Regex

Andrew Rosen

April 1, 2019

Complete each of the following sections.

1 Regex Dictionary

For this section, I want you to answer questions about the dictionary file `words.txt`. If you're using a Unix system, feel free to use the dictionary file included on your computer. Some words begin with capital letters! Please include each regex with your answer.

1. I like animals. How many words contain the word 'cat' or 'dog' in them?
2. Four letters words are supposedly naughty. How many four letter words are there?
3. I am scared of pyramid schemes and mlms. How many words contain 'hun' in them?
4. Do more words end in "ing" or "ion?"
5. How many words contain a "q" not immediately followed by a "u."
6. How many words have no vowels?
7. How many words consist of only vowels?
8. How many words are contracted with "not," meaning they end with "'nt"
9. How many words with two vowels **in a row** are there?
10. How many words with at least two vowels are there? The vowels need not be adjacent, like in the previous problem.

If you are having trouble, please see this link.

2 More Regex

1. What is the difference between `.*` and `.*?` ?

2. How would you write a regex that matches the full name of someone whose last name is Nakamoto? You can assume that the first name that comes before it will always be one word that begins with a capital letter. The regex must match the following:

   - 'Satoshi Nakamoto'
   - 'Alice Nakamoto'
   - 'RoboCop Nakamoto'

   but not the following:

   - 'satoshi Nakamoto' (where the first name is not capitalized)
   - 'Mr. Nakamoto' (where the preceding word has a nonletter character)
   - 'Nakamoto' (which has no first name)
   - 'Satoshi nakamoto' (where Nakamoto is not capitalized)

3. Create a regex which matches the strings twenty, twenty-one, twenty-two,. . . ,ninety-nine

4. Create a regex that will match dollar values, such as the following.

   - $100.00
   - $10,000.00
   - $1234
   - $5000.00
   - $1,000,000

3 Strong Password Detection

Write a function that uses regular expressions to make sure the password string it is passed is strong. A strong password is defined as one that is at least eight characters long, contains both uppercase and lowercase characters, and has at least one digit. You may need to test the string against multiple regex patterns to validate its strength.

4 Kinda a Regex Problem

The above kind of password is actually bad design. Read the comic below.

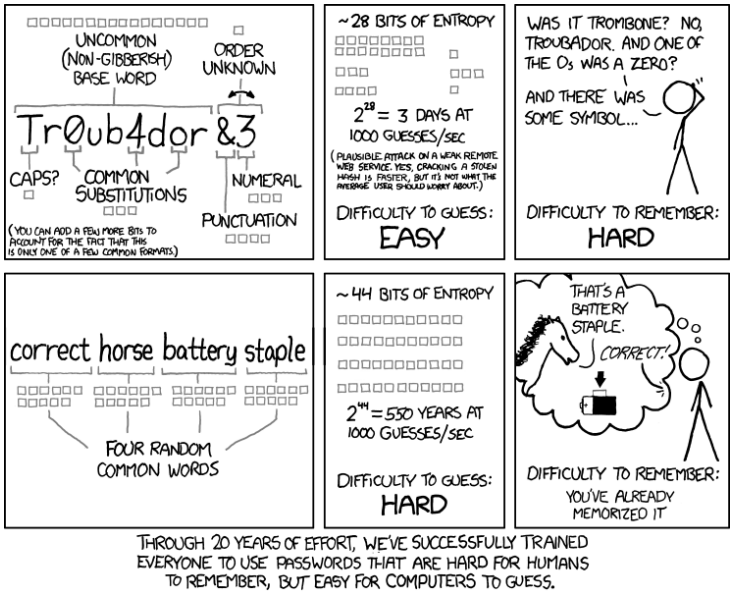

Figure 1: XKCD using stick figures to explain the concept better than your professor.

Write a program that creates passwords similar to the comic by using the `words.txt` file. Combine four random words that consist of only lowercase characters and are each at least four letters long into a single password.