CMPU 240 · Spring 2026

Assignment 1

Submissions due: 3 February, 1:30 p.m.
Corrections due: 5 February, 1:30 p.m.

Exercise 1

In language theory, languages are represented as sets of strings, and we use set operations to manipulate them. This exercise is to refresh your memory of set theory from CMPU 145. Consider the following languages:

$$A = \{\mathtt{a}, \mathtt{b}\}$$
$$B = \{\mathtt{c}, \mathtt{d}\}$$
$$C = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}\}$$

a. Is $B \in C$? Is $B \subseteq C$?

b. What is $A \cap B$? How about $A \cup B$?

c. What is $C - A$?

d. What is $A \times B$?

Exercise 2

How would you represent each of the following sets as a formal language? (There are many possible right answers!) For each set,

- write what the alphabet (Σ) would be,
- describe informally how to encode an element as a string, and
- give an example of a string over Σ that belongs to the language and a string over Σ that does not belong to the language.

You don't need to explain how you would actually decide whether a string belongs to the language. As an example, for "the set of all episode IDs for a TV show", you might write:

> The alphabet $\Sigma = \{\texttt{0}, \ldots, \texttt{9}, \texttt{s}, \texttt{e}\}$.
>
> An episode ID would consist of the season number – `s` followed by one or more digits – and the episode number – `e` followed by one or more digits.
>
> `s01e23` would be in the language, while `es0123` would not.

a. The set of all valid games of tic-tac-toe.[1]

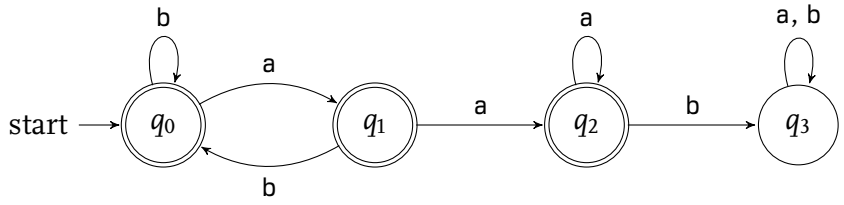
[1] You might know it as "noughts and crosses" instead.

b. The set of all syntactically correct Python programs.

Hint: Your answer for this part only needs to be three sentences. The problem does *not* ask you to describe the syntax of Python, only to say how to encode a syntactically correct Python program as a string. Your examples should only be one line long.

Exercise 3

Consider the following deterministic finite automaton (DFA), M:



a. What is the start state of M?

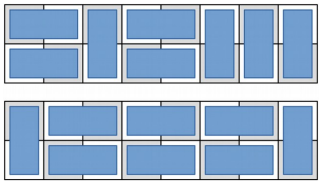
b. What is the set of accept states of M?

c. What sequence of states does M go through on input `bbaaa`? Does M accept this input? Why or why not?

d. Describe the set of strings accepted by M.

Exercise 4

There are many ways to tile a 2×8 checkerboard with dominoes, two of which are shown to the right.



Two domino tilings

Notice that the horizontal dominoes must appear as stacked pairs. We can encode each tiling from left to right as a string made from the characters `I` and `B`, where `I` denotes a vertical domino and `B` denotes two horizontal dominoes. The top tiling here would be represented as `BIBIII` and the bottom tiling as `IBBBI`.

Let $\Sigma = \{\texttt{B}, \texttt{I}\}$. Design the state diagram of a DFA for the language $\{w \in \Sigma^* \mid w \text{ represents a domino tiling of a } 2 \times 8 \text{ checkerboard}\}$.

Your solution for this exercise should be submitted as a CSV file generated by `tock`, using the provided Colab notebook.

Exercise 5

You're taking a walk with your dog, Dug. Dug is a vicious beast, so he's on a short leash, which keeps the distance between you at most two steps. You both start at the same position. (Dug's sitting on your feet.)

Consider the alphabet $\Sigma = \{\texttt{y}, \texttt{d}\}$. A string in Σ^* can be thought of as a sequence of events in which either you or Dug moves forward one step. For example, the string `ddy` means Dug takes two step forward, then you take one step forward.

Design the state diagram of a DFA for the language

$$L = \{w \in \Sigma^* \mid w \text{ describes a sequence of steps where you and Dug are never more than two steps apart}\}.$$

Dug

Your solution for this exercise should be submitted as a CSV file generated by `tock`, using the provided Colab notebook.

Exercise 6

When you upload the PDF of your assignment to Gradescope, you'll be asked to select the pages that have your answers to each exercise. This is *immensely* helpful when I'm grading your work. If you select the right page(s), then when I go to grade a particular exercise, I'll see your solution automatically. Conversely, if you *don't* select your pages, when I go to grade a particular exercise, Gradescope says something to the effect of "the student might not have answered this question at all". How sad!

So, please make sure to tag your pages when submitting. Gradescope has a guide to submitting PDFs if you need help. I'll give you credit for this exercise if you do so! (Aside from tagging your pages, you do not need to do anything to answer this exercise.)

Acknowledgments

This assignment includes exercises adapted from Keith Schwarz, Stanford University, and David Chiang, University of Notre Dame.