Reading from Internet

CS 1111
Introduction to Programming

Spring 2019

File Processing (Local)

Repeated process

| Open a file | → | Read contents of a file | → | Process the content |

Write/append content to another file

Close a file
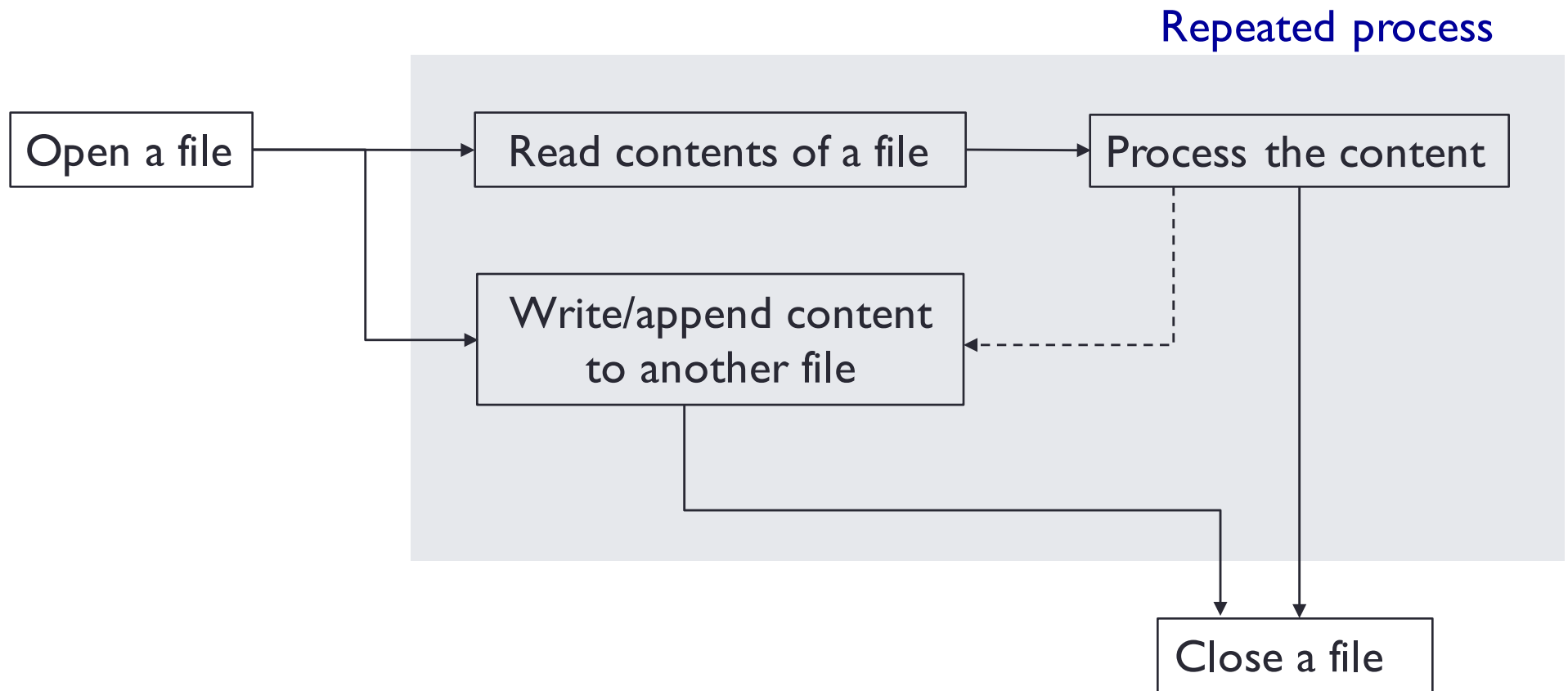
File Processing (Internet)

Repeated process

Open a file → Read contents of a file → Decode the content → Process the content
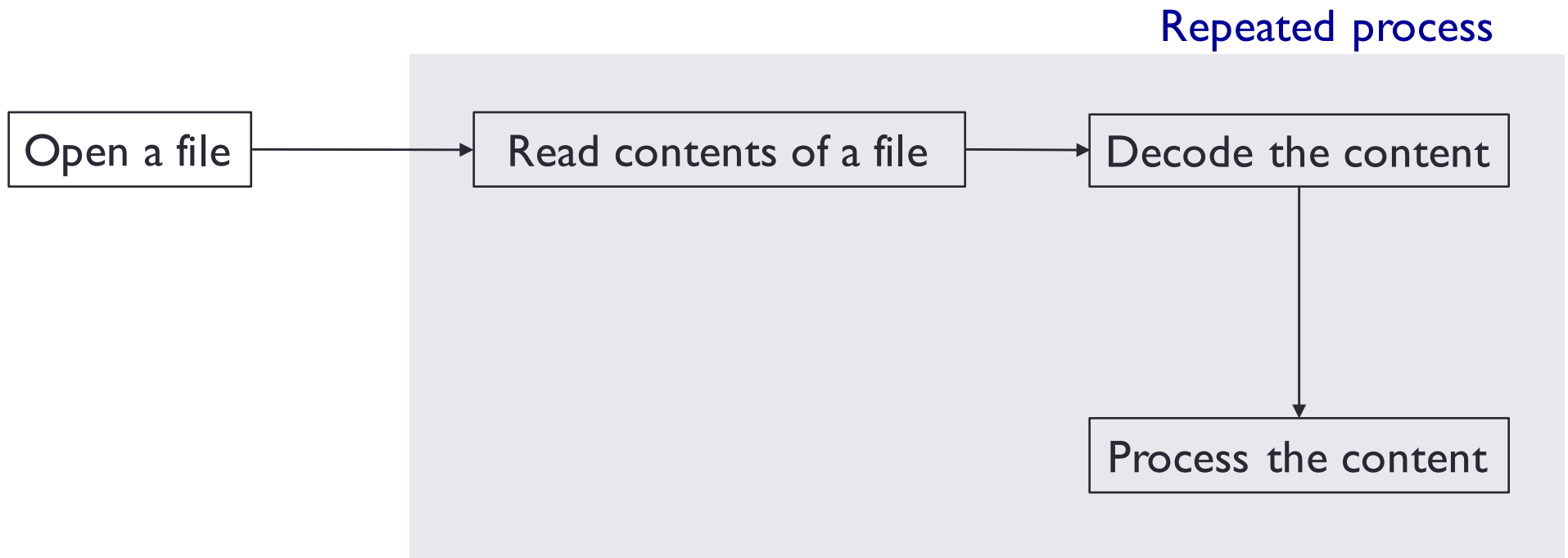
Opening Files (Internet, via URLs)

```
import urllib.request

url = some-url-to-the-file
stream = urllib.request.urlopen( url)

for line in stream:
    decoded = line.decode("UTF-8")
    print(decoded.strip())
```

Summary

- Read from files -- Must know (based on exam2 topic list, as of 03/27/2019)
    - `open(filename)`
    - `connection.read()`
    - `connection.readline()`
    - ways of iterating lines
        - `connection.readlines()`
        - `connection.read().split('\n')`
        - `for line in connection`

- Read from the Internet -- Must know (based on exam3 topic list, as of 03/27/2019)
    - `import urllib.request`
    - `urllib.request.urlopen(url)`
    - `stream.read().decode('utf-8')`