Homework 1

1. **(Optional)** If you want to run the code presented in the lectures, install the Scala programming language available (for free) from

   <center>http://www.scala-lang.org</center>

   and the Ammonite REPL from

   <center>https://ammonite.io</center>

   If you want to follow the code I present during the lectures, it might be useful to install VS Code or Codium. Please have a look at the handout about Ammonite and if you need a refresher for Scala - I linked on KEATS the Scala handout from PEP.

2. **(Optional)** Have a look at the catastrophic backtracking programs uploaded on KEATS. Convince yourself that they really require a lot of computation time. If you have similar examples in your own favourite programming language, I am happy to hear about it.

3. Read the handout of the first lecture and the handout about notation. Make sure you understand the concepts of strings and languages. In the context of the CFL-course, what is meant by the term *language*?

4. Give the definition for regular expressions—this is an inductive datatype. What is the meaning of a regular expression? (Hint: The meaning is defined recursively.)

5. Assume the concatenation operation of two strings is written as $s_1@s_2$. Define the operation of *concatenating* two sets of strings. This operation is also written as $\_@\_$. According to this definition, what is $A@\{\}$ equal to? Is in general $A@B$ equal to $B@A$?

6. Assume a set A contains 4 strings and a set B contains 7 strings. None of the strings is the empty string. How many strings are in $A@B$?

7. How is the power of a language defined? (Hint: There are two rules, one for $\_^0$ and one for $\_^{n+1}$.)

8. Let $A = \{[a], [b], [c], [d]\}$. (1) How many strings are in A^4? (2) Consider also the case of A^4 where one of the strings in A is the empty string, for example $A = \{[a], [b], [c], []\}$.

9. (1) How many basic regular expressions are there to match **only** the string *abcd*? (2) How many if they cannot include **1** and **0**? (3) How many if they are also not allowed to contain stars? (4) How many if they are also not allowed to contain $\_+\_$?

10. When are two regular expressions equivalent? Can you think of instances where two regular expressions match the same strings, but it is not so obvious that they do? For example $a + b$ and $b + a$ do not count…they obviously match the same strings, namely $[a]$ and $[b]$.

11. What is meant by the notions *evil regular expressions* and by *catastrophic backtracking*?

12. Given the regular expression $(a + b)^* \cdot b \cdot (a + b)^*$, which of the following regular expressions are equivalent

    1) $(ab + bb)^* \cdot (a + b)^*$
    2) $(a + b)^* \cdot (ba + bb + b) \cdot (a + b)^*$
    3) $(a + b)^* \cdot (a + b) \cdot (a + b)^*$

13. Given the extended regular expression `[b-d]a?e+`, what does the equivalent basic regular expression look like?

14. **(Optional)** This question is for you to provide regular feedback to me: for example what were the most interesting, least interesting, or confusing parts in this lecture? Any problems with my Scala code? Please feel free to share any other questions or concerns. Also, all my material is ~~crap~~ imperfect. If you have any suggestions for improvement, I am very grateful to hear.

If *you* want to share anything (code, videos, links), you are encouraged to do so. Just drop me an email or send a message to the Forum.