# Chapter 8

# Context-free and Noncontext-free Languages

8.1. (a) Suppose $L$ is a CFL, and let $n$ be the integer in the pumping lemma. Let $u = a^n b^{n+1} c^{n+2}$. Then $u = vwxyz$ for some $v$, $w$, $x$, $y$, and $z$ satisfying $|wy| > 0$, $|wxy| \leq n$, and $vw^ixy^iz \in L$ for every $i \geq 0$.

First we consider the case when $wy$ contains at least one $a$. Then since $|wxy| \leq n$, $wy$ can contain no $c$'s. Therefore, $vw^2xy^2z$ has at least $n+1$ $a$'s and exactly $n+2$ $c$'s, which is impossible if the string is in $L$.

If $wy$ contains no $a$'s, then it must contain either $b$ or $c$. In this case, $vw^0xy^0z = vxz$ has either fewer than $n+1$ $b$'s or fewer than $n+2$ $c$'s, but in either case exactly $n$ $a$'s. This is also impossible if this string is in $L$. Thus in either case we have derived a contradiction, and we conclude that $L$ cannot be a context-free language.

(b) Suppose $L$ is a CFL, and let $n$ be the integer in the pumping lemma. Let $u = a^n b^{n^2}$. Then $u = vwxyz$, where the same three conditions on $v$, $w$, $x$, $y$, and $z$ hold. Let $n_a(wy) = p$ and $n_b(wy) = q$. Then for each $i$, we have

$$n_a(vw^{i+1}xy^{i+1}z) = n + ip \qquad\qquad n_b(vw^{i+1}xy^{i+1}z) = n^2 + iq$$

Using the definition of $L$, we may conclude that $n^2 + iq = (n + ip)^2$ for every $i \geq 0$. However, it is easy to see that this is impossible. The numbers $p$ and $q$ cannot both be 0, since $|wy| > 0$, and if one of them is 0 we have a contradiction. Suppose both are positive. Then using $i = 1$, we obtain $q = 2pn + p^2$, and using $i = 2$ we obtain $q = 2pn + 2p^2$. These two equations imply that $p = 0$, and we obtain a contradiction. Therefore, $L$ is not a CFL.

(c) Suppose $L$ is a CFL, and let $n$ be the integer in the pumping lemma. Let $u = a^n b^{2n} a^n$. Then $u = vwxyz$, where the same three conditions on $v$, $w$, $x$, $y$, and $z$ hold. If $wy$ contains at least one $a$ from the first group, then since $|wxy| \leq n$, $wy$ can contain no $a$'s from the second group. Then $vxz$ contains $n$ $a$'s at the end and fewer than $n$ at the beginning, so that it is not in $L$. Similarly if $wy$ contains at least one $a$ from the second group. The only remaining case is when $wy$ contains no $a$'s, and in this case $vxz$ contains fewer than $2n$ $b$'s. In each case we have a contradiction.

(d) Suppose $L$ is a CFL, and let $n$ be the integer in the pumping lemma. Let $u = a^n b^n c^n$. Then $u = vwxyz$, where the same three conditions on $v$, $w$, $x$, $y$, and $z$ hold.

If $wy$ contains at least one $a$, then since $|wxy| \leq n$, $wy$ can contain no $c$'s. Therefore, $vw^0xy^0z$ contains fewer than $n$ $a$'s but exactly $n$ $c$'s, and so it is impossible for this string to be in $L$. If $wy$ contains no $a$'s, then $vw^2xy^2z$ contains either more than $n$ $b$'s or more than $n$ $c$'s, but exactly $n$ $a$'s. In this case also, the string cannot be in $L$. Therefore, we have a contradiction.

8.2. Some choices that would not work include $a^{2n}$, $(ab)^n(ab)^n$, and $a^nba^nb$.

8.4. Yes. In the application of Ogden's lemma, we can designate as distinguished all the positions of $u$ except those containing semicolons. This eliminates the possibility that $wy$ consists of only a a semicolon.

8.5. (a) Yes. A CFG generating $L$ has productions $S \to aSb \mid T \qquad T \to bTa \mid \Lambda$.

(b) Yes. A CFG generating $L$ has productions $S \to Tb \qquad T \to aTa \mid aTb \mid bTa \mid bTb \mid a$.

(c) No. A proof can be constructed involving the pumping lemma that is very similar to the proof in Example 8.2.

(d) No. Suppose $L$ is a CFG, and let $n$ be the integer in the pumping lemma. Let $u$ be the string $a^n b^n a^n b^n$ (i.e., $x = a^n b^n$ and $y = \Lambda$). Then $u = vwxyz$, where $|wy| > 0$, $|wxy| \leq n$, and $vw^ixy^iz \in L$ for every $i \geq 0$.

Suppose first that $wy$ contains either only $a$'s from the first group or only $b$'s from the last group. Then $vw^2xy^2z$ is either $a^{n+i}b^na^nb^n$ or $a^nb^na^nb^{n+i}$ for some $i > 0$, and in neither case can this string be of the form $sts$ for any $s$ with $|s| > 0$.

Otherwise, $wy$ contains either a $b$ from the first group or an $a$ from the second. In this case $vw^0xy^0z$ is either $a^ib^ja^kb^n$ or $a^nb^ia^jb^k$, where in either case $i$ and $k$ are positive and $j < n$. Neither of these strings can be of the form required for $L$ either. Therefore, we have a contradiction, and $L$ is not a CFG.

(e) Yes. See Exercise 7.37(b).

(f) Yes. A PDA (in fact, a counter automaton) can be constructed that accepts the language, following the general model in the solution to Exercise 7.18(b). Roughly speaking, the contents of the stack measures the absolute value of $d(x) = 10n_b(x) - n_a(x)$, and the current state indicates whether the actual value is positive or not.

(g) Yes. The set of balanced strings of parentheses is a DCFL, and it follows from the discussion at the end of Section 8.2 that its complement is also.

8.6. The generalization of Theorem 5.3 says that if $L$ is an infinite CFL, then there are strings $v$, $w$, $x$, $y$, and $z$ with $|wy| > 0$ and $vw^ixy^iz \in L$ for every $i \geq 0$. The generalization of Theorem 5.4 says that if $L$ is an infinite CFL, the set of integers that are lengths of elements of $L$ contains an infinite arithmetic progression.

Let $L_1 = \{x \in \{a,b,c\}^* \mid n_a(x) = n_b(x) = n_c(x)\}$, $L_2 = \{a^ib^ic^i \mid i \geq 0\}$, and $L_3 = \{a^{n^2} \mid n \geq 0\}$. Theorem 8.1a can be used to show that $L_1$ is not a CFL, but the generalization of Theorem 5.3 cannot. (There is no guarantee that the string $wy$ doesn't have equal numbers of $a$'s, $b$'s, and $c$'s.) The generalization of Theorem 5.3 can be used to show $L_2$ is not a CFL, but the generalization of Theorem 5.4 cannot. Finally, the generalization of Theorem 5.4 can be used to show that $L_3$ is not a CFL. (See Exercise 5.27(h).)

8.7. The proof of Theorem 8.4 actually proves that if $M_1$ is a DPDA accepting $L_1$ and $M_2$ is an FA recognizing $L_2$, then the PDA $M$ that we constructed is a DPDA accepting $L_1 \cap L_2$.

8.8. (a) Call this language $L$. Then $L$ is the union of the two languages $\{a^ib^jc^k \mid i \geq j\}$ and $\{a^ib^jc^k \mid i \geq k$, each of which is easily seen to be a CFL. The complement of $L$ in $\{a,b,c\}^*$ is $L' = (\{a\}^*\{b\}^*\{c\}^*)' \cup \{a^ib^jc^k \mid i < j \text{ and } i < k\}$. From the formula $L' \cap \{a\}^*\{b\}^*\{c\}^* = \{a^ib^jc^k \mid i < j \text{ and } i < k\}$ and Theorem 8.4, it follows that if $L$ is a CFL, then so is $\{a^ib^jc^k \mid i < j \text{ and } i < k\}$. However, we can show using the pumping lemma that this is not the case. Suppose $\{a^ib^jc^k \mid i < j \text{ and } i < k\}$ is a CFL and let $n$ be the integer in the pumping lemma. Let $u = a^nb^{n+1}c^{n+1}$. Then $u = vwxyz$, where

the usual conditions hold. If $wy$ contains $a$'s, it can contain no $c$'s, and we may obtain a contradiction by considering $vw^2xy^2z$. If $wy$ contains no $a$'s, we may obtain a contradiction by considering $vxz$.

(b) The proof is similar to the preceding proof.

8.9. (a) Suppose $L = \{a^ib^{i+k}a^k \mid k \neq i\}$ is a CFL. Let $n$ be the integer in Ogden's lemma, let $u = a^nb^{2n+n!}a^{n+n!}$, and designate the first $n$ positions of $u$ as the distinguished positions. Then $u = vwxyz$, where $wy$ contains at least one from the first group of $a$'s and $vw^ixy^iz \in L$ for every $i \geq 0$. If either $w$ or $y$ contained both $a$'s and $b$'s, then clearly $vw^2xy^2z$ would not be in $L$. Also, if neither $w$ nor $y$ contained any $b$'s, then $vw^2xy^2z$ would not preserve the balance between $a$'s and $b$'s required for membership in $L$. Therefore, $w$ contains only $a$'s from the first group, and $y$ contains only $b$'s, and in fact the lengths of these strings are equal, say $p$. Now, however, let $i = 1 + n!/p$. Then $vw^ixy^iz = a^{n+(i-1)p}b^{2n+n!+(i-1)p}a^{n+n!} = a^{n+n!}b^{2n+2n!}a^{n+n!}$, which is clearly not an element of $L$. This contradiction implies that $L$ is not a CFL.

(c) Suppose $L = \{a^ib^ja^i \mid j \neq i\}$ is a CFL. Let $n$ be the integer in Ogden's lemma, let $u = a^nb^{n+n!}a^n$, and designate the first $n$ positions of $u$ as the distinguished positions. Then $u = vwxyz$, where $wy$ contains at least one from the first group of $a$'s and $vw^ixy^iz \in L$ for every $i \geq 0$. The only case in which considering $vw^2xy^2z$ does not easily produce a contradiction is that in which $w$ contains $k$ $a$'s from the first group for some $k > 0$ and $y$ contains $k$ $a$'s from the second group. In this case, however, we can let $i = 1 + n!/k$, so that $vw^ixy^iz = a^{n+n!}b^{n+n!}a^{n+n!}$. This string is not in $L$, and so we also have a contradiction in this case.

8.10. (a) Since finite sets are regular, $F$ is regular, and since the complement of a regular set is regular, $F'$ is regular. Therefore, by Theorem 8.4, $L - F = L \cap F'$ is a CFL.

(b) We use the formula $(L - F) \cup (L \cap F) = L$. If $L - F$ were a CFL, then since $L \cap F$ is finite and therefore a CFL, $L$ would be a CFL.

(c) Here we use the formula $L \cup F - (F - L) = L$. If $L \cup F$ were a CFL, then since $F - L$ is finite, it would follow from (a) that $L$ is a CFL.

8.11. Part (a) is still true in the more general case, but (b) and (c) are not, since in each case we could take $F$ to be $\Sigma^*$.

8.12. All three statements are true.

8.13. The DPDA given in the solution to Exercise 7.13(c) has this property. For the input string $ab$, for example, we have

$$(q_0, ab, Z_0) \vdash (q_0, b, aZ_0) \vdash (q_t, \Lambda, Z_0) \vdash (q_1, \Lambda, Z_0)$$

At the point when the PDA enters $q_t$, the string is accepted. However, $ab$ is an element of $L$, not $L'$.

8.14. Yes. It is not difficult to show that if $G$ is a CFG generating $L$, then the grammar obtained from $G$ by reversing the right sides of all the productions generates $rev(L)$.

8.15. (a) No. Suppose $L$ is a CFG, and let $n$ be the integer in the pumping lemma. Let $u = b^p a^{2np}$, where $p$ is an integer large enough so that $2n^2/p < 1$ (for reasons that will be clear shortly). Then $u = vwxyz$, where $|wy| > 0$, $|wxy| \leq n$, and $vw^ixy^iz \in L$ for every $i \geq 0$.

The string $vw^2xy^2z$ has $2np + j$ $a$'s and $p + k$ $b$'s, where $0 \leq j \leq n$ and $0 \leq k \leq n$ Consider the ratio $n_a/n_b$ for this string. It is

$$\frac{2np + j}{p + k} = 2n + \frac{j - 2nk}{p + k}$$

The fraction on the right side has absolute value no larger than $2n^2/p$, which by assumption is less than 1. However, it is nonzero. ($j$ and $k$ cannot both be 0; if one is zero, the numerator is clearly nonzero; and if both are nonzero, $j - 2nk \neq 0$ because $j < n$.) Therefore, the original ratio cannot be an integer, which means that the string $vw^2xy^2z$ can't be in $L$. This contradiction implies that $L$ is not a CFL.

(b) Yes. Given a PDA $M = (Q, \Sigma, \Gamma, q_0, Z_0, A, \delta)$ accepting $L$, we can construct another PDA $M_1 = (Q_1, \Sigma, \Gamma, q_1, Z_0, A_1, \delta_1)$ accepting the set of prefixes of elements of $L$ as follows. $Q_1$ can be taken to be $Q \times \{0, 1\}$—in other words, a set containing two copies $(q, 0)$ and $(q, 1)$ of each element of $Q$. The initial state $q_1$ is $(q_0, 0)$, and the set $A_1$ is $A \times \{1\}$. For each combination $(q, 0)$ (where $q \in Q$), $a \in \Sigma$, and $X \in \Gamma$, the set $\delta_1((q, 0), a, X)$ is simply $\{((p, 0), \alpha) \mid (p, \alpha) \in \delta(q, a, X)\}$; however, $\delta_1((q, 0), \Lambda, X)$ contains not only the elements $((p, 0), \alpha)$ for which $(p, \alpha) \in \delta(q, \Lambda, X)$, but one additional element, $((q, 1), X)$. For each $q \in Q$, $a \in \Sigma$, and $X \in \Gamma$, $\delta_1((q, 1), a, X) = \emptyset$. Finally, for each $q \in Q$ and each $X \in \Gamma$, $\delta_1((q, 1), \Lambda, X)$ is the union of the sets $\{((p, 1), \alpha) \mid (p, \alpha) \in \delta(q, a, X)\}$ over all $a \in \Sigma \cup \{\Lambda\}$.

The idea here is that $M_1$ starts out in state $(q_0, 0)$ and acts on the states $(q, 0)$ exactly the same way $M$ acts on the states $q$, until such time as $M_1$ takes a $\Lambda$-transition from its current state $(q, 0)$ to $(q, 1)$. From this point on, $M_1$ can make the same moves on the states $(q, 1)$ that $M$ can make on the states $q$, changing the stack in the same way, but without reading any input. If $xy \in L$, then for any sequence of moves $M$ makes corresponding to $xy$, ending in the state $r$, one possible sequence of moves $M_1$ can make on $xy$ is to copy the moves of $M$ while processing $x$, reaching some state $(q_x, 0)$, then make a $\Lambda$-transition to $(q_x, 1)$, then continue simulating the sequence on the states $(p, 1)$ but making only $\Lambda$-transitions, ending at the state $(r, 1)$ having processed the string $x$. Therefore, $x$ is accepted by $M_1$. Conversely, if $x$ is accepted by $M_1$, there is a sequence of moves corresponding to $x$ that ends at a state $(p, 1)$, where $p \in A$. In this case, the moves of the sequence that involve states $(q, 0)$, ending at $(q_x, 0)$, correspond to moves $M$ can make processing $x$, ending at the state $q_x$; and the remaining $\Lambda$-transitions that end at $(p, 1)$ correspond to transitions that $M$ can make using the symbols of some string $y$. Therefore, if $M_1$ accepts $x$, then there is a string $y$ so that $M$ accepts $xy$.

(c) and (d) Yes. The argument in part (i) can be adapted for each of these parts.

(f) No. Suppose $L$ is a CFG, and let $n$ be the integer in the pumping lemma. Let $u = a^p b^p c^p$, where $p$ is a prime larger than $n$. Then $u = vwxyz$, where $|wy| > 0$, $|wxy| \leq n$,

and $vw^ixy^iz \in L$ for every $i \geq 0$. These conditions imply that $wy$ contains no more than two distinct symbols. Therefore, if $j = n_a(vw^0xy^0z = n_a(vxz)$, $k = n_b(vxz)$, and $m = n_c(vxz)$, at least one the three integers $j$, $k$, $m$ is $p$, at least one is less than $p$, and all three are positive. Since $p$ is prime, it is not possible for all three to have a nontrivial common factor. Therefore, $L$ is not a CFG.

8.16. As suggested in the hint, let $L_1 = \{x\#y\#z \mid xyz \in L\}$. Then it is easy to see that if there is a PDA $M$ accepting $L$, there is another one $M_1$ accepting $L_1$. ($M_1$ can simply ignore the # symbols during the processing, except that it enters an accepting state only if it has seen exactly two such symbols during the processing.)

Let $n$ be the integer in Ogden's lemma applied to $L_1$, and let $u_1u_2u_3$ be any string in $L$, where $|u_2| \geq n$. Then consider the string $u = u_1\#u_2\#u_3 \in L_1$, and specify all the positions in the substring $u_2$ to be distinguished. Then according to Ogden's lemma, $u = v'w'x'y'z'$, where $w'y'$ contains at least one distinguished position and $v'(w')^ix'(y')^iz' \in L_1$ for every $i \geq 0$. Suppose that $w'$ contains one of the distinguished positions. Then $w'$ must be a substring of $u_2$, for otherwise we would have strings with more than two #'s in $L_1$; similarly if $y'$ contains a distinguished position. Now let $v$, $w$, $x$, $y$, and $z$ be $v'$, $w'$, $x'$, $y'$, $z'$, respectively, except that any occurrences of # are omitted. Then the strings $v$, $w$, $x$, $y$, and $z$ satisfy the desired conditions.

8.17. (c) Suppose $L = \{a^ib^ja^i \mid j \neq i\}$ is a CFL. Let $n$ be the integer in Exercise 8.16, let $u = a^nb^{n+n!}a^n$, and let $u_1 = \Lambda$, $u_2 = a^n$, and $u_3 = b^{n+n!}a^n$. Then $u = vwxyz$, where $w$ or $y$ is a nonempty substring of the first group of $a$'s and $vw^ixy^iz \in L$ for every $i \geq 0$. From this point on, the contradiction is obtained in almost the same way as in the solution to Exercise 8.9.

8.18. Let $L$ be the language in Example 8.5, $\{a^ib^ic^j \mid i, j \geq 0 \text{ and } j \neq i\}$. Suppose $L$ is a CFL, and let $n$ be the integer in the statement in Exercise 8.16. Let $u = a^nb^nc^{n+n!}$, as in Example 8.5, and let the string $u_2$ be $a^n$. Then there are strings $v$, $w$, $x$, $y$, and $z$ so that either $w$ or $y$ is a nonnull substring of $a^n$ and $vw^ixy^iz \in L$ for every $i \geq 0$. (Note: the statement in Exercise 8.16 says only that $w$ or $y$ is a substring of $u_2$, but the proof given in the solution implies that we can require it to be nonnull.) From this point on, the argument follows that in Example 8.5.

Let $L_1$ be the language in Example 8.6, $\{a^pb^qc^rd^s \mid p = 0 \text{ or } q = r = s\}$. Let $n$ be the integer in Exercise 8.16, let $u = ab^nc^nd^n$, and let $u_2$ be the substring $c^n$. Then $u = vwxyz$, where either $w$ or $y$ is a nonnull substring of $c^n$ and $vw^ixy^iz \in L_1$ for every $i \geq 0$. It follows that $wy$ cannot contain all three of the symbols $b$, $c$, and $d$. We can now finish the argument as in Example 8.6.

8.19. The class of DCFLs is not closed under any of these operations.

(a) The languages $L_1$ and $L_2$ defined in the proof of Theorem 8.3 are both DCFLs. The language $(L_1 \cup L_2)' \cap \{a\}^*\{b\}^*\{c\}^*$ is $\{a^ib^jc^k \mid i \geq j \text{ and } i \geq k\}$, which can be shown by the pumping lemma not to be a CFL. It follows from Theorem 8.4 that $(L_1 \cup L_2)'$ is not a CFL, and therefore that $L_1 \cup L_2$ is not a DCFL.

(b) We have the formula $L_1 \cup L_2 = (L_1' \cap L_2')'$. If the set of DCFLs were closed under intersection, it would follow from this formula that it was also closed under union.

(c) Again look at the DCFLs $L_1$ and $L_2$ in the proof of Theorem 8.3. Let $L_3 = \{d\}L_1 \cup L_2$. Then $L_3$ is a DCFL, because in order to accept it, the presence or absence of an initial $d$ is all that needs to be checked before executing the appropriate DPDA to accept either $L_1$ or $L_2$. The language $\{d\}^*$ is also a DCFL. However, we can check that $\{d\}^*L_3$ is not a DCFL. The reason is that $\{d\}^*L_3 \cap \{d\}\{a\}^*\{b\}^*\{c\}^* = \{d\}L_1 \cup \{d\}L_2$. If $\{d\}^*L_3$ were a DCFL, then this intersection would be also (see Exercise 8.7), and it follows easily that $L_1 \cup L_2$ would be as well.

(d) Let $L_4 = \{d\} \cup \{d\}L_1 \cup L_2$, where $L_1$ and $L_2$ are as above. Then $L_4$ is a DCFL. It is not hard to see that

$$L_4^* \cap \{d\}\{a\}^*\{b\}^*\{c\}^* = \{d\}L_1 \cup \{d\}L_2$$

Therefore, the same argument used in (c) implies that $L_4^*$ is not a DCFL.

(e) We have the formula $A \cap B = A - B'$. If the set of DCFLs were closed under the difference operation, then since it is closed under complements, it would be closed under intersections.

8.20. (a) Let $L_1 = \{x\#y \mid x \in \textit{pal} \text{ and } xy \in \textit{pal}\}$. Then if *pal* were a DCFL, $L_1$ would be, and therefore, by Exercise 8.7, $L_1 \cap \{0\}^*\{1\}^*\{0\}^*\{\#\}\{1\}^*\{0\}^*$ would be also. However, this intersection is $\{0^i1^j0^i\#1^j0^i \mid i, j \geq 0\}$, and it's easy to show using the pumping lemma that this language is not even a CFL.

(b) Let this language be $L$, and let $L_1 = \{x\#y \mid x \in L \text{ and } xy \in L\}$. If $L$ were a DCFL, then $L_1$ would be, and so $L_1 \cap \{a\}^+\{b\}^*\{\#\}\{a\}^+$ would be. However, consider what it means for a string $a^ib^j\#a^k$ (with $i, j, k > 0$) to be in this intersection. Saying that the prefix up to the $\#$ is in $L$ means that $j$ is either $i$ or $2i$, and saying that $a^ib^ja^k$ is in $L$ means that $j$ is either $i + k$ or $2(i + k)$. Since $i$, $j$, and $k$ are all positive, the only way both conditions can be satisfied is for $j$ to be equal to both $2i$ and $i + k$—in other words, for the string to be of the form $a^ib^{2i}\#a^i$. The set of all such strings is not a CFL, and so $L$ cannot be a DCFL.

8.21. Every string accepted by $M_1$ is of the form $x\#y$, where $x$ and $xy$ are in $L$. If the PDA is nondeterministic, however, the converse may not be true. If $x$ and $xy$ are in $L$, $M$ can make a sequence of moves on $x$ that leads to an accepting state, and it can make a sequence of moves on $xy$ that leads to an accepting state, but the second sequence may not cause $M$ to be in an accepting state after $x$.

8.22. An example is the language *pal*. If $L_1 = \{x\#y \mid x, y \in L\}$ were a CFL, then $L_1 \cap \{0\}^*\{1\}^*\{0\}^*\{\#\}\{1\}^*\{0\}^*$ would also be a CFL, by Theorem 8.4. As we observed in the solution to Exercise 8.20, however, it is not.

8.23. To simplify things slightly, let $D = \{i \mid a^i \in L\}$. Then it will be sufficient to show that $D$ is a finite union of (not necessarily infinite) arithmetic progressions of the form

$\{m + ip \mid i \geq 0\}$. (See Exercise 5.51.)

Let $n$ be the integer in the pumping lemma applied to $L$. Then for any $m \in D$ with $m \geq n$, there is an integer $p_m$ with $0 < p_m \leq n$ for which $m + ip_m \in D$ for every $i \geq 0$. There are only a finite number of distinct $p_m$'s; let $p$ be the least common multiple of all of them. Then for every $m \in D$ with $m \geq n$, $m + ip \in D$ for every $i \geq 0$.

Now, for every $k$ satisfying $0 \leq k < p$, if there is an $m \in D$ with $m \geq n$ and $m \equiv k$ mod $p$, then because of the preceding paragraph, the set of all such $m$ form an infinite arithmetic progression. Whether there is or not, the set $\{m \in D \mid m < n \text{ and } m \equiv k \text{ mod } p\}$ is a finite union of (finite) arithmetic progressions. Therefore, the set $\{m \in D \mid m \equiv k \text{ mod } p\}$ is a finite union of arithmetic progressions. Therefore, by taking the union over all $k$ with $0 \leq k < p$, the entire set $D$ has this property.

8.24. We show the "only if" part. Suppose $L$ is a CFL, and let $n$ be the integer in the pumping lemma. For any $m \geq n$, let $u = a^{f(m)}b^m$. Then there are strings $v$, $w$, $x$, $y$, and $z$, with $w$ and $y$ not both null and $|wy| \leq n$, so that $vw^ixy^iz \in L$ for every $i \geq 0$. Suppose $wy$ contains $p_m$ $b$'s and $q_m$ $a$'s. Then $vw^{i+1}xy^{i+1}z$ contains $iq_m + f(m)$ $a$'s and $ip_m + m$ $b$'s, and so, since this string is in $L$, $f(ip_m + m) = iq_m + f(m)$ for every $i \geq 0$. Since $p_m$ and $q_m$ can't both be 0, it follows that $p_m > 0$. Since each $p_m$ satisfies $p_m \leq n$, there are only a finite number of them; let $p$ be the least common multiple of these values. Then for each $m$, $p = k_m p_m$ for some $k_m$, so that $f(ip + m) = ik_m q_m + f(m)$.

Now let $r$ be any integer with $0 \leq r < p$. Pick an integer $\alpha$ so that $m = \alpha p + r \geq n$. If we let $j' = j - \alpha$, then for $j' \geq 0$, $f(jp + r) = f(j'p + m) = j'k_m q_m + f(m) = cj + d$, where $c = k_m q_m$ and $d = f(m) - \alpha k_m q_m$.