Arrays for Tabulation

Ryan Eberhardt
CS 106AJ
November 2, 2018
slides courtesy of Eric Roberts

Using Arrays for Tabulation

- Arrays turn out to be useful when you have a set of data values and need to count how many values fall into each of a set of ranges. This process is called *tabulation*.

- Tabulation uses arrays in a slightly different way from those applications that use them to store a list of data. When you implement a tabulation program, you use each data value to compute an index into an array of integers that keeps track of how many values fall into that category.

- The example of tabulation used in the text is a program that counts how many times each of the 26 letters appears in a sequence of text lines. Such a program would be very useful in decoding a letter-substitution ciphers, such as the one from Edgar Allan Poe's short story "The Gold Bug," which we showed in class this past Monday.

Poe's Cryptogram Puzzle

- The first step in solving Poe's cryptogram was to construct a table showing how often each character appeared in the plaintext message.
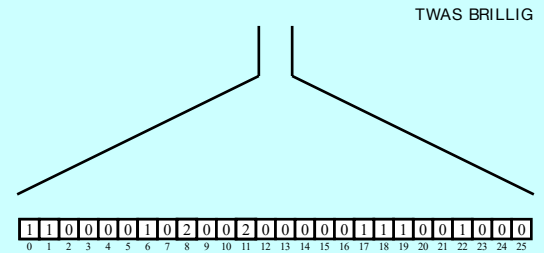
```
53‡‡†305))6*;4826)4‡•)4‡);806*;48†8¶
60))85;1‡(;:‡*8†83(88)5*†;46(;88*96*
?;8)*‡(;485);5*†2:*‡(;4956*2(5*−4)8¶
8*;4069285);)6†8)4‡‡;1(‡9;48081;8:8‡
1;48†85;4)485†528806*81(‡9;48;(88;4(
‡?34;48)4‡;161;:188;‡?;
```

8	33
;	26
4	19
‡	16
)	16
*	13
5	12
6	11
(	10
†	8
1	8
0	6
9	5
2	5
:	4
3	4
?	3
¶	2
−	1
.	1

- The problem here is to write a program that creates a letter-frequency table similar to the one that Poe used in his story.

Implementation Strategy

The basic idea behind the program to count letter frequencies is to use an array with 26 elements to keep track of how many times each letter appears. As the program reads the text, it increments the array element that corresponds to each letter.

TWAS BRILLIG

1	1	0	0	0	0	1	0	2	0	0	2	0	0	0	0	0	1	1	1	0	0	1	0	0	0
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

CountLetterFrequencies.js

```javascript
/*
 * File: CountLetterFrequencies.js
 * -------------------------------
 * This function displays a letter-frequency table for a file.
 */

/* Lets the user choose a file and displays the frequency table */

function CountLetterFrequencies() {
   let callback = function(text) {
      displayFrequencyTable(createFrequencyTable(text));
   };
   JSFileChooser.chooseTextFile(callback)
}

/* Creates and returns an array with n copies of value */

function createArray(n, value) {
   let array = [];
   for (let i = 0; i < n; i++) array.push(value);
   return array;
}
```
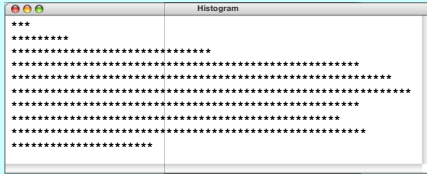

CountLetterFrequencies.js

```javascript
/* Creates a letter frequency table from the specified string */
function createFrequencyTable(str) {
   let letterCounts = createArray(26, 0);
   for (let i = 0; i < str.length; i++) {
      let ch = str.charAt(i).toUpperCase();
      if (ch >= "A" && ch <= "Z") {
         letterCounts[ch.charCodeAt(0) - "A".charCodeAt(0)]++;
      }
   }
   return letterCounts;
}

/* Displays the frequency table omitting letters whose count is 0 */

function displayFrequencyTable(letterCounts) {
   for (let i = 0; i < 26; i++) {
      let count = letterCounts[i];
      if (count > 0) {
         let ch = String.fromCharCode("A".charCodeAt(0) + i);
         console.log(ch + ": " + count);
      }
   }
}
```

skip code
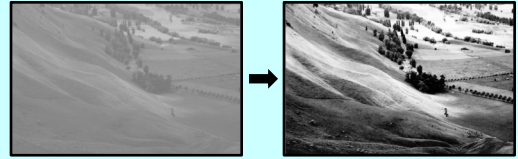
1

Exercise: Display a Histogram

- Write a program that reads a file called `MidtermGrades.txt` containing one score (0-100) per line and displays a histogram of the scores divided into ranges 0-9, 10-19, and so on.

```
000                          Histogram
***
*********
**********************************
***************************************************
*****************************************************
*******************************************************
*****************************************************
*****************************************************
*******************************************************
***********************
```

- Histograms are usually presented vertically. The one in this exercise is drawn horizontally because that program is so much easier to write.
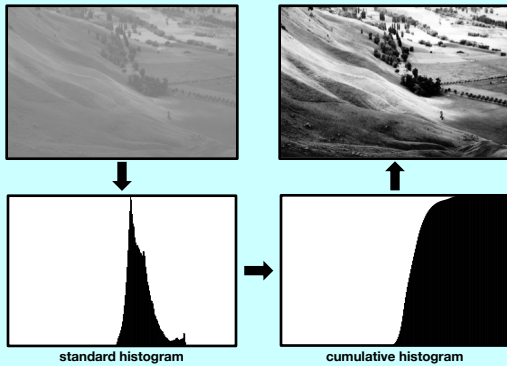
Image Histograms

- The statistical notions behind the histogram program are important in image processing.

- A few years ago, Keith Schwarz developed an assignment in which one of the exercises was to use histograms to equalize a photograph with insufficient contrast, as illustrated below:

low-contrast image   the same image after equalization

- The steps in the process are are illustrated on the next slide.

Equalization

standard histogram   cumulative histogram

The End