COMP41100

EXPLORING PROGRAMMING IN RUBY

PRACTICAL SEVEN

07th November 2013

Andrew Doyle
Student Number: 12252388

University College Dublin

Q1

Task:

In the SixDegrees program, the find_films_by_actor method is fairly crap. If you look at any web---page it clearly misses loads of the films that the actor is in. Your task is to modify it to find more films (n.b., everything it finds should be a film---link, dealing with garbage would cause too many problems in loading).

Solution:

This method was improved by searching through the table identified as the next element after the **Filmography** heading. A hashmap is used to store the contents of the table.

```ruby
def find_films_by_actor_new(doca, out = [])
  doc = Nokogiri::HTML(open(doca))

  table = doc.at('#Filmography').parent.next_element
  films = table.search('tr')[1..-1].map{ |tr|
    tds = tr.search('td')
    year = tds.shift.text

    movie = tds.shift
    movie_url = movie.at('a')['href']
    movie_title = movie.at('a').text

    role = tds.shift.text

    {
      year: year,
      movie_url: movie_url,
      movie_title: movie_title,
      role: role
    }
  }

  puts films.map{|h|h[:movie_url]}
end
```

Q2

The way sixdegrees.rb searches through the space of actors and films is ok, but probably not the most efficient. Can you propose two modifications to the program that might search in a different way. Just propose the algorithm, you don't have to implement it.

It would be better to search the films first, and then the actors; rather than the other way around.