

## CIS 110: Introduction to Computer Programming

Lecture 15  
Process them lines  
(§ 6.3-6.4)

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

1

## Outline

- Line-based text processing

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

2

## Line-based text processing

(Live demo, ReceiptMethods.java)

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

3

## Token-based processing review

- We walk over a chunk of text using the nextX methods of the Scanner class.

```
Scanner file = new Scanner(new File ("num.txt"));
while(file.hasNextDouble()) {
    System.out.print(file.nextDouble() + " ");
}
```

0.41	410.1	10.1					
	13	31.5					
0.8							
		0.5					

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

4

## Files with lines of text

- Many files use each line to represent an *individual record* of a data set.

```
Lastname Firstname weight year ppg
Allen Lavoy 225 R 0.0
Brackins Craig 230 R 2.7
Brand Elton 254 11 15.0
```

- Token-based processing throws away line information which makes it unsuitable for these kinds of situations

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

5

## Line-based processing template

```
// while file has lines left to process
// process the line using token-based
processing
```

- We can decompose processing a file into
  - Reading in each line of a file
  - Processing an individual line

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

6

## Tokenizing individual lines

- Once we have a line, how do we tokenize it?

```
Scanner line = new Scanner("Hello world!");
```

- Scanners can be made over Strings!
  - e.g., the line Scanner above tokenizes "Hello world!" into two tokens

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

7

## Line-based processing template expanded

```
Scanner file = new Scanner(new File("file.txt"));
// While file has lines left to process
while (file.hasNextLine()) {
    // Process the line using token-based processing
    Scanner line = new Scanner("Hello world!");
    // token-based processing here...
}
```

- Idiom: one scanner for file, one for each line
- Follow this pattern for the homework!

11/5/2011

CIS 110 (11fa) - University of Pennsylvania

8