

STUDENT ID:300338937

TOTAL MARKS: 77.5

WHOLE DAYS LATE:

COMMENT: please follow submission guidelines next time and package your code as required

CORE:

- 10 Tested against a target html file (10 marks)
  - downloads html file to current directory.
  - hostname not retained in the name of the downloaded resource.
- 10 Tested against a target jpg file (10 marks)
  - downloads jpg file to current directory.
  - hostname not retained in the name of the downloaded resource.
- 5 Tested by running twice and checking that filename collisions are handled correctly (5 marks).
- 20 Inspected code to ensure that have implemented a general solution (20 marks).
  - extraction of filename from URL
  - collision handling
  - error handling (generally)
  - error handling (network errors)
- 0 Included README file assessing what did or didn't work (5 marks)

subtotal 45/50

COMPLETION:

- 10 Tested against a target html file to ensure that only resources on the same host as the root html file are downloaded (10 marks)
    - index.html
    - felids.html
    - felines.html
    - GrumpyCat.jpg
    - ignores commons.wikimedia.org and creativecommons.org
  - 5 Tested against a target html file to check that links are rewritten as required (10 marks).
    - absolute link (felines.html)
    - X relative link (images/GrumpyCat.jpg)
    - X relative link (felids.html) NOT rewritten
    - ignores commons.wikimedia.org and creativecommons.org
  - 2.5 Inspected code to ensure that have implemented a general solution (10 marks).
    - identification of which resources to download
    - X rewriting of absolute links
    - X rewriting of relative links
    - X ignoring resources located on a different server from the root html file
- COMMENT: your code which rewrites links was unsuccessful/not implemented

subtotal 17.5/30

CHALLENGE (20 marks):

- 10 Tested against a target html file to check that downloads the target web to a depth of two hops (10 marks).
  - creates appropriate subdirectory structure
  - downloads first level of the target website
  - downloads second level of the target website
  - ignores commons.wikimedia.org and creativecommons.org
  - does not treat the same file being downloaded twice as a

collision when the cause is a loop

- 5 Inspected code to ensure that have implemented a general solution (10 marks).  
- recursive downloading of resources

X file and directory collisions

COMMENTS: Directory collisions are not correctly handled

subtotal 15/20