
57 child = 2 * rootIndex + 1;
58 }
59 wordArr[rootIndex] = rootkey;
60 }
61
62 void heap_sort(arr *wordArr, int n) {
63     int i;
64     arr temp;
65     // Build the initial heap
66     for (i = n / 2 - 1; i >= 0; i--) {
67         adjust(wordArr, i, n);
68     }
69     // Perform the sorting
70     for (i = n - 1; i > 0; i--) {
71         SWAP(wordArr[0], wordArr[i], temp);
72         adjust(wordArr, 0, i);
73     }
74 }
```

zucchetto : n.(카톨릭)모관
zucchini : n.서양 호박
zwieback : n.리스크의 일종
twitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
Execution time: 0.034830 seconds

(3)_P 힙 정렬 : 실행시간 = 1.022476 s



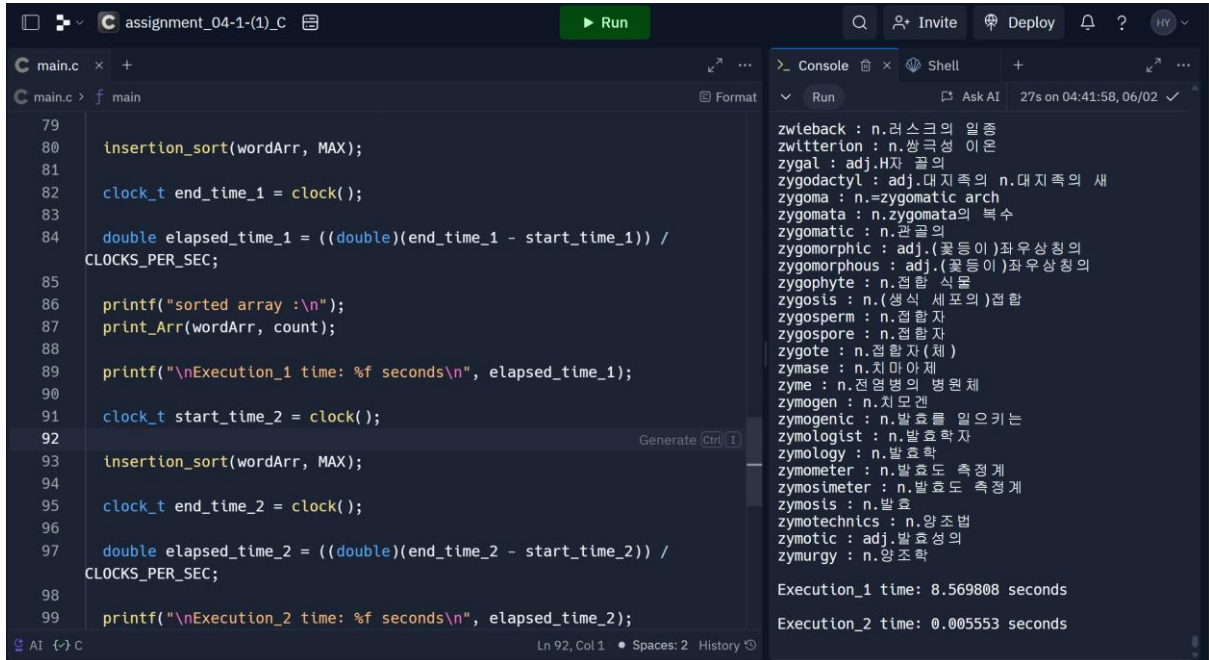
```
1 import time
2
3
4 class Arr:
5
6     def __init__(self, word, mean):
7         self.word = word
8         self.mean = mean
9
10
11 def print_arr(word_arr):
12     for item in word_arr:
13         print(f"{item.word} : {item.mean}")
14
15
16 def adjust(word_arr, root, n):
17     root_key = word_arr[root].word
18     root_index = root
19
20     child = 2 * root + 1
21     while child < n:
22         if child + 1 < n and word_arr[child].word < word_arr[child + 1].word:
```

sounds : int.에잇
zucchetto : n.(카톨릭)모관
zucchini : n.서양 호박
zwieback : n.리스크의 일종
twitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
Execution time: 1.0224766731262207 seconds

II 사전 파일 정렬(2)

:I 에서 정렬된 데이터(이미 정렬된 데이터)를 다시 정렬할 때 걸리는 시간 측정

(1) 삽입 정렬 : 다시 정렬했을 때의 실행시간 = 0.005553 s

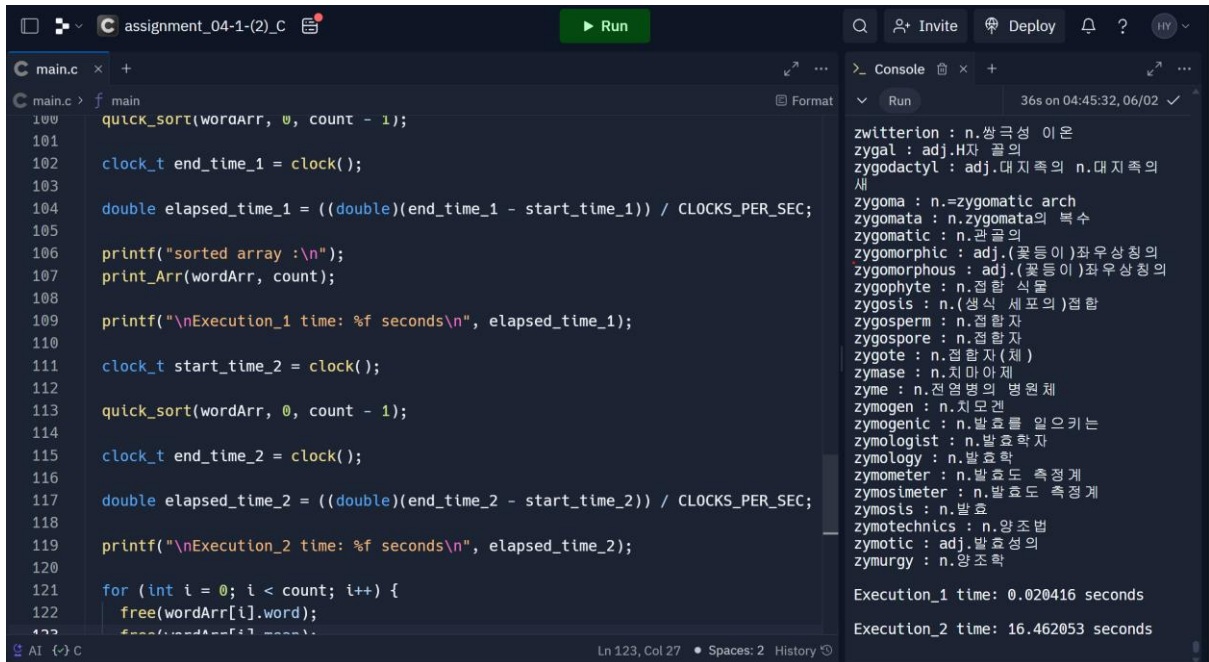


The screenshot shows a C program in a code editor. The program uses `insertion_sort` on an array `wordArr` of size `MAX`. It measures the execution time for two consecutive sorts. The first sort takes 8.569808 seconds, and the second sort takes 0.005553 seconds. The console output on the right shows the sorted array and the execution times.

```
79
80 insertion_sort(wordArr, MAX);
81
82 clock_t end_time_1 = clock();
83
84 double elapsed_time_1 = ((double)(end_time_1 - start_time_1)) /
CLOCKS_PER_SEC;
85
86 printf("sorted array :\n");
87 print_Arr(wordArr, count);
88
89 printf("\nExecution_1 time: %f seconds\n", elapsed_time_1);
90
91 clock_t start_time_2 = clock();
92
93 insertion_sort(wordArr, MAX);
94
95 clock_t end_time_2 = clock();
96
97 double elapsed_time_2 = ((double)(end_time_2 - start_time_2)) /
CLOCKS_PER_SEC;
98
99 printf("\nExecution_2 time: %f seconds\n", elapsed_time_2);
```

Execution_1 time: 8.569808 seconds
Execution_2 time: 0.005553 seconds

(2) 퀵 정렬 : 다시 정렬했을 때의 실행시간 = 16.462053 s

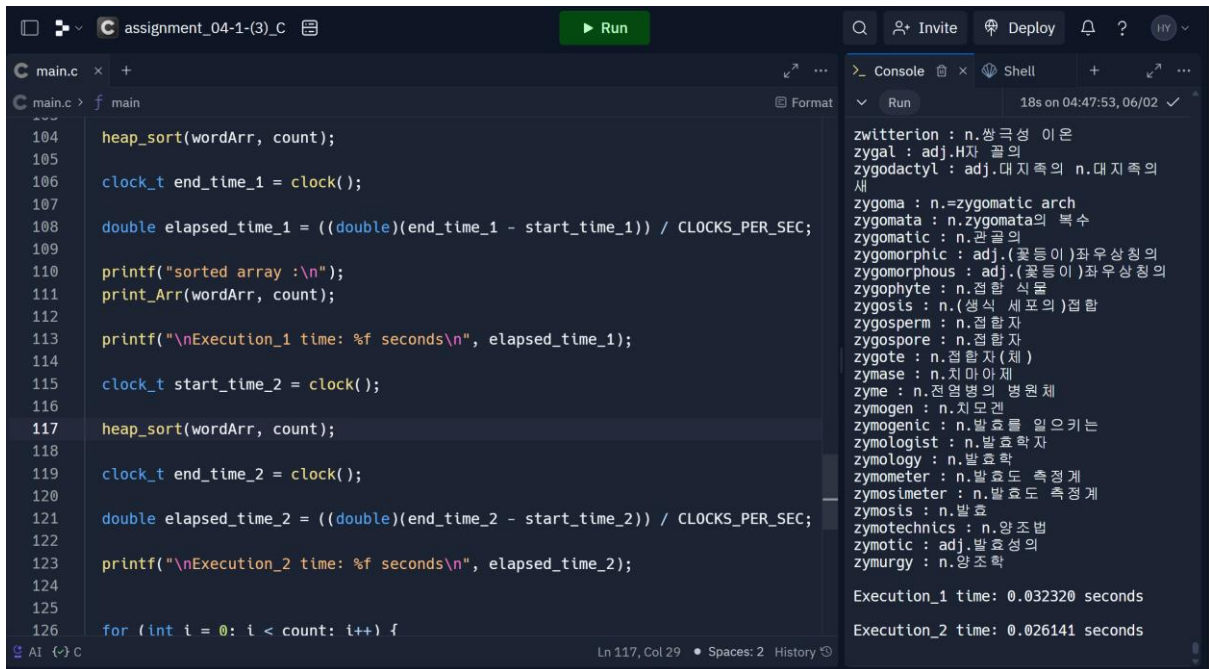


The screenshot shows a C program in a code editor. The program uses `quick_sort` on an array `wordArr` of size `count - 1`. It measures the execution time for two consecutive sorts. The first sort takes 0.020416 seconds, and the second sort takes 16.462053 seconds. The console output on the right shows the sorted array and the execution times.

```
100 quick_sort(wordArr, 0, count - 1);
101
102 clock_t end_time_1 = clock();
103
104 double elapsed_time_1 = ((double)(end_time_1 - start_time_1)) / CLOCKS_PER_SEC;
105
106 printf("sorted array :\n");
107 print_Arr(wordArr, count);
108
109 printf("\nExecution_1 time: %f seconds\n", elapsed_time_1);
110
111 clock_t start_time_2 = clock();
112
113 quick_sort(wordArr, 0, count - 1);
114
115 clock_t end_time_2 = clock();
116
117 double elapsed_time_2 = ((double)(end_time_2 - start_time_2)) / CLOCKS_PER_SEC;
118
119 printf("\nExecution_2 time: %f seconds\n", elapsed_time_2);
120
121 for (int i = 0; i < count; i++) {
122     free(wordArr[i].word);
123 }
```

Execution_1 time: 0.020416 seconds
Execution_2 time: 16.462053 seconds

(3) 힙 정렬 : 다시 정렬했을 때의 실행시간 = 0.026141 s



```
main.c x +
main.c > f main
104 heap_sort(wordArr, count);
105
106 clock_t end_time_1 = clock();
107
108 double elapsed_time_1 = ((double)(end_time_1 - start_time_1)) / CLOCKS_PER_SEC;
109
110 printf("sorted array :\n");
111 print_Arr(wordArr, count);
112
113 printf("\nExecution_1 time: %f seconds\n", elapsed_time_1);
114
115 clock_t start_time_2 = clock();
116
117 heap_sort(wordArr, count);
118
119 clock_t end_time_2 = clock();
120
121 double elapsed_time_2 = ((double)(end_time_2 - start_time_2)) / CLOCKS_PER_SEC;
122
123 printf("\nExecution_2 time: %f seconds\n", elapsed_time_2);
124
125
126 for (int i = 0; i < count; i++) {
```

Console

Run

18s on 04:47:53, 06/02 ✓

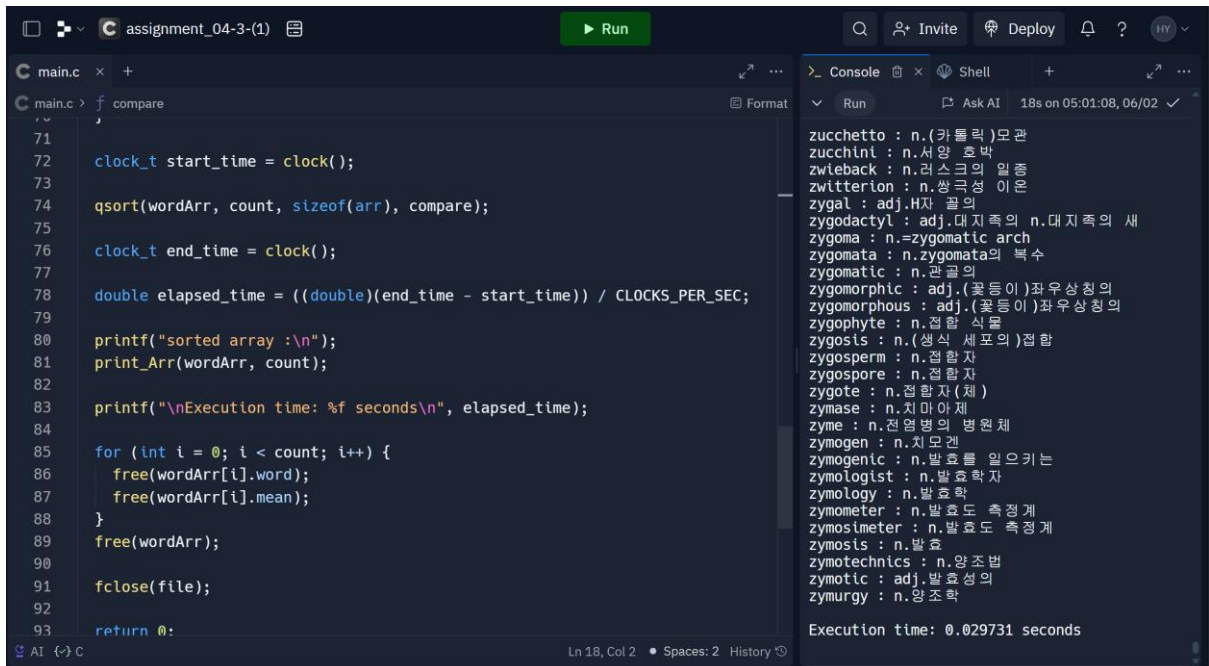
twitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygosperm : n.접합자
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학

Execution_1 time: 0.032320 seconds
Execution_2 time: 0.026141 seconds

Ⅲ 내장된 정렬 방법과 비교

: C와 Python에 내장된 정렬 함수로 정렬해보고 시간 밀리초단위로 측정

(1) C의 내장된 정렬 함수 qsort를 이용하여 정렬 : 실행시간 = 0.029731 s



```
main.c x +
main.c > f compare
71
72 clock_t start_time = clock();
73
74 qsort(wordArr, count, sizeof(arr), compare);
75
76 clock_t end_time = clock();
77
78 double elapsed_time = ((double)(end_time - start_time)) / CLOCKS_PER_SEC;
79
80 printf("sorted array :\n");
81 print_Arr(wordArr, count);
82
83 printf("\nExecution time: %f seconds\n", elapsed_time);
84
85 for (int i = 0; i < count; i++) {
86     free(wordArr[i].word);
87     free(wordArr[i].mean);
88 }
89 free(wordArr);
90
91 fclose(file);
92
93 return 0;
```

Console

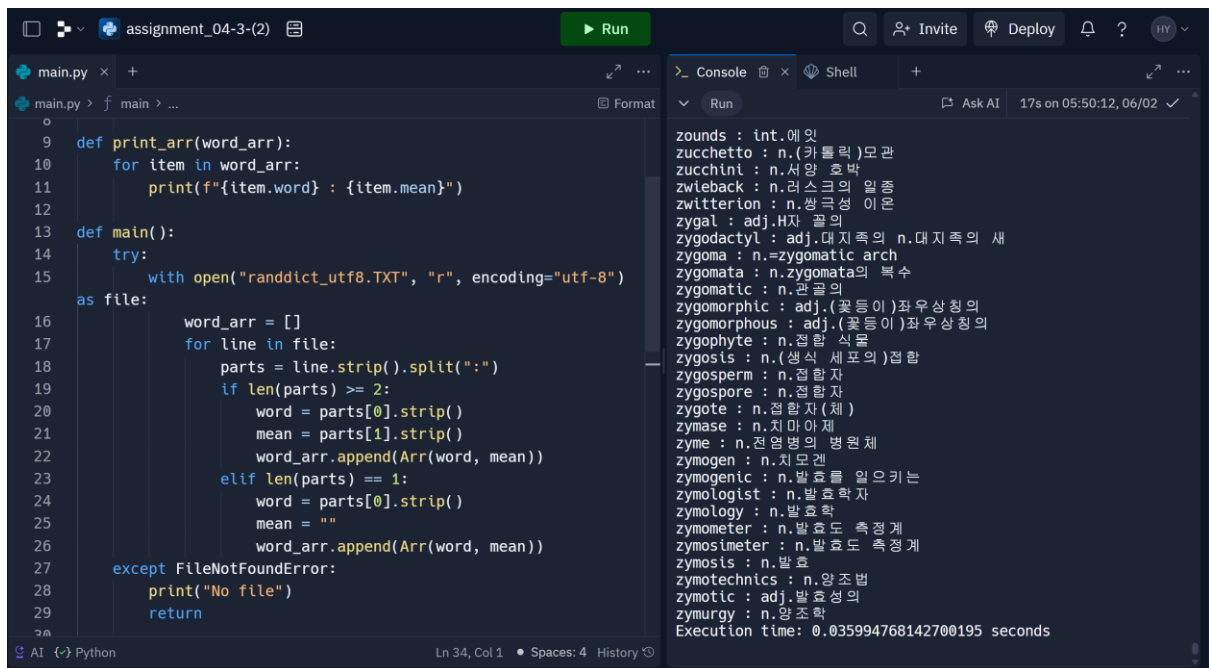
Run

18s on 05:01:08, 06/02 ✓

zucchetto : n.(카톨릭)모관
zucchini : n.서양 호박
zwieback : n.러스크의 일종
twitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygosperm : n.접합자
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학

Execution time: 0.029731 seconds

(2) Python의 내장된 정렬 함수 sorted를 이용하여 정렬 : 실행시간 = 0.0359947 s



```
main.py x +
main.py > f main > ...
9 def print_arr(word_arr):
10     for item in word_arr:
11         print(f"{item.word} : {item.mean}")
12
13 def main():
14     try:
15         with open("randdict_utf8.TXT", "r", encoding="utf-8")
16     as file:
17         word_arr = []
18         for line in file:
19             parts = line.strip().split(":")
20             if len(parts) >= 2:
21                 word = parts[0].strip()
22                 mean = parts[1].strip()
23                 word_arr.append(Arr(word, mean))
24             elif len(parts) == 1:
25                 word = parts[0].strip()
26                 mean = ""
27                 word_arr.append(Arr(word, mean))
28     except FileNotFoundError:
29         print("No file")
30     return
```

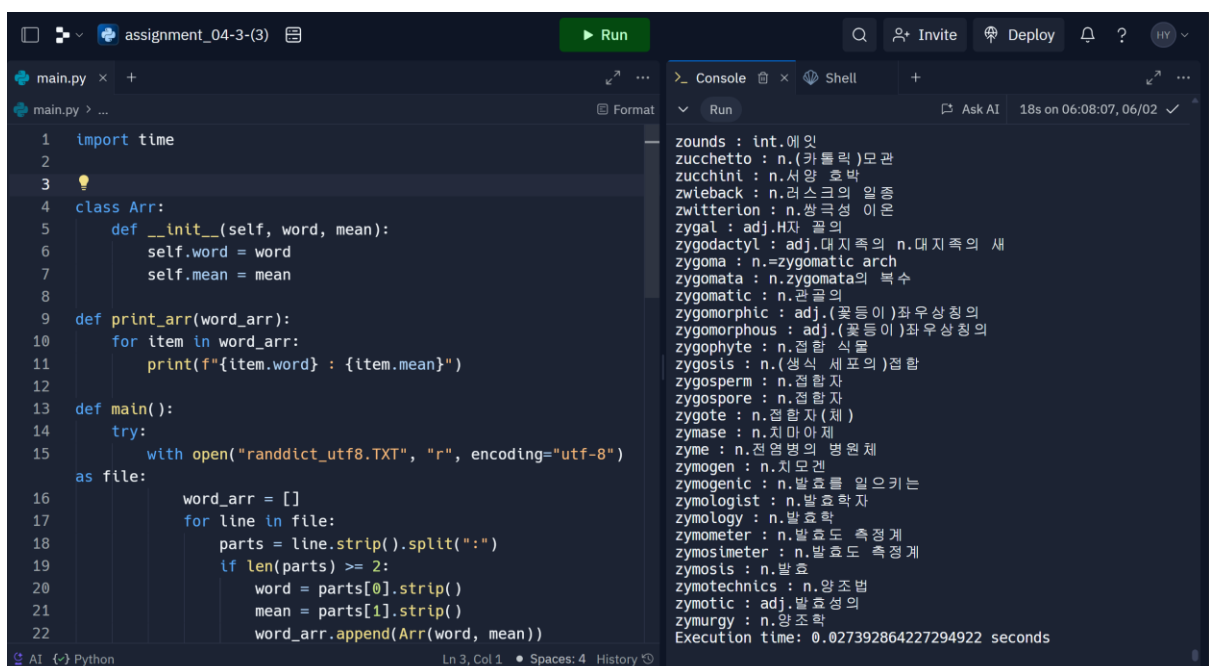
Console output:

```
zounds : int.에잇
zucchetto : n.(카를릭 )모관
zucchini : n.서양 호박
zwieback : n.러스크의 일종
zwitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
Execution time: 0.035994768142700195 seconds
```

(3) 파이썬에서 더 빠른 정렬 결과를 얻는 방법 : Timsort

=> 실행시간 = 0.027392 s

Timsort는 Python의 정렬 알고리즘 중 하나로, 병합정렬과 삽입정렬을 결합한 하이브리드 알고리즘이다. Python의 list.sort() 메서드는 Timsort를 기본 정렬 알고리즘으로 사용한다. 다만, 이 방법은 리스트의 정렬 결과를 직접 수정하므로 주의해야 한다.



```
main.py x +
main.py > ...
1 import time
2
3
4 class Arr:
5     def __init__(self, word, mean):
6         self.word = word
7         self.mean = mean
8
9     def print_arr(word_arr):
10         for item in word_arr:
11             print(f"{item.word} : {item.mean}")
12
13 def main():
14     try:
15         with open("randdict_utf8.TXT", "r", encoding="utf-8")
16     as file:
17         word_arr = []
18         for line in file:
19             parts = line.strip().split(":")
20             if len(parts) >= 2:
21                 word = parts[0].strip()
22                 mean = parts[1].strip()
23                 word_arr.append(Arr(word, mean))
```

Console output:

```
zounds : int.에잇
zucchetto : n.(카를릭 )모관
zucchini : n.서양 호박
zwieback : n.러스크의 일종
zwitterion : n.쌍극성 이온
zygal : adj.H자 골의
zygodactyl : adj.대지족의 n.대지족의 새
zygoma : n.=zygomatic arch
zygomata : n.zygomata의 복수
zygomatic : n.관골의
zygomorphic : adj.(꽃등이)좌우상칭의
zygomorphous : adj.(꽃등이)좌우상칭의
zygophyte : n.접합 식물
zygosis : n.(생식 세포의)접합
zygospore : n.접합자
zygote : n.접합자(체)
zymase : n.치마아제
zyme : n.전염병의 병원체
zymogen : n.치모겐
zymogenic : n.발효를 일으키는
zymologist : n.발효학자
zymology : n.발효학
zymometer : n.발효도 측정계
zymosimeter : n.발효도 측정계
zymosis : n.발효
zymotechnics : n.양조법
zymotic : adj.발효성의
zymurgy : n.양조학
Execution time: 0.027392864227294922 seconds
```