Reading Log Files

Finding Unique IP Addresses

Introduction 1 min

Developing an Algorithm 3 min

Translating to Code 3 min

Equality 4 min

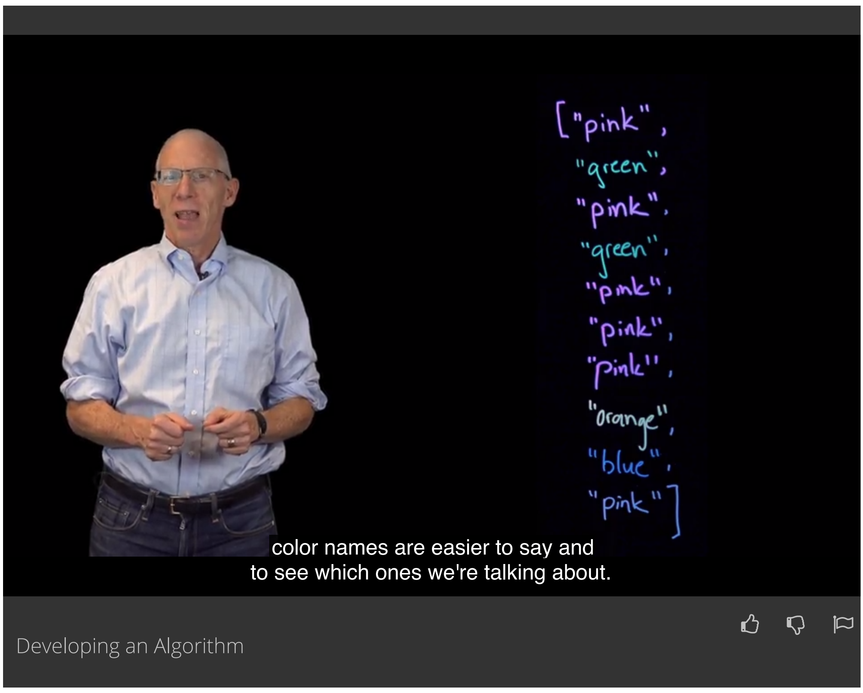
Summary 43 sec

Programming Exercise: Finding Unique IP Addresses 10 min

Practice Quiz: Finding Unique IP Addresses 4 questions

Counting Website Visits

Review



Developing an Algorithm

Have a question? Discuss this lecture in the week forums.

Interactive Transcript

Search Transcript English

0:04
You've already written code to read in the entire contents of a web server log. And now you wanna find out how many unique IP addresses it contains. As always, you're going to want to approach this problem with the seven steps you use for every problem. We're going to walk through this problem, starting at step one. And we're gonna use color names instead of IP addresses. This problem is really the same problem. You're finding out how many unique values there are in an array list of strings, but color names are easier to say and to see which ones we're talking about.

0:38
When we look at this list of ten values. You might be able to tell just by looking at it that there are only four unique values. But if this list had a million elements in it, we'd like to develop a method that would still work.

0:52
There are many ways of solving problems in general. And in this case, there are a couple that I'll mention and one that we'll realize in code. One method that we won't do is to as we look down the list, cross out values that we've seen before. Crossing out values in a list is a problem in Java programming because if we replace the values in our parameter with other values we've caused a side effect, something we've spoken about before as a thing to avoid if you can. So instead we'll use an idea of visiting every value in the list in turn, which is typical for array problems, and if we've haven't seen a value before we'll copy it into a new list. So first I'm copying the value pink, because I haven't seen it before, in my new list. Then I'm gonna visit green. I haven't seen it before. So I'll copy it over into my new list. Now I'm gonna visit pink. I've already seen it, so I don't copy it. Then green. I have seen it, so I don't copy it. Pink, pink, pink, I've seen them all. Orange I haven't seen before, so I'll copy that into my new list.

2:02
The next value's blue. I haven't seen.

2:06
So I'm copying it into my list.

2:08
Finally, I get to pink, the last value in my original list. I have seen it before, so I don't copy it over. Now, I look at my list that I've created. It has four values, and so I'm gonna return the value, four, for my method that I'm writing that determines the number of unique values in the array parameter.

2:29
Developing this algorithm follows a lot of the same patterns you've seen before. You'll see that sometimes you add an element to the copy, and sometimes you do not, and you'll want to think through the conditions under which you add. Once you've done that, you can express the main portion of the algorithm in terms of steps to do for each element of the input. We'll let you work through steps two and three.

2:54
Hopefully, you came up with pseudocode that looks like this.

2:59
Now for this particular problem, you'll want to do things just a little bit differently. Remember that you have log entries in the field records. And you want to use the get IP address method to get the string out of the log entry object. There are a couple ways to deal with this difference. The simplest is to just use the same algorithm but adjust it slightly to reflect the fact that you want to use records, the fields in the class and you want to get the IP address out of each element of records, check to see if that IP address is in copy, and if not, add the IP address to copy.

3:36
At this point, you'd wanna test out your pseudo code, then you're ready to turn it into Java code.

Downloads

Lecture Video mp4

Subtitles (English) WebVTT

Transcript (English) txt

Would you like to help us translate the transcript and subtitles into additional languages?