CS 591 S2—Formal Language Theory: Integrating Experimentation and Proof—Fall 2019

Problem Set 6

Model Answers

Problem 1

Easy mathematical inductions show that for all $n \in \mathbb{N}$, $\mathbf{diff}(\mathsf{1}^n) = n$ and $\mathbf{diff}(\mathsf{0}^n) = -2n$.

Let X be the least subset of $\{\mathsf{0}, \mathsf{1}\}^*$ such that:

(1) $\% \in X$;

(2) $\mathsf{1} \in X$;

(3) for all $x, y \in X$, $\mathsf{1}x\mathsf{1}y\mathsf{0} \in X$;

(4) for all $x, y \in X$, $xy \in X$.

In Problem Set 2, we proved $X = Y$.

Lemma PS6.1.1
For all $n \in \mathbb{N}$, $\mathsf{1}^{2n}\mathsf{0}^n \in Y$.

Proof. Because $Y = X$, it will suffice to show that, for all $n \in \mathbb{N}$, $\mathsf{1}^{2n}\mathsf{0}^n \in X$. We proceed by mathematical induction.

- **(basis step)** We have that $\mathsf{1}^{2*0}\mathsf{0}^0 = \mathsf{1}^0\mathsf{0}^0 = \%\% = \% \in X$, by Rule (1) of X's definition.
- **(inductive step)** Suppose $n \in \mathbb{N}$, and assume the inductive hypothesis: $\mathsf{1}^{2n}\mathsf{0}^n \in X$. Then $\mathsf{1}^{2(n+1)}\mathsf{0}^{n+1} = \mathsf{1}^{2n+2}\mathsf{0}^n\mathsf{0} = \mathsf{1}^{1+1+2n}\mathsf{0}^n\mathsf{0} = \mathsf{1}\mathsf{1}\mathsf{1}^{2n}\mathsf{0}^n\mathsf{0} = \mathsf{1}(\%)\mathsf{1}(\mathsf{1}^{2n}\mathsf{0}^n)\mathsf{0} \in X$, by Rule (3) of X's definition, since $\% \in X$ (by Rule (1) of X's definition) and $\mathsf{1}^{2n}\mathsf{0}^n \in X$ (by the inductive hypothesis).

$\square$

Suppose, toward a contradiction, that Y is regular. Thus there is an $n \in \mathbb{N} - \{0\}$ with the property of the Pumping Lemma, where Y has been substituted for L. Suppose $z = \mathsf{1}^{2n}\mathsf{0}^n$. By Lemma PS6.1.1, we have that $z \in Y$. Thus, since $|z| = 2n+n = 3n \geq n$, it follows there are $u, v, w \in$ **Str** such that $z = uvw$ and properties (1)–(3) of the lemma hold. Since $uvw = z = \mathsf{1}^{2n}\mathsf{0}^n = \mathsf{1}^n\mathsf{1}^n\mathsf{0}^n$, (1) tells us that there are $i, j, k \in \mathbb{N}$ such that

$$u = \mathsf{1}^i, \quad v = \mathsf{1}^j, \quad w = \mathsf{1}^k\mathsf{1}^n\mathsf{0}^n, \quad i+j+k = n.$$

By (2), we have that $j \geq 1$, and thus that $i + k = n - j < n$. By (3), we have that $\mathsf{1}^{i+k+n}\mathsf{0}^n = \mathsf{1}^i\mathsf{1}^k\mathsf{1}^n\mathsf{0}^n = uw = u\%w = uv^0w \in Y$. Because $\mathsf{1}^{i+k+n}\mathsf{0}^n$ is a prefix of itself, we have that $i+k-n = i+k+n+-n+-n = (i+k+n)+-2n = \mathbf{diff}(\mathsf{1}^{i+k+n}) + \mathbf{diff}(\mathsf{0}^n) = \mathbf{diff}(\mathsf{1}^{i+k+n}\mathsf{0}^n) \geq 0$. But since $i+k < n$, we have that $i+k-n < 0$—contradiction. Thus Y is not regular.

Problem 2

Let G be the grammar

$$\mathsf{A} \to \% \mid \mathsf{1} \mid \mathsf{1A1A0} \mid \mathsf{AA}$$

Because $Y = X$ (see Problem 1), it will suffice to show that $L(G) = X$.

Lemma PS6.2.1
For all $w \in \Pi_{\mathsf{A}}$, $w \in X$.

Proof. We proceed by induction on Π. There are four productions to consider.

($\mathsf{A} \to \%$) We must show that $\% \in X$, and this follows by Rule (1) of X's definition.

($\mathsf{A} \to \mathsf{1}$) We must show that $\mathsf{1} \in X$, and this follows by Rule (2) of X's definition.

($\mathsf{A} \to \mathsf{1A1A0}$) Suppose $x, y \in \Pi_{\mathsf{A}}$, and assume the inductive hypothesis: $x, y \in X$. Then $\mathsf{1}x\mathsf{1}y\mathsf{0} \in X$ by Rule 3 of X's definition.

($\mathsf{A} \to \mathsf{AA}$) Suppose $x, y \in \Pi_{\mathsf{A}}$, and assume the inductive hypothesis: $x, y \in X$. Then $xy \in X$ by Rule 4 of X's definition.

□

Lemma PS6.2.2
For all $w \in X$, $w \in \Pi_{\mathsf{A}}$.

Proof. We proceed by induction on X. There are four steps to show.

(1) We must show that $\% \in \Pi_{\mathsf{A}}$. And this follows because of the production $\mathsf{A} \to \%$ of G.

(2) We must show that $\mathsf{1} \in \Pi_{\mathsf{A}}$. And this follows because of the production $\mathsf{A} \to \mathsf{1}$ of G.

(3) Suppose $x, y \in X$, and assume the inductive hypothesis: $x, y \in \Pi_{\mathsf{A}}$. We must show that $\mathsf{1}x\mathsf{1}y\mathsf{0} \in \Pi_{\mathsf{A}}$, and this follows because of the production $\mathsf{A} \to \mathsf{1A1A0}$ and the inductive hypothesis.

(4) Suppose $x, y \in X$, and assume the inductive hypothesis: $x, y \in \Pi_{\mathsf{A}}$. We must show that $xy \in \Pi_{\mathsf{A}}$, and this follows because of the production $\mathsf{A} \to \mathsf{AA}$ and the inductive hypothesis.

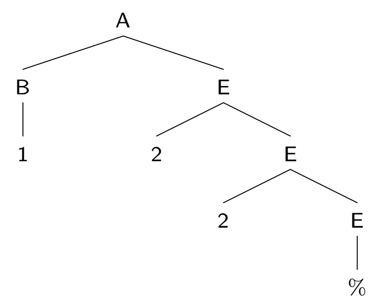
□

Lemma PS6.2.1 tells us that $L(G) = \Pi_{\mathsf{A}} \subseteq X$, and Lemma PS6.2.2 tells us that $X \subseteq \Pi_{\mathsf{A}} = L(G)$. Thus $L(G) = X$.

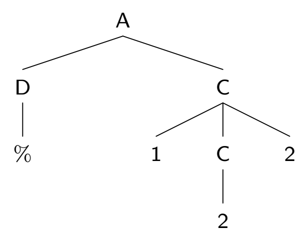
Problem 3

(a) $L(G) = \{\, 0^i 1^j 2^k \mid i, j, k \in \mathbb{N} \text{ and } (i < j \text{ or } j < k) \,\}$.

(b) Let pt_1 be the parse tree



And let pt_2 be the parse tree



To check that our answer is correct, we proceed as follows:

```
- fun amb(gram, pt1, pt2) =
=       not(PT.equal(pt1, pt2))                                         andalso
=       Gram.validPT gram pt1 andalso Gram.validPT gram pt2             andalso
=       Sym.equal(Gram.startVariable gram, PT.rootLabel pt1)            andalso
=       Sym.equal(Gram.startVariable gram, PT.rootLabel pt2)            andalso
=       Str.equal(PT.yield pt1, PT.yield pt2)                           andalso
=       SymSet.subset(Str.alphabet(PT.yield pt1), Gram.alphabet gram);
val amb = fn : gram * pt * pt -> bool
- val gram = Gram.input "";
@ {variables} A, B, C, D, E {start variable} A
@ {productions}
@ A -> BE | DC; B -> 1 | B1 | 0B1; C -> 2 | C2 | 1C2;
@ D -> % | 0D; E -> % | 2E
@ .
val gram = - : gram
- val pt1 = PT.fromString "A(B(1), E(2, E(2, E(%))))";
val pt1 = - : pt
- val pt2 = PT.fromString "A(D(%), C(1, C(2), 2))";
val pt2 = - : pt
- amb(gram, pt1, pt2);
val it = true : bool
```