A resposta é sim para ambas as perguntas. Na realidade, existem várias caracterizações adicionais que podem servir de auxílio em uma ou ambas as situações. Nesta seção, três tipos de resultados serão apresentados. Os dois primeiros explicitam características importantes de toda linguagem regular, que são úteis, por exemplo, para mostrar que uma linguagem não é regular. Em seguida, serão indicadas algumas propriedades de fechamento[6] da classe das linguagens regulares. Essas propriedades são úteis para as duas aplicações referidas anteriormente. Outras caracterizações serão exibidas na Seção 2.6.

2.4.1 Invariância à direita

O seguinte teorema apresenta uma condição necessária para uma linguagem ser regular, ou seja, uma propriedade que toda linguagem regular tem.

Teorema 7[7] *Se L é uma linguagem regular de alfabeto Σ, então para qualquer $R \subseteq \Sigma^*$ infinito existem $x, y \in R$ tais que $x \neq y$ e para toda $z \in \Sigma^*$ $xz \in L \leftrightarrow yz \in L$.*

Prova

Suponha que L seja regular. Seja, então, um AFD $(E, \Sigma, \delta, i, F)$ que reconheça L. Seja $R \subseteq \Sigma^*$ um conjunto infinito arbitrário. Como existem mais de $|E|$ palavras em R (pois R é infinito), pelo princípio da casa de pombos existem palavras $x, y \in R$ tais que $x \neq y$ e $\hat{\delta}(i, x) = \hat{\delta}(i, y)$. Para mostrar que para toda $z \in \Sigma^*$ $(xz \in L \leftrightarrow yz \in L)$, basta mostrar que para toda $z \in \Sigma^*$ $\hat{\delta}(i, xz) = \hat{\delta}(i, yz)$. De fato:

$$\hat{\delta}(i, xz) = \hat{\delta}(\hat{\delta}(i, x), z) = \hat{\delta}(\hat{\delta}(i, y), z) = \hat{\delta}(i, yz).$$

qualquer que seja $z \in \Sigma^*$. ■

A contrapositiva do enunciado do Teorema 7 pode ser usada para provar que uma linguagem L, de alfabeto Σ, não é regular: basta encontrar um conjunto infinito R tal que para cada par $x, y \in R$, $x \neq y$, existe $z \in \Sigma^*$ tal que $xz \in L$ e $yz \notin L$, ou vice-versa, $yz \in L$ e $xz \notin L$.

Exemplo 41 A linguagem $L = \{\mathtt{a}^n\mathtt{b}^n \mid n \in \mathbf{N}\}$ não é regular, como mostrado a seguir usando o Teorema 7.

Seja $R = \{\mathtt{a}\}^*$. Sejam duas palavras arbitrárias $\mathtt{a}^i, \mathtt{a}^j \in R$, $i \neq j$. Como $\mathtt{a}^i\mathtt{b}^i \in L$ e $\mathtt{a}^j\mathtt{b}^i \notin L$, segue-se, pelo Teorema 7, que L não é regular. ■

Note que, nesse tipo de demonstração, o problema é (1) encontrar um conjunto infinito de palavras R tal que (2) para quaisquer duas palavras diferentes x e y de R exista uma palavra z (possivelmente dependente de x ou y) tal que ou $xz \in L$ ou $yz \in L$, mas não ambos. Segue mais um exemplo.

[6] Em inglês, *closure properties*.

[7] Mais formalmente:
L é regular $\rightarrow \forall R \subseteq \Sigma^* R$ é infinito $\rightarrow \exists x, y \in R\,[x \neq y \wedge \forall z \in \Sigma^*\,(xz \in L \leftrightarrow yz \in L)]$.

Exemplo 42 Seja $L = \{0^m 1^n \mid m > n\}$. Tome $R = \{0\}^*$ e sejam $0^i, 0^j \in R$, $i < j$, palavras arbitrárias de R. Tem-se que $0^i 1^i \notin L$ e $0^j 1^i \in L$. Logo, pelo Teorema 7, L não é regular. □

Exemplo 43 Seja $L = \{w \in \{0,1\}^* \mid w = w^R\}$. Tome $R = \{0\}^*\{1\}$ e sejam $0^i 1, 0^j 1 \in R$, $i < j$, palavras arbitrárias de R. Tem-se que $0^i 1 0^i \in L$ e $0^j 1 0^i \notin L$. Assim, pelo Teorema 7, L não é regular. □

Exemplo 44 Seja a linguagem $L = \{0^{n^2} \mid n \in \mathbb{N}\}$. Seja o conjunto $R = L$. Sejam $0^{i^2}, 0^{j^2} \in R$, $j > i$, quaisquer. Segue-se que $0^{i^2} 0^{2i+1} = 0^{(i+1)^2} \in L$, mas $0^{j^2} 0^{2i+1} \notin L$, pois $j^2 < j^2 + 2i + 1 < (j+1)^2$). Portanto, pelo Teorema 7, L não é regular. □

O problema de encontrar um conjunto R adequado nem sempre é lá tão trivial.

Exemplo 45 Seja a linguagem $L = \{0^n \mid \ n \text{ é primo}\}$. Seja o conjunto $R = \{0^{p_0}, 0^{p_1}, 0^{p_2}, \ldots\}$ definido assim:

- $p_0 = 2$;
- seja k o menor número primo maior do que p_{n-1}; para $n \geq 1$, p_n é o menor número primo maior ou igual a k tal que entre p_n e o menor número primo maior do que p_n existam pelo menos $k - p_{n-1}$ números consecutivos não primos.[8]

Sejam $0^{p_i}, 0^{p_j} \in R$, $j > i$, quaisquer. Seja k o menor número primo maior do que p_i. Como $p_j > p_i$ existem no mínimo $k - p_i$ números consecutivos não primos entre o número primo p_j e o próximo número primo maior do que p_j. Logo, $p_i + (k - p_i) = k$ é primo, mas $p_j + (k - p_i)$ não é. Portanto, $0^{p_i} 0^{k-p_i} \in L$ e $0^{p_j} 0^{k-p_i} \notin L$. Pelo Teorema 7, L não é regular. □

2.4.2 O lema do bombeamento

A visão de um AFD como um grafo facilita o raciocínio para a obtenção de diversas propriedades da classe das linguagens reconhecidas por esse tipo de máquina. Uma das principais é dada pelo chamado lema do bombeamento, que especifica uma propriedade que qualquer linguagem regular possui. Além dessas, outras linguagens (não reconhecidas por AFDs) também têm essa mesma propriedade. Assim, ao mostrar que uma linguagem L satisfaz o lema do bombeamento, isso não implica que L seja regular. Mas ao demonstrar que L *não* satisfaz o lema do bombeamento, pode-se concluir que L *não* é linguagem regular.

Segue o lema do bombeamento (LB), uma extensão do Teorema 3.

Lema 4[9] *Seja L uma linguagem regular. Então existe uma constante $k > 0$ tal que para qualquer palavra $z \in L$ com $|z| \geq k$ existem u, v e w que satisfazem as seguintes condições:*

---

[8] Veja o Exercício 9 da Seção A.1.

[9] Mais formalmente:
L é regular $\rightarrow \exists k > 0 \, \forall z \in L \, [|z| \geq k \rightarrow \exists u, v, w \, (z = uvw \wedge |uv| \leq k \wedge |v| \geq 1 \wedge \, \forall i \in \mathbf{N} \, uv^i w \in L)]$.