

CS 591 S2—Formal Language Theory: Integrating Experimentation and Proof—Fall 2018

Problem Set 7

Model Answers

Problem 1

First, we put the Forlan code

```
val regToGram' = faToGram o regToFA;
val faToGram' = regToGram o faToReg Reg.weaklySimplify;
```

in the file `ps7-p1-convert.sml`. Then we load `ps7-p1-convert.sml` into Forlan:

```
- use "ps7-p1-convert.sml";
[opening ps7-p1-convert.sml]
val regToGram' = fn : reg -> gram
val faToGram' = fn : fa -> gram
val it = () : unit
```

To compare `regToGram'` with `regToGram`, we proceed as follows:

```
- val reg = Reg.fromString "(01*)*0";
val reg = - : reg
- val gram1 = Gram.renameVariablesCanonically(regToGram reg);
val gram1 = - : gram
- val gram1' = Gram.renameVariablesCanonically(regToGram' reg);
val gram1' = - : gram
- Gram.output("", gram1);
{variables} A, B, C, D, E, F, G {start variable} A
{productions} A -> BC; B -> % | DB; C -> 0; D -> EF; E -> 0; F -> % | GF; G -> 1
val it = () : unit
- Gram.output("", gram1');
{variables} A, B, C, D, E, F, G, H {start variable} A
{productions}
A -> B | D; B -> 0C; C -> %; D -> 0E; E -> F; F -> A | G; G -> 1H; H -> F
val it = () : unit
```

To compare `faToGram'` with `faToGram`, we put the text

```
{states} A, B {start state} A {accepting states} B
{transitions} A, 0 -> B; A, 1 -> A; B, 0 -> A; B, 1 -> B
```

in the file `ps7-p1-fa.txt`, and then proceed as follows:

```
- val fa = FA.input "ps7-p1-fa.txt";
val fa = - : fa
- val gram2 = Gram.renameVariablesCanonically(faToGram fa);
```

```

val gram2 = - : gram
- val gram2' = Gram.renameVariablesCanonically(faToGram' fa);
val gram2'' = - : gram
- Gram.output("", gram2);
{variables} A, B {start variable} A {productions} A -> 0B | 1A; B -> % | 0A | 1B
val it = () : unit
- Gram.output("", gram2'');
{variables} A, B, C, D, E, F, G, H, I, J, K, L, M, N {start variable} A
{productions}
A -> BC; B -> % | DB; C -> EF; D -> 1; E -> 0; F -> % | GF; G -> H | I; H -> 1;
I -> JK; J -> 0; K -> LM; L -> % | NL; M -> 0; N -> 1
val it = () : unit

```

Problem 2

Suppose, toward a contradiction, that X is context-free. Thus there is an $n \in \mathbb{N}$ with the property of the Pumping Lemma for Context-free Languages, where X has been substituted for L . Let $z = 0^n 2^n 1^n 3^n$. Since $z \in \Sigma^*$ and $\mathbf{zeros} z = n = \mathbf{ones} z$ and $\mathbf{twos} z = n = \mathbf{threes} z$, we have that $z \in X$. Because $z \in X$ and $|z| = 4n \geq n$, the property of the lemma tells us there are $u, v, w, x, y \in \mathbf{Str}$ such that $z = uvwxy$ and

- (1) $|vwx| \leq n$; and
- (2) $vx \neq \%$; and
- (3) $uv^iwx^iy \in X$, for all $i \in \mathbb{N}$.

Because $0^n 2^n 1^n 3^n = z = uvwxy$, (1) tells us that vwx does not have both 0's and 1's, and vwx does not have both 2's and 3's. Thus vx does not have both 0's and 1's, and vx does not have both 2's and 3's. By (2), it will suffice to consider the following cases:

($\mathbf{zeros}(vx) \geq 1$) Thus $\mathbf{ones}(vx) = 0$, so that $\mathbf{ones}(uwy) = \mathbf{ones}(uvwxy) - \mathbf{ones}(vx) = \mathbf{ones} z - 0 = n - 0 = n$. And $\mathbf{zeros}(uwy) = \mathbf{zeros}(uvwxy) - \mathbf{zeros}(vx) = \mathbf{zeros} z - \mathbf{zeros}(vx) = n - \mathbf{zeros}(vx) < n$. But (3) tells us that $uwy = uv^0wx^0y \in X$, so that $n > \mathbf{zeros}(uwy) = \mathbf{ones}(uwy) = n$ —contradiction.

($\mathbf{ones}(vx) \geq 1$) Thus $\mathbf{zeros}(vx) = 0$, so that $\mathbf{zeros}(uwy) = \mathbf{zeros}(uvwxy) - \mathbf{zeros}(vx) = \mathbf{zeros} z - 0 = n - 0 = n$. And $\mathbf{ones}(uwy) = \mathbf{ones}(uvwxy) - \mathbf{ones}(vx) = \mathbf{ones} z - \mathbf{ones}(vx) = n - \mathbf{ones}(vx) < n$. But (3) tells us that $uwy = uv^0wx^0y \in X$, so that $n > \mathbf{ones}(uwy) = \mathbf{zeros}(uwy) = n$ —contradiction.

($\mathbf{twos}(vx) \geq 1$) Thus $\mathbf{threes}(vx) = 0$, so that $\mathbf{threes}(uwy) = \mathbf{threes}(uvwxy) - \mathbf{threes}(vx) = \mathbf{threes} z - 0 = n - 0 = n$. And $\mathbf{twos}(uwy) = \mathbf{twos}(uvwxy) - \mathbf{twos}(vx) = \mathbf{twos} z - \mathbf{twos}(vx) = n - \mathbf{twos}(vx) < n$. But (3) tells us that $uwy = uv^0wx^0y \in X$, so that $n > \mathbf{twos}(uwy) = \mathbf{threes}(uwy) = n$ —contradiction.

($\mathbf{threes}(vx) \geq 1$) Thus $\mathbf{twos}(vx) = 0$, so that $\mathbf{twos}(uwy) = \mathbf{twos}(uvwxy) - \mathbf{twos}(vx) = \mathbf{twos} z - 0 = n - 0 = n$. And $\mathbf{threes}(uwy) = \mathbf{threes}(uvwxy) - \mathbf{threes}(vx) = \mathbf{threes} z - \mathbf{threes}(vx) =$

$n - \mathbf{threes}(vx) < n$. But (3) tells us that $uwy = uv^0wx^0y \in X$, so that $n > \mathbf{threes}(uwy) = \mathbf{twos}(uwy) = n$ —contradiction.

Because we obtained a contradiction in each case, we have an overall contradiction. Thus X is not context-free.

Problem 3

(a) First, we put the text

```
{variables} A, B, <1>, <3> {start variable} A
{productions}
A -> 0A3 | B<3>;
B -> 1B2 | <1>;
<1> -> 1 | 1<1>;
<3> -> 3 | 3<3>
```

for a grammar generating $\{0^i1^j2^k3^l \mid i, j, k, l \in \mathbb{N} \text{ and } i < l \text{ and } j > k\}$ in the file `ps7-p3-orig-gram.txt`. Next we put the text

```
{states} A, B {start state} A {accepting states} A
{transitions}
A, 0 -> B; A, 1 -> B; A, 2 -> A; A, 3 -> A;
B, 0 -> A; B, 1 -> A; B, 2 -> B; B, 3 -> B
```

for a DFA accepting all elements of $\{0, 1, 2, 3\}^*$ in which the sum of the numbers of 0's and 1's is even in the file `ps7-p3-even-01-dfa.txt`. Then we load the grammar and DFA into Forlan:

```
- val origGram = Gram.input "ps7-p3-orig-gram.txt";
val origGram = - : gram
- val even01DFA = DFA.input "ps7-p3-even-01-dfa.txt";
val even01DFA = - : dfa
```

Next, we put the Forlan program

```
(* standard definitions *)

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

(* the alphabet {0, 1, 2, 3} *)

val syms0123 = SymSet.fromString "0, 1, 2, 3";

(* regular expression and DFA generating/accepting {0, 1, 2, 3}* *)

val allStrReg = Reg.closure(Reg.allSym syms0123);
val allStrDFA = minAndRen(regToDFA allStrReg);

(* symbolic relation swapping 2/3 for 0/1 *)
```

```

val swap23for01 = SymRel.fromString "(0, 2), (1, 3), (2, 0), (3, 1)";

(* DFA accepting all elements of {0, 1, 2, 3}* in which the sum of
   the numbers of 2s and 3s is odd *)

val odd23DFA =
  minAndRen
    (DFA.minus
      (allStrDFA,
        DFA.renameAlphabet(even01DFA, swap23for01)));

(* DFA accepting all elements of {0, 1, 2, 3}* in which sum of the
   numbers of 0s and 1s is even, and sum of the numbers of 2s and 3s
   is odd *)

val paritiesDFA = minAndRen(DFA.inter(even01DFA, odd23DFA));

(* grammar generating X *)

val gram0 =
  Gram.renameVariablesCanonically
    (Gram.inter(origGram, injDFAToEFA paritiesDFA));

```

in the file ps7-p3-find.sml, and proceed as follows:

```

- use "ps7-p3-find.sml";
[opening ps7-p3-find.sml]
val regToDFA = fn : reg -> dfa
val minAndRen = fn : dfa -> dfa
val syms0123 = - : sym set
val allStrReg = - : reg
val allStrDFA = - : dfa
val swap23for01 = - : sym_rel
val odd23DFA = - : dfa
val paritiesDFA = - : dfa
val gram0 = - : gram
val it = () : unit
- Gram.output("", gram0);
{variables} A, B, C, D, E, F, G, H, I, J, K, L, M {start variable} A
{productions}
A -> B; B -> DK | EM | OC3; C -> FJ | GL | OB3; D -> H | 1G2; E -> 1F2;
F -> I | 1E2; G -> 1D2; H -> 1I; I -> 1 | 1H; J -> 3L; K -> 3 | 3M; L -> 3 | 3J;
M -> 3K
val it = () : unit
- Gram.numVariables gram0;
val it = 13 : int
- Gram.numProductions gram0;
val it = 22 : int

```

In the grammar `gram0`, there are opportunities for hand-simplification using Forlan. We put the text

```

fun elimVars(gram, nil)      = gram
  | elimVars(gram, x :: xs) =
      elimVars(Gram.eliminateVariable(gram, Sym.fromString x), xs);

(* hand-simplified grammar generating X *)

val gram1 =
  Gram.renameVariablesCanonically
    (elimVars(Gram.restart gram0, ["E", "G", "H", "J", "M"]));

```

in the file `ps7-p3-elim.sml`, and proceed as follows:

```

- use "ps7-p3-elim.sml";
[opening ps7-p3-elim.sml]
val elimVars = fn : gram * string list -> gram
val gram1 = - : gram
val it = () : unit
- Gram.output("", gram1);
{variables} A, B, C, D, E, F, G {start variable} A
{productions}
A -> CF | 0B3 | 1D23F; B -> 0A3 | D3G | 1C2G; C -> 1E | 11C22; D -> E | 11D22;
E -> 1 | 11E; F -> 3 | 33F; G -> 3 | 33G
val it = () : unit
- Gram.numVariables gram1;
val it = 7 : int
- Gram.numProductions gram1;
val it = 16 : int

```

We finish up by making some final improvements to `gram1`, working outside of Forlan. First, note that in `gram1`, F and its productions, and G and its productions, have the same form. There is no reason to have both of them, and so we can remove G and its productions, replacing all occurrences of G in the remaining productions by F. This gives us the grammar:

$$\begin{aligned}
A &\rightarrow CF \mid 0B3 \mid 1D23F, \\
B &\rightarrow 0A3 \mid D3F \mid 1C2F, \\
C &\rightarrow 1E \mid 11C22, \\
D &\rightarrow E \mid 11D22, \\
E &\rightarrow 1 \mid 11E, \\
F &\rightarrow 3 \mid 33F.
\end{aligned}$$

Next, we note that

$$\begin{aligned}
\Pi_E &= \{1^m \mid m \in \mathbb{N} \text{ and } m \text{ is odd}\}, \\
\Pi_D &= \{1^{2n}1^m2^{2n} \mid n, m \in \mathbb{N} \text{ and } m \text{ is odd}\}, \\
\Pi_C &= \{1^{2n}11^m2^{2n} \mid n, m \in \mathbb{N} \text{ and } m \text{ is odd}\} = \{11^{2n}1^m2^{2n} \mid n, m \in \mathbb{N} \text{ and } m \text{ is odd}\} = \{1\}\Pi_D.
\end{aligned}$$

Thus we can remove C and its productions, replacing all occurrences of C by 1D:

$$\begin{aligned} A &\rightarrow 1DF \mid 0B3 \mid 1D23F, \\ B &\rightarrow 0A3 \mid D3F \mid 11D2F, \\ D &\rightarrow E \mid 11D22, \\ E &\rightarrow 1 \mid 11E, \\ F &\rightarrow 3 \mid 33F. \end{aligned}$$

Since E is only used in a production of D, we can combine the productions of D and E, yielding

$$D \rightarrow 1 \mid 11D \mid 11D22.$$

Then, because B's productions don't involve self-recursion, we can get rid of B and its productions, by substituting its right sides into the only production using B, $A \rightarrow 0B3$, obtaining:

$$\begin{aligned} A &\rightarrow 1DF \mid 00A33 \mid 0D3F3 \mid 011D2F3 \mid 1D23F, \\ D &\rightarrow 1 \mid 11D \mid 11D22, \\ F &\rightarrow 3 \mid 33F. \end{aligned}$$

And, because $\Pi_F = \{3^m \mid m \in \mathbb{N} \text{ and } m \text{ is odd}\}$, we can replace F3 by 3F in the right sides of productions, yielding:

$$\begin{aligned} A &\rightarrow 1DF \mid 00A33 \mid 0D33F \mid 011D23F \mid 1D23F, \\ D &\rightarrow 1 \mid 11D \mid 11D22, \\ F &\rightarrow 3 \mid 33F. \end{aligned}$$

To rename the variables of this grammar canonically, we can return to Forlan:

```
- val gram = Gram.renameVariablesCanonically(Gram.input "");
@ {variables} A, D, F {start variable} A
@ {productions}
@ A -> 1DF | 00A33 | 0D33F | 011D23F | 1D23F;
@ D -> 1 | 11D | 11D22;
@ F -> 3 | 33F
@ .
val gram = - : gram
- Gram.output("", gram);
{variables} A, B, C {start variable} A
{productions}
A -> 1BC | 00A33 | 0B33C | 1B23C | 011B23C; B -> 1 | 11B | 11B22; C -> 3 | 33C
val it = () : unit
- Gram.numVariables gram;
val it = 3 : int
- Gram.numProductions gram;
val it = 10 : int
```

The grammar `gram` is our answer, but `gram0` will be useful in part (b).

(b) We put the Forlan program


```
1111111111113333333333333333, 11111111112222222223333333, 11111111112222222233333333,
1111111111222222233333333333, 1111111111222222333333333333, 1111111111222223333333333333,
1111111111222333333333333333, 1111111111222333333333333333, 1111111111223333333333333333,
1111111111233333333333333333, 1111111111333333333333333333, 11111111222222333333333333,
1111111122222333333333333333, 11111111222233333333333333, 11111111222333333333333333,
11111111223333333333333333, 11111111233333333333333333, 11111112333333333333333333,
11111113333333333333333333, 11111222233333333333333333, 111112223333333333333333,
111112233333333333333333, 111112333333333333333333, 111112333333333333333333,
111113333333333333333333, 111122233333333333333333, 111123333333333333333333,
111123333333333333333333, 111133333333333333333333, 112333333333333333333333,
113333333333333333333333
val it = () : unit
```

On my laptop, it took about five minutes for `genLen` to return. I don't know how long the computation using `gram` would have taken—maybe too long to get an answer?