CIS 110: Introduction to Computer Programming

Lecture 15

Process them lines

(§ 6.3-6.4)

Outline

- Line-based text processing

Line-based text processing

(Live demo, ReceiptMethods.java)

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
```

```
0.41 410.1 10.1 13. 31.5 0.8 0.5
```

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

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
  1. Reading in each line of a file
  2. Processing an individual line

Tokenizing individual lines

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

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

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

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!