web

Jon Sporring

November 21, 2019

1 Lærervejledningn

Emne Input/output from the web

Sværhedsgrad medium

2 Introduktion

The internet is a great source of information, and many files are published as html-files. In the following assignment(s), you are to work with html-files on the internet.

Note that most internet pages requires a valid certificate before they will allow your program to access it. By default, Mono has no certificates installed. One way to install useful certificates is to use `mozroots`, which is part of the Mono package. On Linux/MacOS you do the following from the console:

```
mozroots --import --sync
```

On Windows you type the following (on one line)

```
mono "C:\Program Files (x86)\Mono\lib\mono\4.5\mozroots.exe" --import
--sync
```

Note that your installation of `mozroots` may be in a different path, and you may have to adapt the above path to your installation. After running the above, your program should be able to read most pages without being rejected.

3 Opgave(r)

1. Make a program with the function,

```
printWebPage : url:string -> string
```

which reads the content of the internetpage `url` and returns its content as a string option.

2. In the html-standard links are given by the `<a></a>` tags. For example, a link to Google's homepage is written as `<a href="https://google.com">Press to go to Google</a>`.

   Make a program

```
countLinks : url:string -> int
```

which reads the page given in `url` and counts how many links that page has to other pages. You should count by counting the number of `<a` substrings.