CS 516—Software Foundations via Formal Languages—Spring 2022

Problem Set 6

Due by 5pm on Friday, April 22
Submission via Gradescope and GitHub

Problem 1 (20 points)

Let $X = \{\, 0^i 1^j 2^k 3^l \mid i, j, k, l \in \mathbb{N} \text{ and } i + j \leq k + l \,\}$. Thus $0012223 = 0^2 1^1 2^3 3^1 \in X$, because $2 + 1 = 3 \leq 4 = 3 + 1$, and $11112233 = 0^0 1^4 2^2 3^2 \in X$, because $0 + 4 = 4 \leq 4 = 2 + 2$. On the other hand, any element of $\{0, 1, 2, 3\}^*$ with the digits out of order (e.g., 3201) is not in X, and, e.g., $00023 = 0^3 1^0 2^1 3^1$ is not in X, because $3 + 0 = 3 > 2 = 1 + 1$.

Prove that X is not regular.

Problem 2 (30 points)

Let X be as in Problem 1.

(a) Find a grammar G such that **alphabet** $G = \{0, 1, 2, 3\}$ and $L(G) = X$. [15 points]

(b) Express G in Forlan's syntax in the file `ps6-p2-gram` of the subdirectory `CS516-PS6` of your private GitHub repository. Use Forlan to find and display the alphabet of G. Use Forlan to find and display parse trees showing why the strings 0012223 and 11112233 are generated by G. Draw those parse trees (e.g., using JForlan). Include a transcript of your Forlan session. [5 points]

(c) Test your grammar on all elements of $\{0, 1, 2, 3\}^*$ of length no more than 9. Include a transcript of your Forlan session. Your testing code should reside in the file `ps6-p2-testing.sml` of the subdirectory `CS516-PS6` of your private GitHub repository [10 points]

Problem 3 (50 points)

Given an $m \in \mathbb{N}$, let X_m be the set of all $w \in \{0,1,2\}^*$ such that

- neither 02, 10 nor 21 are substrings of w; and
- for all substrings v of w, if $|v| = m$, then $\{0,1,2\} \subseteq \textbf{alphabet}\, v$.

For example, $0112011 \in X_4$, because the digits come in an allowed order and all of the substrings of length 4 (0112, 1120, 1201 and 2011) have occurrence of all three digits. On the other hand, $021 \notin X_4$, because it has the forbidden substring 02. And $011220 \notin X_4$, because the substring 1122 has length 4 but doesn't include 0.

(a) Use the functions/algorithms given in the book/lecture slides to define a function/algorithm **ansDFA** $\in \mathbb{N} \to$ **DFA** such that, for all $m \in \mathbb{N}$:

- **alphabet**(**ansDFA** m) $\subseteq \{0,1,2\}$;
- L(**ansDFA** m) $= X_m$; and
- **minimize**(**ansDFA** m) is isomorphic to **ansDFA** m.

[20 points]

(b) In the file `ps6-p3.sml`, define a Forlan/SML function

```
val ansDFA : int -> dfa
```

that implements your definition of **ansDFA**. You should assume that `ansDFA` will only be called with non-negative integers.

Run your definition of `ansDFA` on the elements of $\{0,1,2,3,4,5\}$, displaying the alphabets and numbers of states of the resulting six DFAs. Test these six DFAs on all elements of $\{0,1,2\}^*$ of length no more than 12. Include a transcript of your Forlan session.

You should put your `ps6-p3.sml` in the subdirectory `CS516-PS6` of your private GitHub repository. Your testing code should reside in the file `ps6-p3-testing.sml` in this directory.

[15 points]

(c) Prove that, for all $m \in \mathbb{N}$, L(**ansDFA** m) $= X_m$. [15 points]