

File at this location, same text as below.

<https://ccd-school.de/en/coding-doj/application-katas/bankocr/>

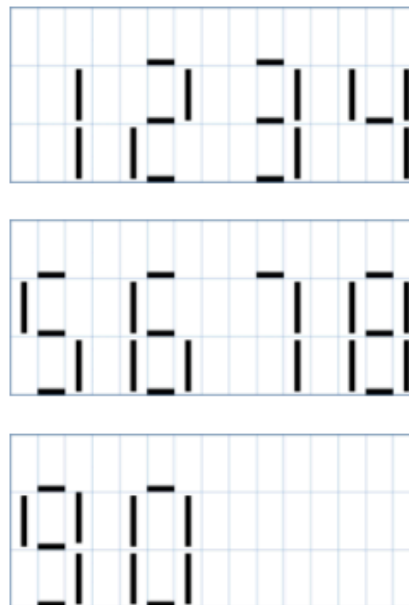
## BankOCR

### Application Kata „BankOCR“

Write a program that scans account numbers from ASCII files.

OCR means *optical character recognition*. Of course it would be hard to implement a real ORC algorithm as an exercise. But lets see if we can reduce the complexity of such an algorithm.

The ASCII files contain numbers that each are encoded on three lines. The following picture shows how the digits 1234567890 are encoded in the file.



Each digit is three characters wide and three characters high. Consecutive digits are delimited by spaces. An empty line delimits consecutive rows with numbers. Each digit is build from „\_“ (underscore) and „I“ (uppercase letter I).

The program is started with a file name as a parameter. It prints the recognized numbers on the console. Please note that the input file is considered well formed. There are no errors in the input files.

```
C:> bankocr file1.txt
1234567890
815
42
07
```

Test data can be found [here](#).

### Variation #1

Files may contain errors. Although the structure of the rows is correct: three rows build one number, a blank line delimits multiple numbers. The structure of the digits is correct too, so each digit consists of a 3 x 3 matrix of characters. But the characters inside the 3 x 3 matrix may not be valid. There may be illegal characters and there may be characters at an illegal position. Each number that could not be recognized because of such errors should be printed as „Error in data“.

```
C:> bankocr file2.txt
1234567890
Error in data
42
07
```