7 Cybersecurity (fms27)

The following Python program (`imports` omitted for brevity) takes a user-supplied `fields.txt` input file containing two text fields, one per line. By concatenating them with template fragments (not fully shown for brevity), it writes out a temporary LaTeX file that places those fields in specific positions on the page. It then compiles the LaTeX into a pdf, intended to be overprinted onto an existing form. Assume the user has no direct access to the file system: the user interacts with a web page that writes the user-supplied fields, as supplied, to the `fields.txt` file. [*Note:* In Python, when the end of the file is reached, calling `readline` returns empty string. The default encoding used by `readline` is UTF-8.]

```
1  t_before = "\\documentclass{article} \\begin{document} ..."
2  t_middle = " ... "
3  t_after = " ... \\end{document}"
4  with open("fields.txt", "r") as fields:
5       f0 = fields.readline()
6       f1 = fields.readline()
7  with open("form-content.tex", "w") as latex:
8     latex.write(t_before + f0 + t_middle + f1 + t_after)
9  os.system("pdflatex form-content.tex")
```

(*a*) Is it possible to produce a `fields.txt` input file that will cause a malfunction while executing lines 1–8? If yes, produce such input; otherwise justify why this is not possible. State your assumptions explicitly. [3 marks]

(*b*) Is it possible to produce a `fields.txt` input file that will cause a malfunction while executing line 9? If yes, produce such input; otherwise justify why this is not possible. State your assumptions explicitly. [3 marks]

(*c*) Amend the program so that it won't malfunction on the specific inputs you provided in parts (*a*) or (*b*). [*Note:* Ad-hoc fixes that won't fix the general problem are deemed sufficient for full marks here—but see part (*e*).] [*Hint:* At least one of parts (*a*) or (*b*) admits a "yes" answer.] [3 marks]

(*d*) Is it possible for an attacker to exploit this program to execute an arbitrary command, and under which circumstances? If yes, produce an input file that will execute `/tmp/payload`. If not, justify why this is not possible. State your assumptions explicitly. [3 marks]

(*e*) After the above program caused embarrassment, you are called in to offer security training to the company's programmers. State the two most important points you are going to make. For each, clearly explain the remedial approaches you recommend, discussing them in the context of the above listing. [8 marks]