

Use the Pumping Lemma to show that  $\{xx \mid x \in \{a, b\}^*\}$  is not regular.

Answer: Let  $N$  be the threshold length and pick  $xx = a^N b^N a^N b^N$ . This belongs to our language ( $x = a^N b^N$ ) and  $|xx| \geq N$  as required.

Assume that  $L = \{xx \mid x \in \{a, b\}^*\}$  is regular; then Theorem 3.6 says that  $a^N b^N a^N b^N = xyz$  where  $|xy| \leq N$  and  $|y| \geq 1$ . We observe that  $xy$  must consist entirely of  $as$  since  $a^N b^N a^N b^N$  starts with  $N$   $as$ . Let  $1 \leq k \leq N$  be the number of  $as$  in  $y$ , then we can write  $xyz$  as  $a^{N-k} a^k b^N a^N b^N$  where  $y = a^k$ . But then  $xz = a^{N-k} b^N a^N b^N$ , which does not have the form  $xx$ . So the assumption that  $L$  was regular must be false.

Discussion: Use the same strategy used for  $\{a^n b^n \mid n \geq 0\}$ : let  $N$  be the threshold length, then pick a string of the desired form that is long enough to ensure that  $xy$  can only cover one of the symbols. Then  $xz$  omits some of those symbols, disrupting the balance between the two halves of the string.

---

Use the Pumping Lemma to show that  $\{x \mid n_a(x) = n_b(x)\}$  is not regular.

Answer: Let  $N$  be the threshold length and pick  $x = a^N b^N$ . This belongs to our language ( $n_a(x) = n_b(x) = N$ ) and  $|x| \geq N$  as required.

Assume that  $L = \{x \mid n_a(x) = n_b(x)\}$  is regular; then Theorem 3.6 says that  $a^N b^N = xyz$  where  $|xy| \leq N$  and  $|y| \geq 1$ . We observe that  $xy$  must consist entirely of  $as$  since  $a^N b^N$  starts with  $N$   $as$ . Let  $1 \leq k \leq N$  be the number of  $as$  in  $y$ , then we can write  $xyz$  as  $a^{N-k} a^k b^N$  where  $y = a^k$ . But then  $xz = a^{N-k} b^N$ , where the number of  $as$  and  $bs$  are not the same. So the assumption that  $L$  was regular must be false.

Discussion: Use the same strategy used for  $\{a^n b^n \mid n \geq 0\}$ : let  $N$  be the threshold length, then pick a string of the desired form that is long enough to ensure that  $xy$  can only cover one of the symbols. Then  $xz$  omits some of those symbols, disrupting the balance between the two halves of the string.