THEOREM 3.19

Let G be LR(k). Then for any chain set C in G, (G,C) is CFLR(k).

PROOF. Because it is LR(k), G must be reduced and $S \xrightarrow{+} S$ cannot occur in G. Also, by Theorem 2.2, G must be unambiguous. Let C be any chain set in G. Then (G,C) is cf-unambiguous by Theorem 3.4. These observations will be needed in the proof.

Now in order to prove that (G,C) is CFLR(k), we suppose that α and β are rorcsf's of (G,C) with cf-handles (p,m) and (q,n) respectively such that :

$$m/\beta \quad \epsilon \quad V_T^* \tag{1a}$$

and $$(m+k):\alpha \quad = \quad (m+k):\beta \tag{1b}$$

and proceed to show that (p,m) = (q,n).

If (p,m) is a cf-handle for α, then by Corollary 3.8 there exist $\gamma, \delta \ \epsilon \ V^*$ such that

$$S \xrightarrow[R]{*} \gamma \dashv\!(p,m)\!\xrightarrow[RCF]{} \delta \xrightarrow[c]{*} \alpha. \tag{2}$$

Owing to the definition of the relation $\xrightarrow[RCF]{}$ we must have $m/\delta \ \epsilon \ V_T^*$ and so we may write $\delta = \mu x$ and $\alpha = \theta x$ where $m = len(\mu) = len(\theta)$ and $\mu \xrightarrow[c]{*} \theta$. We may therefore rewrite (2) as :

$$S \xrightarrow[R]{*} \gamma \dashv\!(p,m)\!\xrightarrow[RCF]{} \mu x \xrightarrow[c]{*} \theta x. \tag{3}$$

Since G is reduced, there must exist $z \ \epsilon \ V_T^*$ such that $\theta \xrightarrow[R]{*} z$. Let the cf-distance from θ to z be d. Because $\mu \xrightarrow[c]{*} \theta$ and $\theta \xrightarrow[R]{*} z$ we must also have $\mu \xrightarrow[R]{*} z$ and the cf-distance from μ to z must be d also (this is because the cf-distance from μ to θ is zero). We may now extend (3) to give :

$$S \xrightarrow[R]{*} \gamma \dashv\!(p,m)\!\xrightarrow[RCF]{} \mu x \xrightarrow[R]{*} zx. \tag{4}$$

We also have that (q,n) is a cf-handle for β and so, again by Corollary 3.8, there exist $\rho, \pi \in V^*$ such that

$$S \xrightarrow[R]{*} \rho \xrightarrow[RCF]{(q,n)} \pi \xrightarrow[C]{*} \beta . \qquad (5)$$

From (1a) and (1b) it follows that β has the form $\beta = \theta y$ where $y \in V_T^*$ satisfies $k:y = k:x$ and since we already have $\theta \xrightarrow[R]{*} z$, it follows that $\pi \xrightarrow[R]{*} zy$. We can therefore extend (5) to give

$$S \xrightarrow[R]{*} \rho \xrightarrow[RCF]{(q,n)} \pi \xrightarrow[R]{*} zy. \qquad (6)$$

Note that the cf-distance from π to zy is again d. (This is because the cf-distance from π to θy is zero, and the cf-distance from θ to z is d.)

Now (4) gives $S \xrightarrow[R]{*} \mu x$

and (6) gives $S \xrightarrow[R]{*} zy$

and we know that $\mu \xrightarrow[R]{*} z$ and that $k:x = k:y$.

Therefore Lemma 3.17 gives $S \xrightarrow[R]{*} \mu y$ and so we obtain

$$S \xrightarrow[R]{*} \mu y \xrightarrow[R]{*} zy. \qquad (7)$$

From (4) we see that the handle of μx is (p,m). Recall that $m = \text{len}(\mu)$ and $k:x = k:y$. It follows that $m/\mu y \in V_T^*$ and that $(m+k):\mu x = (m+k):\mu y$. Therefore, since G is LR(k), the handle of μy must be (p,m) also. Thus for some $\sigma \in V^*$, (7) becomes

$$S \xrightarrow[R]{*} \sigma \xrightarrow[RCF]{(p,m)} \mu y \xrightarrow[R]{*} zy. \qquad (8)$$

But then both (6) and (8) are r-derivations of zy. Since G is unambiguous there can be only one r-derivation of zy and so (6) and (8) must simply be different ways of writing this unique r-derivation. Let $\langle (q_i, m_i) \rangle_{i=1}^{r}$ be the explicit r-derivation of zy from S and let $\langle \psi_i \rangle_{i=0}^{r}$ be the corresponding implicit derivation. From (6) and (8) it follows that there exist i, j in the range $1 \leq i,j \leq r$ such that

(from 6) : $(q_i, m_i) = (q,n)$, $\psi_{i-1} = \rho$, $\psi_i = \pi$ and (9)

(from 8) : $(q_j, m_j) = (p,m)$ $\psi_{j-1} = \sigma$, $\psi_j = \mu y$.

Remember that our goal is to show that (p,m)=(q,n). We can do this by showing that i=j. Clearly exactly one of the relations i=j, i < j and i > j must be true. Suppose first that i < j. Then we have

$$\psi_0 \overset{*}{\underset{R}{\rightarrow}} \psi_i \overset{*}{\underset{R}{\rightarrow}} \psi_{j-1} \dashv(q_j, m_j)\rightarrowtail_R \psi_j \overset{*}{\underset{R}{\rightarrow}} \psi_r .$$

Substituting S for ψ_0 , zy for ψ_r and using the identities from (9) gives :

$$S \overset{*}{\underset{R}{\rightarrow}} \pi \overset{*}{\underset{R}{\rightarrow}} \sigma \dashv(p,m)\rightarrowtail_R \mu y \overset{*}{\underset{R}{\rightarrow}} zy. \qquad (10)$$

Now we know that $p \in P \setminus C$ (since (p,m) is the cf-handle of α) and so it follows from (10) that the cf-distance from π to zy is at least one greater than the cf-distance from μy to zy. But this is not so, for both these cf-distances are known to be d. From this contradiction we conclude that $i \geqslant j$. By an exactly similar argument it may be shown that the supposition i > j is also untenable and so we deduce that i =j. Then (9) gives (p,m) = (q,n) and we may conclude the theorem. □

This result is of great practical significance, for it means that the speed benefits of cf-parsing can be obtained (anticipating for the moment that it is possible to construct cf-parsers for the CFLR(k) cs-grammars) without sacrificing the attractive generality of the LR(k) grammars.