

ex5filehandling

Joachim von Hacht

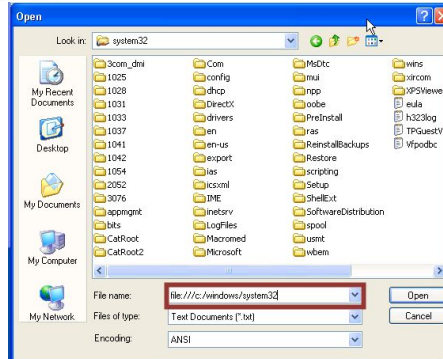
Filhantering

```
hajo/samples$ pwd
/home/hajo/courses/ooit/ass/ass.src/week5/src/s
amples
hajo/samples$ tree
```

```
├── Exceptions.java
├── filehandling
│   ├── BinaryFiles.java
│   ├── ObjectIO.java
│   ├── TextFiles.java
│   └── UseFile.java
├── OrderNumberStatic.java
├── ShortForLoop.java
├── StaticVariablesMethods.java
├── StringMethods.java
├── Strings.java
├── unused
│   ├── CommandLineArgs.class
│   ├── CommandLineArgs.java
│   ├── HigherOrderFunctions.java
│   ├── RemoveFromCollection.java
│   ├── StringBuilderThis.java
│   ├── UseArrayList.java
│   ├── UseAStack.java
│   └── UseCharacter.java
```

```
2 directories, 18 files
hajo/samples$
```

Kommandot tree i Linux



Windows filhanterare

2

En [fil](#) är en sammanhållen informationsmängd.

- Lagras på någon typ av sekundärminne.
- En fil är beständig, d.v.s. den existerar efter det att programmet som skapade filen är avslutat.
- Filer har storlekar (kB, MB, GB), en tidsstämpel, en ägare, rättigheter, m.m.
- Vissa filer kan fungera som behållare för andra filer, kallas bibliotek (eller kataloger eller mappar, directory)
- Det finns många [filformat](#) (alltså hur datan i filen är organiserad)

Filhantering innebär typiskt att: Skapa ny fil, öppna, läsa, skriva, lägga till, ändra, stänga, ta bort, kopiera, flytta fil(er).

- Sköts ofta av användare (människor) men behöver också kunna skötas av program.

Filsystem och Sökvägar

Absoluta

```
// Windows C: is root
C:\Documents\sally\statusReport
C:\Documents\sally\statusreport // Case insensitive, same dir

// Unix, Linux, Mac. Leading / is root
/home/sally/statusReport
/home/sally/statusreport // Case sensitive! This is another dir
```

Relativa

```
// Windows "." = curent dir, ".." = dir one step up
.\sally\statusReport
..\sally\statusReport

// Unix, Linux, Mac
./sally/statusReport
../sally/statusReport
../../sally/statusReport
```

3

Filerna i ett system är normalt ordnade i en (upp och nedvänd) trädstruktur, ett [filsystem](#)

- En sökväg ([path](#)) betecknar en unik position i filsystemet (trädstrukturen)
- Sökvägen kan vara absolut d.v.s. ange hela vägen från trädets rot till positionen ...
- ... eller relativ d.v.s. bara ange vägen utifrån en given plats i trädet.
 - Om vi använder absoluta sökvägar i programmet måste programmet alltid (på alla datorer) ligga på samma ställe i filsystemet.
- Sökvägar byggs upp av en rotbeteckning, biblioteksnamn samt ett avgränsningstecken "/" (för Unix/Linux/Mac) alternativt "\" för Windows.
- Alla filsystem utom Windows är case sensitive.
- Om ett program skall fungera på flera olika operativsystem måste vi se till att sökvägarna/filnamnen fungerar för alla
 - Finns sätt att göra detta i Java t.ex. m.h.a. System API:et.

Sökvägar i IntelliJ utgår från projektkatalogen

- Innebär att om filen finns i någon mapp i src får vi inleda med sökvägen med "src/.../...".

Programmatisk Filhantering



4

Alla programmeringsspråk har stöd för hantering av filer. Det finns API:er för filhantering.

- Ofta påminner API:er i olika språk om varann
- Detaljerna skiljer.
- Inget att lära sig utantill. Man söker dokumentation upp vid behov.
- Vi tittar lite översiktligt på några möjligheter i Java

Allt i bilden är inte språk, vissa saker är ramverk, plattformar.

Klassen File

```
File f = new File("src/samples/tmp");
out.println(f.mkdir());    // Create directory
out.println(f.getName());
out.println(f.exists());
out.println(f.isDirectory());
out.println(f.getAbsolutePath());
try {
    File junk = new File("src/samples/tmp/junk.txt");
    out.println(junk.createNewFile());
} catch (IOException e) {
    e.printStackTrace();
}
File[] files = f.listFiles(); //Content of directory
out.println(files.length);
out.println(files[0]);
```

5

Filhantering i ett Java-program kan skötas m. h. a. den färdiga klassen File.

- Konstruktorn och många av metoderna kastar kontrollerade undantag, måste hanteras.

Textfiler

// Unix/Linux

Det är skimmer i molnen och glitter i sjön,**LF**
det är ljus över stränder och näs,**LF**
och omkring står den härliga skogen grön**LF**
bakom ängarnas gungande gräs.**LF**

// Classic Mac (now LF)

Det är skimmer i molnen och glitter i sjön,**CR**
det är ljus över stränder och näs,**CR**
och omkring står den härliga skogen grön**CR**
bakom ängarnas gungande gräs.**CR**

// Windows

Det är skimmer i molnen och glitter i sjön,**CRLF**
det är ljus över stränder och näs, **CRLF**
och omkring står den härliga skogen grön **CRLF**
bakom ängarnas gungande gräs.**CRLF**

6

Innehållet i en textfil tolkas som tecken och är strukturerat i rader (en Java-fil är en textfil)

- Radsluten beror på operativsystem (vilka tecken (koder) som används för radslut)
 - Rött i bilden.

När filen sparas/läses sker en omvandling från/till bytes till/från tecken (t.ex. Unicode)

- 199 sparas som '1' '9' '9' vilket kodas som 0x31, 0x39, 0x39 (hexadecimalt)

Binärfiler

```
feca beba 0000 3400 5000 000a 0015 072a 2b00 0009 002c
0a2d 0200 2e00 0009 0006 072f 3000 000a 0006
0a2a 0600 3100 0009
002c 0832 3300 000a 0034 0a35 0200 3600 0008 0737 3800
000a 000e 082a 3900 000a 000e 0a3a 0e00 3b00 000a 000e
0a3c 3400
3d00 0007 013e 0200 6373 0001 4c13 616a
6176 752f 6974 2f6c 6353 6e61 656e 3b72
0001 3c06 6e69 7469 013e 0300 2928 0156 0400
6f43 6564 0001 4c0f 6e69 4e65 6d75
6562 5472 6261 656c 0001 4c12 636f 6c61 6156 6972 6261
656c 6154 6c62 0165 0400 6874 7369 0001 4c13 7865 7265
6963 6573 2f73 7845 5f31 4d42 3b49 0001 6d04 6961 016e
1600 5b28 6a4c 7661 2f61 616c 676e 532f 7274 6e69 3b67
5629 0001 6104 6772
0173 1300 4c5b 616a 6176 6c2f 6e61 2f67
...
...
```

Binärfil utskriven på hexadecimal
form (egentligen finns inga rader,
beror från utskriften här)

Binärfiler är ren digital data

- Ej ordnade i rader, ej läsbara för människor
- Ingen omvandling av data sker, ren binär data skrivs och läses

Läsa en Textfil

```
// Method to read a file. NOTE throws
List<String> readFile(Path path) throws IOException {
    List<String> lines = new ArrayList<>();
    // Use "try with resources" style to ensure stream is closed
    try (BufferedReader reader = Files.newBufferedReader(path)) {
        String line;
        while ((line = reader.readLine()) != null) {
            lines.add(line);
        }
        return lines;
    }
}

// Call must use try-catch, method throws checked exception
try {
    Path path = Paths.get("src/samples/filehandling/junk.txt");
    List<String> copy = readFile(path);
} catch (IOException e) {
    e.printStackTrace();
}
```

8

Finns flera sätt (nya och gamla) att läsa och skriva textfiler i Java

- Förvirrande för nybörjare (och andra)
- Dessutom ganska tekniskt komplicerat
- Detta är inget man lär sig utantill
- Sker på olika men liknande sätt i de flesta programmeringsspråk.

Analys av kod:

- Vi ska läsa en textfil
- För detta har vi skapat en metod `readFile`. I metoden kan det kastas checked exceptions (beror på de färdiga Java-metoder vi anropar).
 - Någon måste hantera dessa.
 - Vi väljer att inte göra det i metoden utan skickar ansvaret till den som anropar metoden genom att ange throws i metodhuvudet. Detta gäller bara checked exception!
- Innehållet i filen spar vi i en `List<String>`, varje rad lagras som en sträng i listan
- Vi använder dessutom en speciell form av try, kallas try-with-resources.
 - Problemet är att vi måste se till att inströmmen (som används i bakgrunden) stängs ifall ett undantag uppstår, annars kan programmet "äta" upp en massa resurser (minne).
 - Om vi använder try-with-resources garanterar Java att

- strömmen stängs.
- I try satsen kör vi en loop som läser rad för rad
 - Om det inte finns någon mer rad kommer line-variabeln att innehålla null
 - .. om så bryter vi loopen
 - Annars lägger vi till raden i List.
- Vid anropet till readFile ger vi ett Path-objekt som argument till metoden

Skriva en Textfil

```
void writeFile(Path path, List<String> lines) throws IOException {  
    try (BufferedWriter writer = Files.newBufferedWriter(path)) {  
        for (String line : lines) {  
            writer.write(line);  
            writer.newLine();  
        }  
    }  
}
```

```
List<String> text = Arrays.asList("a", "b", "c", "d", "e");  
try {  
    Path path = Paths.get("src/samples/filehandling/junk.txt");  
    writeFile(path, text);  
} catch (IOException e) {  
    e.printStackTrace();  
}
```

Påminner mycket om att läsa en fil.

Läsa ett Objekt (binärfil)

```
class MyClass implements Serializable {
    private String name;
    public MyClass(String name) { this.name = name;}
    public String getName() { return name;}
    public void setName(String name) {this.name = name;}
}

MyClass readObject(Path path) throws IOException, ClassNotFoundException {
    try (ObjectInputStream is = new ObjectInputStream(Files.newInputStream(path))) {
        return (MyClass) is.readObject(); // Read full object from file
    }
}

// Call
try {
    Path path = Paths.get("src/samples/filehandling/pelle.ser" );
    MyClass o = readObject(path);
} catch (IOException | ClassNotFoundException e) {
    e.printStackTrace();
}
```

10

Man kan skriva och läsa objekt från/till (binär)filer (och andra format t.ex. XML, JSON).

- Klassen måste ange implements Serializable (d.v.s. det går att omvandla objekten till byteströmmar)
- För att kunna läsa objekt måste det finnas en klass (att omvandla bytes:en till)
 - Annars ClassNotFoundException
- Det vi läser måste explicit typomvandlas till klassen för objektet vi vill omvandla till.

Skriva ett Objekt (binärfil)

```
void writeObject(Path path, MyClass object) throws IOException {  
    try (ObjectOutputStream os =  
        new ObjectOutputStream(Files.newOutputStream(path))) {  
        os.writeObject(object); // Write full object to file  
    }  
}  
  
// Create object  
MyClass o = new MyClass("pelle");  
  
// Call  
try {  
    Path path = Paths.get("src/samples/filehandling/pelle.ser" );  
    writeObject(path, o);  
} catch (IOException | ClassNotFoundException e) {  
    e.printStackTrace();  
}
```

Ungefär som att läsa