

Problem Set 6

Model Answers

Problem 1

Suppose, toward a contradiction, that X is regular. Thus there is an $n \in \mathbb{N} - \{0\}$ with the property of the Pumping Lemma, where X has been substituted for L . Suppose $z = 0^n 1^n 2^n 3^n$. Since $n + n \leq n + n$, we have that $z \in X$. Thus, since $|z| = 4n \geq n$, it follows there are $u, v, w \in \mathbf{Str}$ such that $z = uvw$ and properties (1)–(3) of the lemma hold. Since $uvw = z = 0^n 1^n 2^n 3^n$, (1) tells us that there are $i, j, k \in \mathbb{N}$ such that

$$u = 0^i, \quad v = 0^j, \quad w = 0^k 1^n 2^n 3^n, \quad i + j + k = n.$$

By (2), we have that $j \geq 1$, and thus that $n + j > n$. By (3), we have that $0^{i+j+j+k} 1^n 2^n 3^n = 0^i 0^j 0^j 0^k 1^n 2^n 3^n = uvvw = wv^2w \in X$. Thus $(i + j + j + k) + n \leq n + n$, so that $n + j = i + j + k + j = i + j + j + k \leq n$. But $n + j \leq n$ contradicts $n + j > n$. Thus X is regular.

Problem 2

(a) G is

$$\begin{aligned} A &\rightarrow 0A3 \mid A3 \mid B \mid C \\ B &\rightarrow 1B3 \mid D \\ C &\rightarrow 0C2 \mid D \\ D &\rightarrow \% \mid 1D2 \mid D2 \end{aligned}$$

(b) We put the expression

```
{variables} A, B, C, D {start variable} A
{productions}
A -> 0A3 | A3 | B | C;
B -> 1B3 | D;
C -> 0C2 | D;
D -> % | 1D2 | D2
```

of G in the file `ps6-p2-gram`, and load G into Forlan, calling it `gram`:

```
- val gram = Gram.input "ps6-p2-gram";
val gram = - : gram
```

Then we compute and display the alphabet of `gram`, find parse trees `pt1` and `pt2` showing why `0012223` and `11112233` are generated by `gram`, and then display those parse trees:

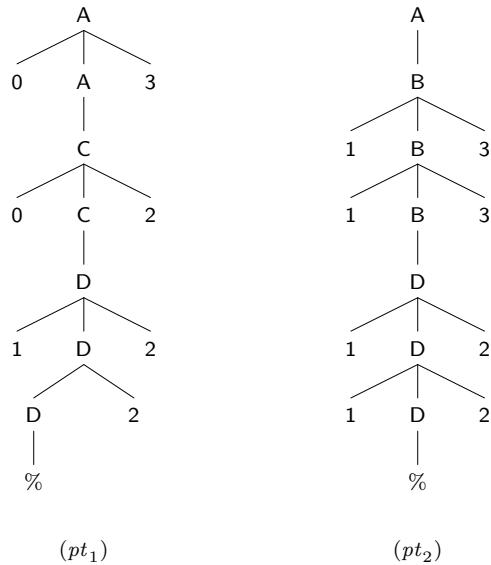
```
- SymSet.output("", Gram.alphabet gram);
```

```

0, 1, 2, 3
val it = () : unit
- val pt1 = Gram.parseAlphabet gram (Str.fromString "0012223");
val pt1 = - : pt
- val pt2 = Gram.parseAlphabet gram (Str.fromString "11112233");
val pt2 = - : pt
- PT.output("", pt1);
A(0, A(C(0, C(D(1, D(D(%), 2), 2)), 2)), 2), 3)
val it = () : unit
- PT.output("", pt2);
A(B(1, B(1, B(D(1, D(1, D(%), 2), 2)), 3), 3))
val it = () : unit

```

Here are drawings of the two parse trees:



(c) Continuing our Forlan session, we put our testing code

```

(* val inOrder : sym list -> bool

    inOrder x tests whether an element of {0, 1, 2, 3}^* is in
    {0}^*{1}^*{2}^*{3}^* *)

fun inOrder (b :: c :: ds) =
    Sym.compare(b, c) <> GREATER andalso inOrder(c :: ds)
  | inOrder _ = true;

(* val count : sym * sym list -> int

    count(a, bs) counts the number of occurrences of a in bs *)

```

```

fun count(_, nil)      = 0
  | count(a, b :: bs) =
    (if Sym.equal(a, b) then 1 else 0) + count(a, bs);

(* val inX : str -> bool

   inX x tests whether an element x of {0, 1, 2, 3}^* is in X *)

fun inX (x : str) =
  inOrder x andalso
  let val i = count(Sym.fromString "0", x)
      val j = count(Sym.fromString "1", x)
      val k = count(Sym.fromString "2", x)
      val l = count(Sym.fromString "3", x)
  in i + j <= k + l end;

(* val upto : int -> str set

   if n >= 0, then upto n returns all strings over alphabet {0, 1, 2,
   3} of length no more than n *)

fun upto 0 : str set = Set.sing nil
  | upto n
    =
    let val xs = upto(n - 1)
        val ys = Set.filter (fn x => length x = n - 1) xs
    in StrSet.union
        (xs, StrSet.concat(StrSet.fromString "0, 1, 2, 3", ys))
    end;

(* val partition : int -> str set * str set

   if n >= 0, then partition n returns (xs, ys) where:

   xs is all elements of upto n that are in X; and

   ys is all elements of upto n that are not in X *)

fun partition n = Set.partition inX (upto n);

(* val test : int -> gram -> str option * str option

   if n >= 0, then test n returns a function f such that, for all grammars
   gram, f gram returns a pair (xOpt, yOpt) such that:

   If there is an element of {0, 1, 2, 3}^* of length no more than n
   that is in X but is not generated by gram, then xOpt = SOME x
   for some such x; otherwise, xOpt = NONE.

```

If there is an element of $\{0, 1, 2, 3\}^*$ of length no more than n that is not in X but is generated by gram , then $y\text{Opt} = \text{SOME } y$ for some such y ; otherwise, $y\text{Opt} = \text{NONE}$. *)

```
fun test n =
  let val (goods, bads) = partition n
  in fn gram =>
    let val generated      = Gram.generated gram
        val goodNotGenOpt = Set.position (not o generated) goods
        val badGenOpt     = Set.position generated bads
    in ((case goodNotGenOpt of
          NONE     => NONE
          | SOME i => SOME(ListAux.sub(Set.toList goods, i))),
       (case badGenOpt of
          NONE     => NONE
          | SOME i => SOME(ListAux.sub(Set.toList bads, i))))
    end
  end;
end;
```

in the file `ps6-p2-testing.sml`, and load it into Forlan:

```
- use "ps6-p2-testing.sml";
[opening ps6-p2-testing.sml]
val inOrder = fn : sym list -> bool
val count   = fn : sym * sym list -> int
val inX     = fn : str -> bool
val upto    = fn : int -> str set
val partition = fn : int -> str set * str set
val test    = fn : int -> gram -> str option * str option
val it     = () : unit
```

Then we carry out the required testing:

```
- test 9 gram;
val it = (NONE,NONE) : str option * str option
```

Problem 3

(a) First, we give these definitions:

$$\begin{aligned} \text{minAndRen} &= \text{renameStatesCanonically} \circ \text{minimize}, \\ \text{faToDFAMar} &= \text{minAndRen} \circ \text{nfaToDFA} \circ \text{efaToNFA} \circ \text{faToEFA}, \\ \text{regToDFAMar} &= \text{faToDFAMar} \circ \text{regToFA}, \\ \text{allStrDFA} &= \text{regToDFAMar}((0 + 1 + 2)^*). \end{aligned}$$

Thus $\text{minAndRen} \in \text{DFA} \rightarrow \text{DFA}$, $\text{faToDFAMar} \in \text{FA} \rightarrow \text{DFA}$, $\text{regToDFAMar} \in \text{Reg} \rightarrow \text{DFA}$ and $\text{allStrDFA} \in \text{DFA}$.

We define $\text{lenDFA} \in \mathbb{N} \rightarrow \text{DFA}$ by: $\text{lenDFA } m = \text{regToDFAMar}((0 + 1 + 2)^m)$.

We define $\text{hasStrDFA} \in \text{Str} \rightarrow \text{DFA}$ and $\text{notHasStrDFA} \in \text{Str} \rightarrow \text{DFA}$ by:

$$\begin{aligned} \text{hasStrDFA } x &= \text{regToDFAMar}((0 + 1 + 2)^*(\text{strToReg } x)(0 + 1 + 2)^*), \\ \text{notHasStrDFA } x &= \text{minAndRen}(\text{minus}(\text{allStrDFA}, \text{hasStrDFA } x)). \end{aligned}$$

We let the DFA hasAllSymsDFA be

$$\text{minAndRen}(\text{inter}(\text{hasStrDFA } 0, \text{inter}(\text{hasStrDFA } 1, \text{hasStrDFA } 2))),$$

and let the DFA notHasAllSymsDFA be

$$\text{minAndRen}(\text{minus}(\text{allStrDFA}, \text{hasAllSymsDFA})).$$

We define $\text{lenAndNotHasAllSymsDFA} \in \mathbb{N} \rightarrow \text{DFA}$ by:

$$\begin{aligned} \text{lenAndNotHasAllSymsDFA } m &= \\ \text{minAndRen}(\text{inter}(\text{lenDFA } m, \text{notHasAllSymsDFA})). \end{aligned}$$

We then define $\text{someLenNotHasAllSymsFA} \in \mathbb{N} \rightarrow \text{FA}$ by:

$$\begin{aligned} \text{someLenNotHasAllSymsFA } m &= \\ \text{concat}(\text{allStrDFA}, \text{concat}(\text{lenAndNotHasAllSymsDFA } m, \text{allStrDFA})). \end{aligned}$$

We then define $\text{someLenNotHasAllSymsDFA} \in \mathbb{N} \rightarrow \text{DFA}$ by:

$$\text{someLenNotHasAllSymsDFA } m = \text{faToDFAMar}(\text{someLenNotHasAllSymsFA } m).$$

And we define $\text{allLenHasAllSymsDFA} \in \mathbb{N} \rightarrow \text{DFA}$ by:

$$\begin{aligned} \text{allLenHasAllSymsDFA } m &= \\ \text{minAndRen}(\text{minus}(\text{allStrDFA}, \text{someLenNotHasAllSymsDFA } m)). \end{aligned}$$

We let the DFA noBadDFA be

$$\text{minAndRen}(\text{inter}(\text{notHasStrDFA } 02, \text{inter}(\text{notHasStrDFA } 10, \text{notHasStrDFA } 21))).$$

Finally, we define $\text{ansDFA} \in \mathbb{N} \rightarrow \text{DFA}$ by:

$$\text{ansDFA } m = \text{minAndRen}(\text{inter}(\text{noBadDFA}, \text{allLenHasAllSymsDFA } m)).$$

(b) We put our definition

```

val minAndRen    = DFA.renameStatesCanonically o DFA.minimize;
val faToDFAMar  = minAndRen o nfaToDFA o efaToNFA o faToEFA;
val regToDFAMar = faToDFAMar o regToFA;

val allStrReg = Reg.fromString "(0 + 1 + 2)*";
val allStrDFA = regToDFAMar allStrReg;
val allStrFA  = injDFAToFA allStrDFA;

```

```

fun lenDFA m = regToDFAMar(Reg.power(Reg.fromString "0 + 1 + 2", m));

fun hasStrReg x =
  Reg.concat(allStrReg, Reg.concat(strToReg x, allStrReg));
fun hasStrDFA x = regToDFAMar(hasStrReg x);
fun notHasStrDFA x = minAndRen(DFA.minus(allStrDFA, hasStrDFA x));

val hasAllSymsDFA =
  minAndRen
  (DFA.inter
   (hasStrDFA(Str.fromString "0"),
    DFA.inter
     (hasStrDFA(Str.fromString "1"),
      hasStrDFA(Str.fromString "2"))));
val notHasAllSymsDFA = minAndRen(DFA.minus(allStrDFA, hasAllSymsDFA));

fun lenAndNotHasAllSymsDFA m =
  minAndRen(DFA.inter(lenDFA m, notHasAllSymsDFA));
fun lenAndNotHasAllSymsFA m = injDFAToFA(lenAndNotHasAllSymsDFA m);

fun someLenNotHasAllSymsFA m =
  FA.concat(allStrFA, FA.concat(lenAndNotHasAllSymsFA m, allStrFA));
fun someLenNotHasAllSymsDFA m = faToDFAMar(someLenNotHasAllSymsFA m);

fun allLenHasAllSymsDFA m =
  minAndRen(DFA.minus(allStrDFA, someLenNotHasAllSymsDFA m));

val noBadDFA =
  minAndRen
  (DFA.inter
   (notHasStrDFA(Str.fromString "02"),
    DFA.inter
     (notHasStrDFA(Str.fromString "10"),
      notHasStrDFA(Str.fromString "21"))));

fun ansDFA m = minAndRen(DFA.inter(noBadDFA, allLenHasAllSymsDFA m));

```

of ansDFA in the file ps6-p3.sml, and load it into Forlan:

```

- use "ps6-p3.sml";
[opening ps6-p3.sml]
val minAndRen = fn : dfa -> dfa
val faToDFAMar = fn : fa -> dfa
val regToDFAMar = fn : reg -> dfa
val allStrReg = - : reg
val allStrDFA = - : dfa
val allStrFA = - : fa
val lenDFA = fn : int -> dfa

```

```

val hasStrReg = fn : str -> reg
val hasStrDFA = fn : str -> dfa
val noHasStrDFA = fn : str -> dfa
val hasAllSymsDFA = - : dfa
val noHasAllSymsDFA = - - : dfa
val lenAndNotHasAllSymsDFA = fn : int -> dfa
val lenAndNotHasAllSymsFA = fn : int -> fa
val someLenNotHasAllSymsFA = fn : int -> fa
val someLenNotHasAllSymsDFA = fn : int -> dfa
val allLenHasAllSymsDFA = fn : int -> dfa
val noBadDFA = - : dfa
val ansDFA = fn : int -> dfa
val it = () : unit

```

Next we put our testing code

```

(* val hasBad2 : str -> bool

   hasBad2 x tests whether x has one or more of 02, 10 and 21 as a
   substring *)

fun hasBad2 x =
  Str.substr(Str.fromString "02", x) orelse
  Str.substr(Str.fromString "10", x) orelse
  Str.substr(Str.fromString "21", x)

(* val hasSym sym * str -> bool

   hasSym(a, x) tests whether a is a symbol in x *)

fun hasSym(a, x) = Str.substr([a], x);

(* val hasAllSyms : str -> bool

   hasAllSyms x tests whether x has at least one occurrence of
   all of 0, 1 and 2 *)

fun hasAllSyms x =
  hasSym(Sym.fromString "0", x) andalso
  hasSym(Sym.fromString "1", x) andalso
  hasSym(Sym.fromString "2", x)

(* val inX : int -> str -> bool

   if m >= 0 and x is in {0, 1, 2}^*, inX m x tests whether x is in X_m *)

fun inX m x =
  not(hasBad2 x) andalso
  Set.all

```

```

    (fn y => if length y = m then hasAllSyms y else true)
    (StrSet.substrings x);

(* val upto : int -> str set

   if n >= 0, then upto n returns all strings over alphabet {0, 1, 2} of
   length no more than n *)

fun upto 0 : str set = Set.sing nil
  | upto n
    =
    let val xs = upto(n - 1)
        val ys = Set.filter (fn x => length x = n - 1) xs
    in StrSet.union
        (xs, StrSet.concat(StrSet.fromString "0, 1, 2", ys))
    end;

(* val partition : int -> int -> str set * str set

   if m >= 0 and n >= 0, then partition m n returns (xs, ys) where:

   xs is all elements of upto n that are in X_m; and

   ys is all elements of upto n that are not in X_m *)

fun partition m n = Set.partition (inX m) (upto n);

(* val test : int -> int -> dfa -> str option * str option

   if m >= 0 and n >= 0, then test m n returns a function f such that,
   for all DFAs dfa, f dfa returns a pair (xOpt, yOpt) such that:

   If there is an element of {0, 1, 2}^* of length no more than n that
   is in X_m but is not accepted by dfa, then xOpt = SOME x for some
   such x; otherwise, xOpt = NONE.

   If there is an element of {0, 1, 2}^* of length no more than n that
   is not in X_m but is accepted by dfa, then yOpt = SOME y for some
   such y; otherwise, yOpt = NONE. *)

fun test m n =
  let val (goods, bads) = partition m n
      in fn dfa =>
        let val determAccepted = DFA.determAccepted dfa
            val goodNotAccOpt = Set.position (not o determAccepted) goods
            val badAccOpt      = Set.position determAccepted bads
        in ((case goodNotAccOpt of
              NONE    => NONE
              | SOME i => SOME(ListAux.sub(Set.toList goods, i))),
          badAccOpt)
        end
      end
end

```



```

                (case badAccOpt of
                  NONE    => NONE
                  | SOME i => SOME(ListAux.sub(Set.toList bads, i)))
            end
        end;

    end;

(* doit m prints the required information about ansDFA m *)

fun doit m =
  let val dfa = ansDFA m
      val bs  = DFA.alphabet dfa
      val n   = DFA.numStates dfa
      val res =
        case test m 12 dfa of
          (NONE, NONE) => "test succeeded"
          | _           => "test failed"
        in print "m = "; print(Int.toString m); print ": ";
          print "alphabet is {"; print(SymSet.toString bs); print "}; ";
          print "number of states is "; print(Int.toString n); print "; ";
          print res; print "\n"
        end;
  end;

```

in the file `ps6-p3-testing.sml`, and load it into Forlan:

```

- use "ps6-p3-testing.sml";
[opening ps6-p3-testing.sml]
val hasBad2 = fn : str -> bool
val hasSym  = fn : sym * str -> bool
val hasAllSyms = fn : str -> bool
val inX     = fn : int -> str -> bool
val upto   = fn : int -> str set
val partition = fn : int -> int -> str set * str set
val test    = fn : int -> int -> dfa -> str option * str option
val doit    = fn : int -> unit
val it     = () : unit

```

Finally, we apply the function `doit` to each of the required values of m , printing the required information about each `ansDFA m`:

```

- app doit [0, 1, 2, 3, 4, 5];
m = 0: alphabet is {}; number of states is 1; test succeeded
m = 1: alphabet is {}; number of states is 1; test succeeded
m = 2: alphabet is {0, 1, 2}; number of states is 3; test succeeded
m = 3: alphabet is {0, 1, 2}; number of states is 9; test succeeded
m = 4: alphabet is {0, 1, 2}; number of states is 18; test succeeded
m = 5: alphabet is {0, 1, 2}; number of states is 30; test succeeded
val it = () : unit

```

Note that the alphabets of the DFAs when m is 0 and 1 are empty. ($X_0 = \emptyset$, $X_1 = \{\%\}$ and $X_2 = \{\%, 0, 1, 2\}$).

(c) Easy calculations using the functions' specifications show that:

- $L(\mathbf{minAndRen} M) = L(M)$, for all DFAs M ;
- $L(\mathbf{faToDFAMar} M) = L(M)$, for all FAs M ;
- $L(\mathbf{regToDFAMar} \alpha) = L(\alpha)$, for all regular expressions α ;
- $L(\mathbf{allStrDFA}) = \{0, 1, 2\}^*$.

Define $\mathbf{Len} \in \mathbb{N} \rightarrow \mathbf{Lan}$ by

$$\mathbf{Len} m = \{ w \in \{0, 1, 2\}^* \mid |w| = m \}.$$

Because $\{0, 1, 2\}^m = \mathbf{Len} m$, for all $m \in \mathbb{N}$, we have that $L(\mathbf{lenDFA} m) = \mathbf{Len} m$, for all $m \in \mathbb{N}$.

Define $\mathbf{HasStr} \in \mathbf{Str} \rightarrow \mathbf{Lan}$ and $\mathbf{NotHasStr} \in \mathbf{Str} \rightarrow \mathbf{Lan}$ by:

$$\begin{aligned} \mathbf{HasStr} x &= \{ w \in \{0, 1, 2\}^* \mid x \text{ is a substring of } w \}, \\ \mathbf{NotHasStr} x &= \{ w \in \{0, 1, 2\}^* \mid x \text{ is not a substring of } w \}. \end{aligned}$$

Because $\mathbf{HasStr} x = \{0, 1, 2\}^* \{x\} \{0, 1, 2\}^*$, for all $x \in \{0, 1, 2\}^*$, we have that $L(\mathbf{hasStrDFA} x) = \mathbf{HasStr} x$, for all $x \in \{0, 1, 2\}^*$. And, because $\{0, 1, 2\}^* - \mathbf{HasStr} x = \mathbf{NotHasStr} x$, for all $x \in \{0, 1, 2\}^*$, and complementation corresponds to negation, we have that $L(\mathbf{notHasStrDFA} x) = \mathbf{NotHasStr} x$, for all $x \in \{0, 1, 2\}^*$.

Let

$$\mathbf{HasAllSyms} = \{ w \in \{0, 1, 2\}^* \mid \{0, 1, 2\} \subseteq \mathbf{alphabet} w \}.$$

Then we have that $L(\mathbf{hasAllSymsDFA}) = \mathbf{HasStr} 0 \cap \mathbf{HasStr} 1 \cap \mathbf{HasStr} 2 = \mathbf{HasAllSyms}$, because intersection corresponds to conjunction (“and”). Let

$$\mathbf{NotHasAllSyms} = \{ w \in \{0, 1, 2\}^* \mid \{0, 1, 2\} \not\subseteq \mathbf{alphabet} w \}.$$

Since $\{0, 1, 2\}^* - \mathbf{HasAllSyms} = \mathbf{NotHasAllSyms}$, we have that $L(\mathbf{notHasAllSymsDFA}) = \mathbf{NotHasAllSyms}$.

Define $\mathbf{LenAndNotHasAllSyms} \in \mathbb{N} \rightarrow \mathbf{Lan}$ by

$$\mathbf{LenAndNotHasAllSyms} m = \{ w \in \{0, 1, 2\}^* \mid |w| = m \text{ and } \{0, 1, 2\} \not\subseteq \mathbf{alphabet} w \}.$$

Because $\mathbf{LenAndNotHasAllSyms} m = \mathbf{Len} m \cap \mathbf{NotHasAllSyms}$, for all $m \in \mathbb{N}$, we have that $L(\mathbf{lenAndNotHasAllSymsDFA} m) = \mathbf{LenAndNotHasAllSyms} m$, for all $m \in \mathbb{N}$.

Define $\mathbf{SomeLenNotHasAllSyms} \in \mathbb{N} \rightarrow \mathbf{Lan}$ by:

$$\begin{aligned} \mathbf{SomeLenNotHasAllSyms} m = \\ \{ w \in \{0, 1, 2\}^* \mid \text{there is a substring } v \text{ of } w \text{ such that } |v| = m \text{ and } \{0, 1, 2\} \not\subseteq \mathbf{alphabet} v \}. \end{aligned}$$

Because

$$\mathbf{SomeLenNotHasAllSyms} m = \{0, 1, 2\}^* (\mathbf{LenAndNotHasAllSyms} m) \{0, 1, 2\}^*,$$

for all $m \in \mathbb{N}$, we have that $L(\mathbf{someLenNotHasAllSymsDFA} m) = \mathbf{SomeLenNotHasAllSyms} m$, for all $m \in \mathbb{N}$.

Define $\mathbf{AllLenHasAllSyms} \in \mathbb{N} \rightarrow \mathbf{Lan}$ by:

$$\mathbf{AllLenHasAllSyms} \ m = \{ w \in \{0, 1, 2\}^* \mid \text{for all substrings } v \text{ of } w, \text{ if } |v| = m, \text{ then } \{0, 1, 2\} \subseteq \mathbf{alphabet} \ v \}.$$

Because $\{0, 1, 2\}^* - \mathbf{SomeLenNotHasAllSyms} \ m = \mathbf{AllLenHasAllSyms}$, for all $m \in \mathbb{N}$, we have that $L(\mathbf{allLenHasAllSymsDFA} \ m) = \mathbf{AllLenHasAllSyms} \ m$, for all $m \in \mathbb{N}$.

Let

$$\mathbf{NoBad} = \{ w \in \{0, 1, 2\}^* \mid \text{neither } 02, 10 \text{ nor } 21 \text{ are substrings of } w \}.$$

Since $\mathbf{NoBad} = \mathbf{NotHasStr} \ 02 \cap \mathbf{NotHasStr} \ 10 \cap \mathbf{NotHasStr} \ 21$, we have that $L(\mathbf{noBadDFA}) = \mathbf{NoBad}$.

Finally, because $\mathbf{NoBad} \cap \mathbf{AllLenHasAllSyms} \ m = X_m$, for all $m \in \mathbb{N}$, we have that $L(\mathbf{ansDFA} \ m) = X_m$, for all $m \in \mathbb{N}$.