1 The Intersection of a CFG and a REG is CFG

It is well known that the intersection of a context free language and a regular language is context free. This theorem is used in several proofs that certain languages are not context free. The usual proof of this theorem is a cross product construction of a PDA and a DFA. This requires the equivalence of PDA's and CFG's. Is there a proof that does not use the equivalence? That is, is there a proof that just uses CFG's? There is and we show it in this note.

This proof is due to Y. Bar-Hillel et al. [1].

Def 1.1 A context free grammar is in *Chomsky Normal Form* if every production is either of the form $X \to YZ$ or $X \to \sigma$ where $\sigma \in \Sigma$.

The following lemmas are well known.

Lemma 1.2 *If L is a context free language without e then there is grammar in Chomsky Normal Form that generates L.*

Lemma 1.3 *If $L \neq \emptyset$ and L is regular then L is the union of regular language $A_1, \ldots, A_n$ where each A_i is accepted by a DFA with exactly one final state.*

We now prove our main theorem.

Theorem 1.4 *If L_1 is a context free language and L_2 is a regular language then $L_1 \cap L_2$ is context free.*

Proof:

We do the case where $e \notin L_1$ and $L_2 \neq \emptyset$. All other cases we leave to the reader.

By Lemma 1.2 we can assume there exists a Chomsky normal form grammar $G = (N, \Sigma, S, P)$ for L_1. By Lemma 1.3 $L_2 = A_1 \cup \cdots \cup A_n$ where each A_i where each A_i is recognized by a DFA with exactly one final state. Note that

$$L_1 \cap L_2 = L_1 \cap (A_1 \cup \cdots \cup A_n) = \bigcup_{i=1}^{n} (L_1 \cap A_i).$$

Since CFL's are closed under union (and this can be proven using CFG's, so this is not a cheat) we need only show that the intersection of L_1 with a regular language recognized by a DFA with one final state is CFL. Let $M = (Q, \Sigma, \delta, s, \{f\})$ be a DFA with exactly one final state.

We construct the CFG $G' = (N', \Sigma, S', P')$ for $L_1 \cap L(M)$.

1. The nonterminals N' are triples $[p, V, r]$ where $V \in N$ and $p, r \in Q$.
2. For each production $A \to BC$ in P, for every $p, q, r \in Q$ we have the production

   $$[p, A, r] \to [p, B, q][q, C, r]$$

   in P'.

3. For every production $A \to \sigma$ in P, for every $(p, \sigma, q) \in Q \times \Sigma \times Q$ such that $\delta(p, \sigma) = q$ we have the production

$$[p, A, q] \to \sigma$$

   in P'

4. $S' = [s, S, f]$

We leave the easy proof that this works to the reader.

$\blacksquare$

References

[1] Y. Bar-Hiller, M. Perles, and E. Shamir. On formal properties of simple phrase structure grammars. *Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationforshung*, 14(2):143–172, 1961.