
Problem 2.39. Refer to Problem 1.42 for the definition of the shuffle operation. Show that the class of context-free languages is not closed under shuffle.

Informal description. For languages A and B, the **shuffle** of A and B is the language

$$\{w \mid w = a_1b_1 \cdots a_kb_k,\ where\ a_1 \cdots a_k \in A\ and\ b_1 \cdots b_k \in B,\ each\ a_i, b_i \in \Sigma^*\}.$$

Let $A_1 = \{x^ky^k \mid k \geq 0\}$, and $B_1 = \{0^k \mid k \geq 0\}$, and let S_1 be the shuffle of A_1 and B_1. Each string in S_1 contains equal number of x's and y's, and any number of 0's.

$S_1 = \{w \mid w = a_1b_1 \cdots a_kb_k,\ where\ a_1 \cdots a_k \in A_1\ and\ b_1 \cdots b_k \in B_1,\ each\ a_i \in \{x, y\}^*,\ b_i \in \{0\}^*\}.$

Now, let $A_2 = \{x^ky^k \mid k \geq 0\}$, and $B_2 = \{0^k1^k \mid k \geq 0\}$. Let S_2 be the shuffle of A_2 and B_2. Every string in S_2 contains equal number of x's and y's, and equal number of 0's and 1's.

$S_2 = \{w \mid w = a_1b_1 \cdots a_kb_k,\ where\ a_1 \cdots a_k \in A_2\ and\ b_1 \cdots b_k \in B_2,\ each\ a_i \in \{x, y\}^*,\ b_i \in \{0, 1\}^*\}.$

The languages A_2 and B_2 are CFLs, but S_2 is not a CFL.

Proof. Let $A = \{x^ky^k \mid k \geq 0\}$, and $B = \{0^k1^k \mid k \geq 0\}$. Languages A and B are easily seen to be CFLs, but their shuffle S is not a CFL.

$S = \{w \mid w = a_1b_1 \cdots a_kb_k,\ where\ a_1 \cdots a_k \in A\ and\ b_1 \cdots b_k \in B,\ each\ a_i \in \{x, y\}^*,\ b_i \in \{0, 1\}^*\}.$

The proof is by contradiction. Assume the language S is context free. Let p be the pumping length given by the pumping lemma. Choose s to be the string $x^p0^py^p1^p$. Clearly s is a member of S and of length at least p. The pumping lemma states that s can be pumped, but we show that it cannot be pumped. $\square$