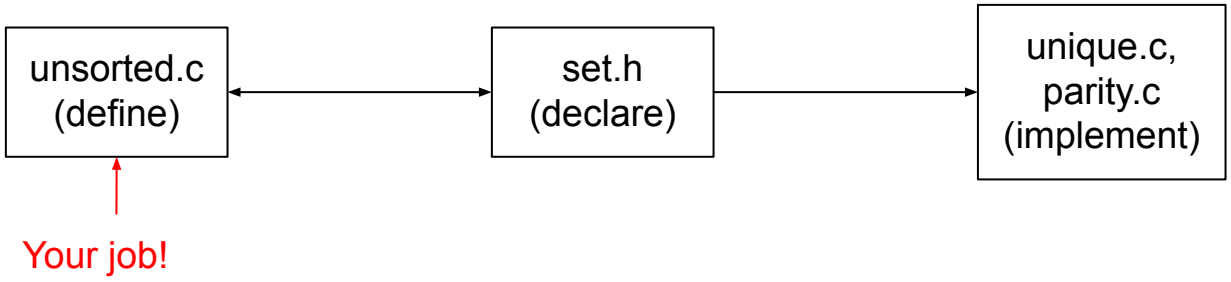
Sets and Arrays

Project 2

Unsorted Array (1st week)

- Download the tar file and untar it with the command:
  tar xvf project2.tar
- Header file (set.h): declaration of the set functions; the linkage between c files.
- Driver programs (unique.c, parity.c) call set functions by including set.h:
  - unique.c: collect unique words.
  - parity.c: collect words that occured odd times.



Functions

- `SET *createSet(int maxElts);`
  return a pointer to a new set with a maximum capacity of *maxElts*
- `void destroySet(SET *sp);`
  deallocate memory associated with the set pointed to by *sp*
- `int numElements(SET *sp);`
  return the number of elements in the set pointed to by *sp*
- `void addElement(SET *sp, char *elt);`
  add *elt* to the set pointed to by *sp*
- `void removeElement(SET *sp, char *elt);`
  remove *elt* from the set pointed to by *sp*
- `char *findElement(SET *sp, char *elt);`
  if *elt* is present in the set pointed to by *sp* then return the matching element, otherwise return `NULL`
- `char **getElements(SET *sp);`
  allocate and return an array of elements in the set pointed to by *sp*

Tricky Parts

- Pointer pointing to pointers
  - Char **
  - Fundamentally, array is a pointer
- search (internal) vs findElement (external)
  - What if not found?
- strdup vs strcpy
  - Strdup returns a new pointer
  - Strcpy works with malloc and copies to the new pointer
- removeElement
  - shifting?
- getElements
  - Memcpy or a loop
- Count (# of elements) vs length (max length)
- Assert pointers != null

Testing Commands

- Compile:

  gcc -o unique_unsorted unique.c unsorted.c

  gcc -o parity_unsorted parity.c unsorted.c

- Execute (always try to debug by yourself first):

  (1) Testing addElement:

  ./unique_unsorted /scratch/coen12/Macbeth.txt

  (2) Testing removeElement if (1) works:

  ./unique_unsorted /scratch/coen12/Macbeth.txt /scratch/coen12/Bible.txt

  (3) Testing getElements if (1) & (2) works:

  ./unique_unsorted -l /scratch/coen12/Macbeth.txt /scratch/coen12/Bible.txt

  (4) Testing findElement if (1) & (2) works:

  ./parity_unsorted /scratch/coen12/Macbeth.txt

Makefiles

- Used to document complex compilation instructions
- Structure
  - target: dependencies
    - compile command
- Example

```
unique_unsorted: unique.c unsorted.c
	gcc -o unique_unsorted unique.c unsorted.c
```

Demo

```
[tzhou@linux10623 demo]$ make clean all
rm -f unique-unsorted unique-sorted parity-unsorted parity-sorted *.o core
gcc -g -Wall   -c -o unique.o unique.c
gcc -g -Wall   -c -o unsorted.o unsorted.c
gcc -o unique-unsorted  unique.o unsorted.o
gcc -g -Wall   -c -o sorted.o sorted.c
gcc -o unique-sorted  unique.o sorted.o
gcc -g -Wall   -c -o parity.o parity.c
gcc -o parity-unsorted  parity.o unsorted.o
gcc -o parity-sorted  parity.o sorted.o
[tzhou@linux10623 demo]$ ./unique-unsorted /scratch/coen12/Macbeth.txt
18464 total words
3894 distinct words
[tzhou@linux10623 demo]$ ./unique-unsorted /scratch/coen12/Macbeth.txt /scratch/coen12/Bible.txt
18464 total words
3894 distinct words
1641 remaining words
```

Demo

```
[tzhou@linux10623 demo]$ ./unique-unsorted -l /scratch/coen12/Macbeth.txt /scratch/coen12/Bible.txt
SCENE
Scotland
Whateer
England
ACT
harbingers
Direness
enrages
underwrit
forsworn
frog
theres
interdiction
Witches
FIRST
WITCH
artificial
speculative
Fate
stared
Pronounce
END
```

Demo

```
profound
Didst
mated
scannd
scream
crickets
howld
Ay
rabbles
sights
Looks
lated
bleed
s
methinks
towering
Ist
addressd
brows
unfortunate
diggd
[tzhou@linux10623 demo]$ ./parity-unsorted /scratch/coen12/Macbeth.txt
18464 total words
2900 words occur an odd number of times
```

report.txt

- Record the “real” execution time for your programs.

```
time ./unique /scratch/coen12/Macbeth.txt
```

```
[tzhou@linux10623 lab2]$ time ./unique-unsorted /scratch/coen12/GreenEggsAndHam.txt
803 total words
67 distinct words

real    0m0.014s
user    0m0.001s
sys     0m0.002s
```

report.txt

Don’t forget big O analysis on functions

```
unique
------
                                unsorted  sorted
GreenEggsAndHam.txt               0.014    ????
Macbeth.txt                       ????     ????
Genesis.txt                       ????     ????
HoundOfTheBaskervilles.txt        ????     ????
TheWarOfTheWorlds.txt             ????     ????
TreasureIsland.txt                ????     ????
TheSecretGarden.txt               ????     ????
TwentyThousandLeagues.txt         ????     ????
TheCountOfMonteCristo.txt         ????     ????
Bible.txt                         ????     ????


parity
------
                                unsorted  sorted
GreenEggsAndHam.txt               ????     ????
Macbeth.txt                       ????     ????
Genesis.txt                       ????     ????
HoundOfTheBaskervilles.txt        ????     ????
TheWarOfTheWorlds.txt             ????     ????
TreasureIsland.txt                ????     ????
TheSecretGarden.txt               ????     ????
TwentyThousandLeagues.txt         ????     ????
TheCountOfMonteCristo.txt         ????     ????
Bible.txt                         ????     ????
~
```