Function Pointers and Abstract Data Types

COS 217
Goals of Today’s Lecture

• Function pointers
  o Sorting an array of integers
  o Sorting an array of strings
  o Sorting an array of any type
    – Void pointers and casting
    – Pointers to functions

• Abstract Data Types
  o Making “array” an ADT
Sorting an Array of Integers

• **Example problem**
  - Input: array \( v \) of \( n \) integers
  - Output: array in sorted order, from smallest to largest

• **Many ways to sort, but three common aspects**
  - Comparison between any two elements
  - Exchange to reverse the order of two elements
  - Algorithm that makes comparisons and exchanges till done

• **Simple approach**
  - Go one by one through the \( n \) array elements
  - By the end of step \( i \), get \( i^{th} \) smallest value in element \( i \)
    - Compare element \( i \) with all elements after it
    - Swap values if the \( i^{th} \) element is larger
## Integer Sorting Example

<table>
<thead>
<tr>
<th>v[0] &gt; v[1]?</th>
<th>7</th>
<th>2</th>
<th>9</th>
<th>6</th>
<th>v[1] &gt; v[2]?</th>
<th>2</th>
<th>7</th>
<th>9</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes, swap</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>v[1] &gt; v[3]?:</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>v[0] &gt; v[2]?</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>Yes, swap</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>v[0] &gt; v[3]?</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
void sort(int *v, int n) {
    int i, j;

    for (i = 0; i < n; i++) {
        for (j = i+1; j < n; j++) {
            if (v[i] > v[j]) {
                int swap = v[i];
                v[i] = v[j];
                v[j] = swap;
            }
        }
    }
}
Sorting an Array of Strings

- **Data types are different**
  - Array elements are `char*`
  - Swap variable is `char*`

- **Comparison operator is different**
  - The greater-than (">") sign does not work
  - Need to use `strcmp()` function instead

```plaintext
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>“the”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>“quick”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>“brown”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>“fox”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```plaintext
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>“brown”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>“fox”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>“quick”</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>“the”</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
void sort(char *v[], int n) {
    int i, j;
    for (i = 0; i < n; i++) {
        for (j = i+1; j < n; j++) {
            if (strcmp(v[i], v[j]) > 0) {
                char* swap = v[i];
                v[i] = v[j];
                v[j] = swap;
            }
        }
    }
}
Creating a Generic Function

• Generic function
  o A single `sort()` function that works for all data types

• C’s notion of data types is getting in our way
  o We need to accept parameters in any type
    - `sort(int *v, int n)` is only good for integer arrays
    - `sort(char *v[], int n)` is only good for string arrays
  o We need to have local variables of any type
    - `int swap` is only good for swapping integers
    - `char* swap` is only good for swapping strings

• Different types need different comparison operators
  o Greater-than sign (">") is only good for numerical types
  o `strcmp()` is only good for strings
  o We need to be able to tell `sort()` what comparison function to use
Generalizing: Void Pointers

- **Generic pointers are the same as any other pointer**
  - Except they point to a variable **with no specific type**
  - Example: `void *datap = "CS217";`

- **Difference:**
  - Regular pointers: compilers "know" what they point to
  - `void` pointers: compilers "don’t know" what they point to

- **Common Uses:**
  - Abstract data types supporting **polymorphism**
  - Pass pointer to function that could be any of several types

* Allowing the same definitions to be used with different types of data
Void Pointers in Sort

• Function parameters
  o Input: array of pointers to some unknown type

    ```c
    void sort(void *v[], int n)
    ```

• Local swap variable
  o Pointer to some unknown type

    ```c
    void *swap = v[i];
    v[i] = v[j];
    v[j] = swap;
    ```

• But, what about the comparison step?
  o Need to be able to pass a function to sort
Casting: Explicit Type Conversions

• Casting
  o As if the expression were assigned to a variable of the specified type
  o E.g., \texttt{int *intp1} cast into void pointer by \((\texttt{void *}) \texttt{intp1}\)

• C does many implicit conversions
  o E.g., function \texttt{double sqrt(double)}
    – Can be called as \texttt{sqrt(2)};
    – Which is treated as \texttt{sqrt((double) 2)};

• Sometimes useful to make conversion explicit
  o Documentation: making implicit type conversions explicit
    – E.g., getting the integer part of a floating-point number
    – Done by \texttt{int_part = (int) float_number};
  o Control: overrule the compile by forcing conversions we want
    – E.g., getting the fractional part of a floating-point number
    – Done by \texttt{frac_part = f - (int) f};
void sort(void *v[], int n, 
    int (*compare)(void *datap1, void *datap2))
{
    int i, j;

    for (i = 0; i < n; i++) {
        for (j = i+1; j < n; j++) {
            if (*((compare)(v[i], v[j]) > 0) {
                void *swap = v[i];
                v[i] = v[i];
                v[i] = swap;
            }
        }
    }
}
#include <stdio.h>
#include <string.h>
#include "sort.h"

int main() {
    char* w[4] = {"the", "quick", "brown", "fox"};

    sort((void **) w, 4, (int (*)(void*,void*)) strcmp);
    ...
}

pointer to a function
#include <stdio.h>
#include "sort.h"

int CompareInts(void *datap1, void *datap2) {
  int *intp1 = (int *) datap1;
  int *intp2 = (int *) datap2;
  return (*intp1 - *intp2);
}

int main() {
  int* w[4];

  w[0] = malloc(sizeof(int));
  *(w[0]) = 7;
  ...
  sort((void **) w, 4, (int (*)(void*,void*))CompareInts);
  ...
}
Making “Array” an ADT

• Arrays in C are error prone
  o Access elements before the array starts (e.g., v[−1])
  o Access elements past the end of array (e.g., v[n])
  o Modify the variable that keeps track of size (e.g., n)

• Protect programmers with an array ADT
  o Create and delete an array
  o Get the current length
  o Read an array element
  o Append, replace, remove
  o Sort
Array ADT: Interface

array.h

client does not know implementation

typedef struct Array *Array_T;

extern Array_T Array_new(void);
extern void Array_free(Array_T array);

extern int Array_getLength(Array_T array);
extern void *Array_getData(Array_T array, int index);

extern void Array_append(Array_T array, void *datap);
extern void Array_replace(Array_T array, int index, void *datap);
extern void Array_remove(Array_T array, int index);

extern void Array_sort(Array_T array, int (*compare)(void *datap1, void *datap2));
#include "array.h"
#include <stdio.h>

int main() {
    Array_T array;
    int i;

    array = Array_new();
    Array_append(array, (void *) "COS217");
    Array_append(array, (void *) "IS");
    Array_append(array, (void *) "FUN");

    for (i = 0; i < Array_getLength(array); i++) {
        char *str = (char *) Array_getData(array, i);
        printf(str);
    }

    Array_free(array);

    return 0;
}
Client Using Array ADT: Integers

```c
#include "array.h"
#include <stdio.h>

int main() {
    Array_T array;
    int one=1, two=2, three=3;
    int i;

    array = Array_new();

    Array_append(array, (void *) &one);
    Array_append(array, (void *) &two);
    Array_append(array, (void *) &three);

    for (i = 0; i < Array_getLength(array); i++) {
        int *datap = (int *) Array_getData(array, i);
        printf("%d ", *datap);
    }

    Array_free(array);

    return 0;
}
```
#include "array.h"

#define MAX_ELEMENTS 128

struct Array {
    void *elements[MAX_ELEMENTS];
    int num_elements;
};

Array_T Array_new(void) {
    Array_T array = malloc(sizeof(struct Array));
    array->num_elements = 0;
    return array;
}

void Array_free(Array_T array) {
    free(array);
}
int Array_getLength(Array_T array) {
    return array->num_elements;
}

void *Array_getData(Array_T array, int index) {
    return array->elements[index];
}

void Array_append(Array_T array, void *datap) {
    int index = array->num_elements;
    array->elements[index] = datap;
    array->num_elements++;
}

void Array_replace(Array_T array, int index, void *datap) {
    array->elements[index] = datap;
}
void Array_insert(Array_T array, int index, void *datap) {
    int i;
    /* Shift elements to the right to make room for new entry */
    for (i = array->num_elements; i > index; i--)
        array->elements[i] = array->elements[i-1];
    /* Add the new element in the now-free location */
    array->elements[index] = str;
    array->num_elements++;
}

void Array_remove(Array_T array, int index) {
    int i;
    /* Shift elements to the left to overwrite freed spot */
    for (i = index+1; i < array->num_elements; i++)
        array->elements[i-1] = array->elements[i];
    array->num_elements--;
}
void Array_sort(Array_T array,
    int (*compare)(void *datap1, void *datap2)) {
    int i, j;
    for (i = 0; i < array->num_elements; i++) {
        for (j = i+1; j < array->num_elements; j++) {
            if ((*compare)(array->elements[i], array->elements[j]) > 0) {
                void *swap = array->elements[i];
                array->elements[i] = array->elements[j];
                array->elements[j] = swap;
            }
        }
    }
}
Stupid Programmer Tricks

• qsort takes int (*compar)(const void *, const void *)
  o Comparison function returns integer greater than, equal, less than zero if first argument is greater than, equal, less than second

• Common approach:

```
int
ItemCompare(const void *pA, const void *pB)
{
    Item *a = pA, *b = pB;
    return(a->field - b->field);
}
```

• Bad idea when field is float or “long long” (64 bit)
Summary

• Module supporting operations on single data structure
  o Interface declares operations, not data structure
  o Interface provides access to simple, complete set of operations
  o Interface provides flexibility and extensibility

• Trick is providing functionality AND generality
  o Take advantage of features of programming language
    – void pointers
    – function pointers

• Advantages
  o Provide complete set of commonly used functions (re-use)
  o Implementation is hidden from client (encapsulation)
  o Can use for multiple types (polymorphism)