| University of Illinois at Urbana-Champaign <br> Dept. of Electrical and Computer Engineering |
| :---: |
|  |
| Programming |
| Review: |
| Letter Frequency Decomposition |
| Eccr 20: Compter System\& \& Programming |

## Let's Decompose the Problem

The task:
${ }^{\circ}$ given an ASCII string (terminated by NUL)

- count the occurrences of each letter
(regardless of case), and
- the number of non-alphabetic characters.

The high-level approach:
initialize histogram to all 0 s
for each character in the string
increment the appropriate histogram bin

ECE 220: Computer Systems \& Programming © 2016 Steven S. Lumetta. All rights reserved.

The First Step: Break the Task into a Sequence of Two
The Second Step: Count by Iterating Over Characters
The first decomposition step is clear from the high-level approach.


Next, break down counting into an iteration.


## Break Down Counting a Character into Two Steps

Counting one character involves two steps.
First, we must increment one bin in the histogram.
Then we must advance our pointer to the next character in the string.

How to Choose a Bin: Use a Conditional Construct

How can we determine which
histogram bin to increment?
The answer depends on the character.
We need to use conditional constructs.

## But how?

Let's take a look at the ASCII sequence.

ECE 220: Computer Systems \& Programming $\quad 2016$ Steven S. Lumetta. All rights reserved.

Start by Breaking Off the Left Region


Notice that

- if a character is less than 'A',
othe character is not a letter.
Let's start with the leftmost region.

ECE 220: Computer Systems \& Programming © 2016 Steven S. Lumetta. All rights reserved.


Continue by Breaking Off Capital Letters


We know that the character is not below 'A.'
What's left?
Let's handle capital letters next.

ECE 220: Computer Systems\& Programming © 2016 Steven S. Lumetta. All rights reserved.
slide 10


Continue with Characters in the Middle Region


We know that the character is at least '[.'
What's left?
Let's handle the middle region next.

ECE 220: Computer Systems \& Programming © 2016 Steven S. Lumetta. All rights reserved.


Continue with Characters in the Middle Region
©EE 220: Computer Systems \& Programming 2016 Steven S. Lumetta. All rights reserved.

## Initialization is a Sequence

What about initialization?

We need to do three things:

- fill the histogram with 0s,
- load any useful values (such as ASCII characters to check the region boundaries).
${ }^{\circ}$ and point to the start of the string.

ECE 220: Computer Systems \& Programming © 2016 Steven S. Lumetta. All rights reserved.

Filling the Histogram: a Sequence and an Iteration
How do we fill the histogram?
We have 27 bins ( 26 letters +1 non-alpha). We should use an iteration.
But again, we need a pointer to the histogram.
So:
${ }^{\circ}$ point a register to the histogram, - then iterate over all bins.

The Final Flow Chart


