TEORÍA DE LA COMPUTACIÓN

Grado en Ingeniería Informática

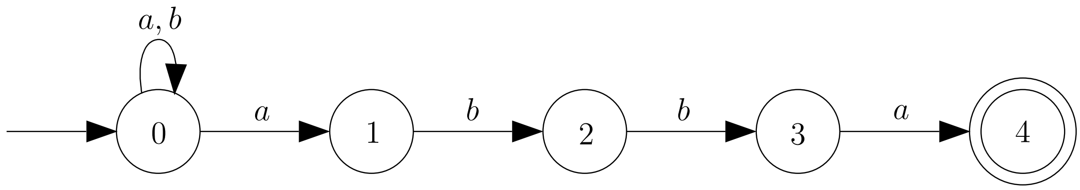
Soluciones del Boletín de Ejercicios nº 2

Autómatas finitos deterministas y no deterministas

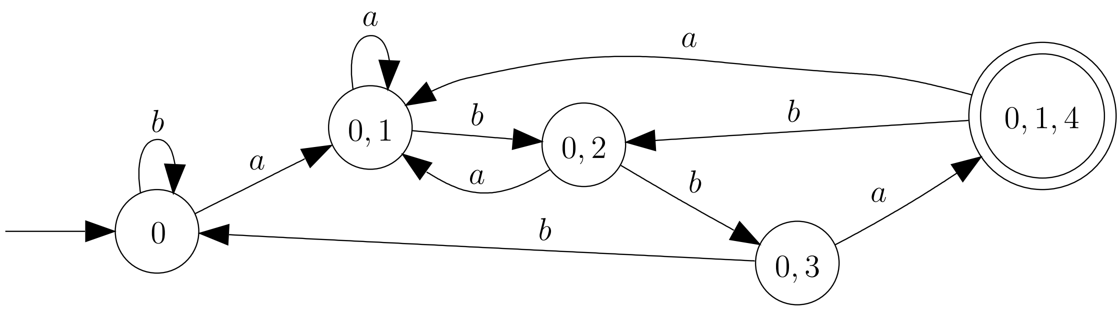
24. Encuentre un AFN que acepte el lenguaje de todas las cadenas sobre $\{a, b\}$ que terminan con la subcadena *abba*. Conviértalo en AFD.

    Solución:

    El AFN podría ser el siguiente:



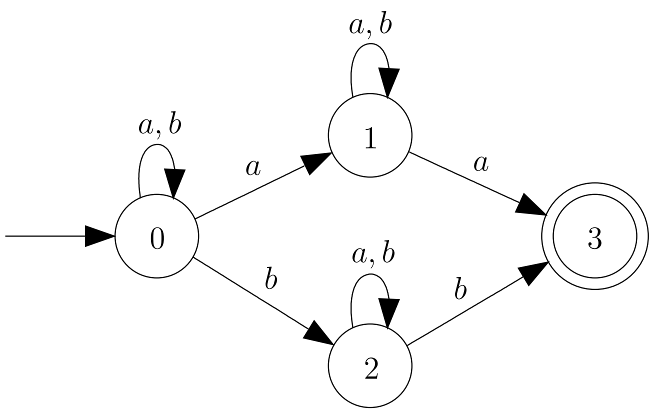
Y su correspondiente AFD es:



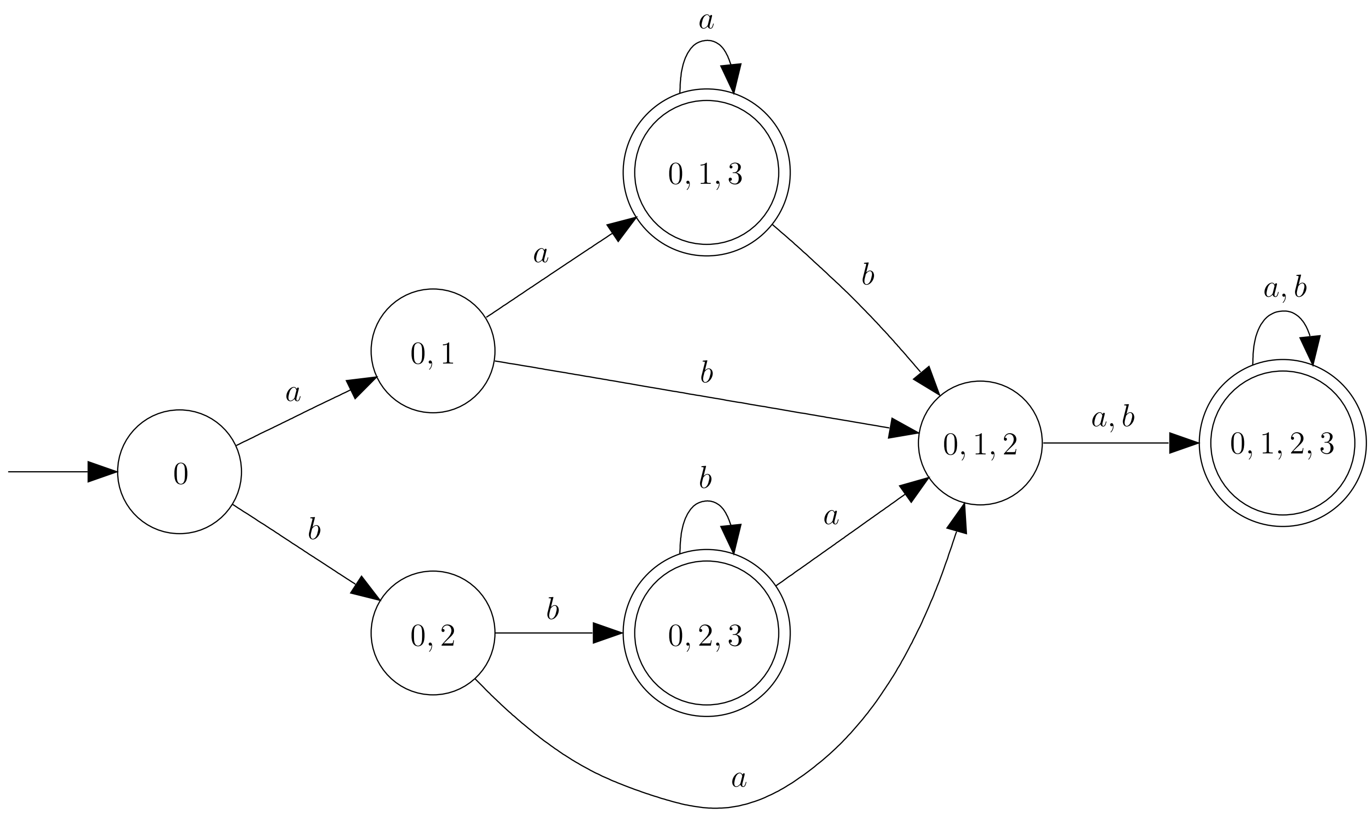
25. Sobre el alfabeto $\{a, b\}$, construya un AFN que acepte cadenas cuyo último símbolo haya aparecido antes en la misma entrada. Conviértalo en AFD.

    Solución:

    El AFN podría ser el siguiente:



Y su correspondiente AFD es:

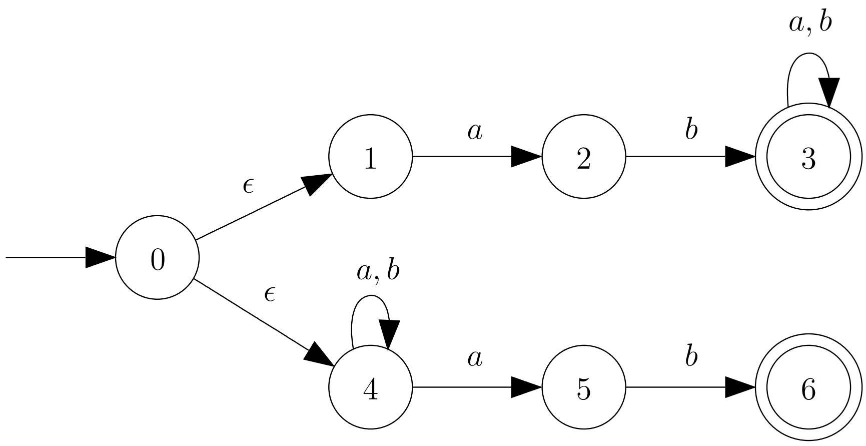


Autómatas finitos no deterministas con ϵ-transiciones

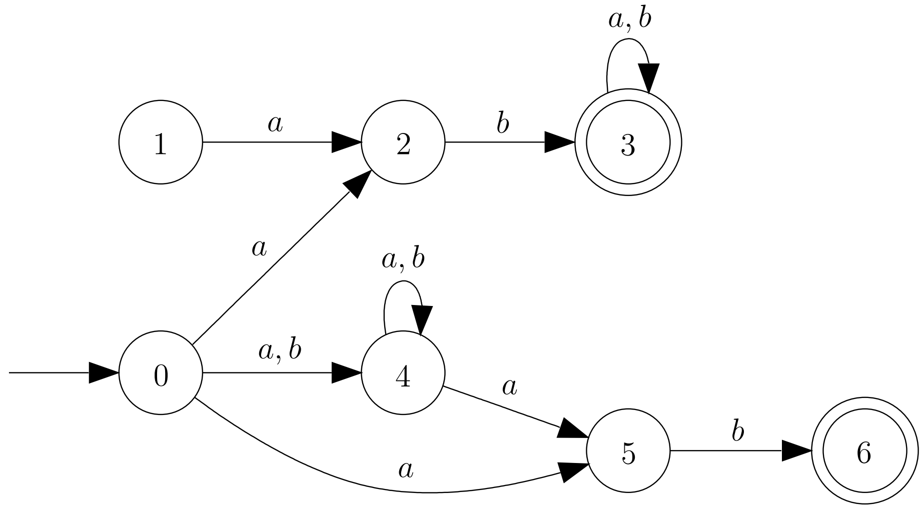
26. Sobre el alfabeto $\{a, b\}$, construya un AFN-ϵ que acepte todas las cadenas que empiezan o terminan con ab, o ambas cosas. Conviértalo primero en AFN y después en AFD.

    Solución:

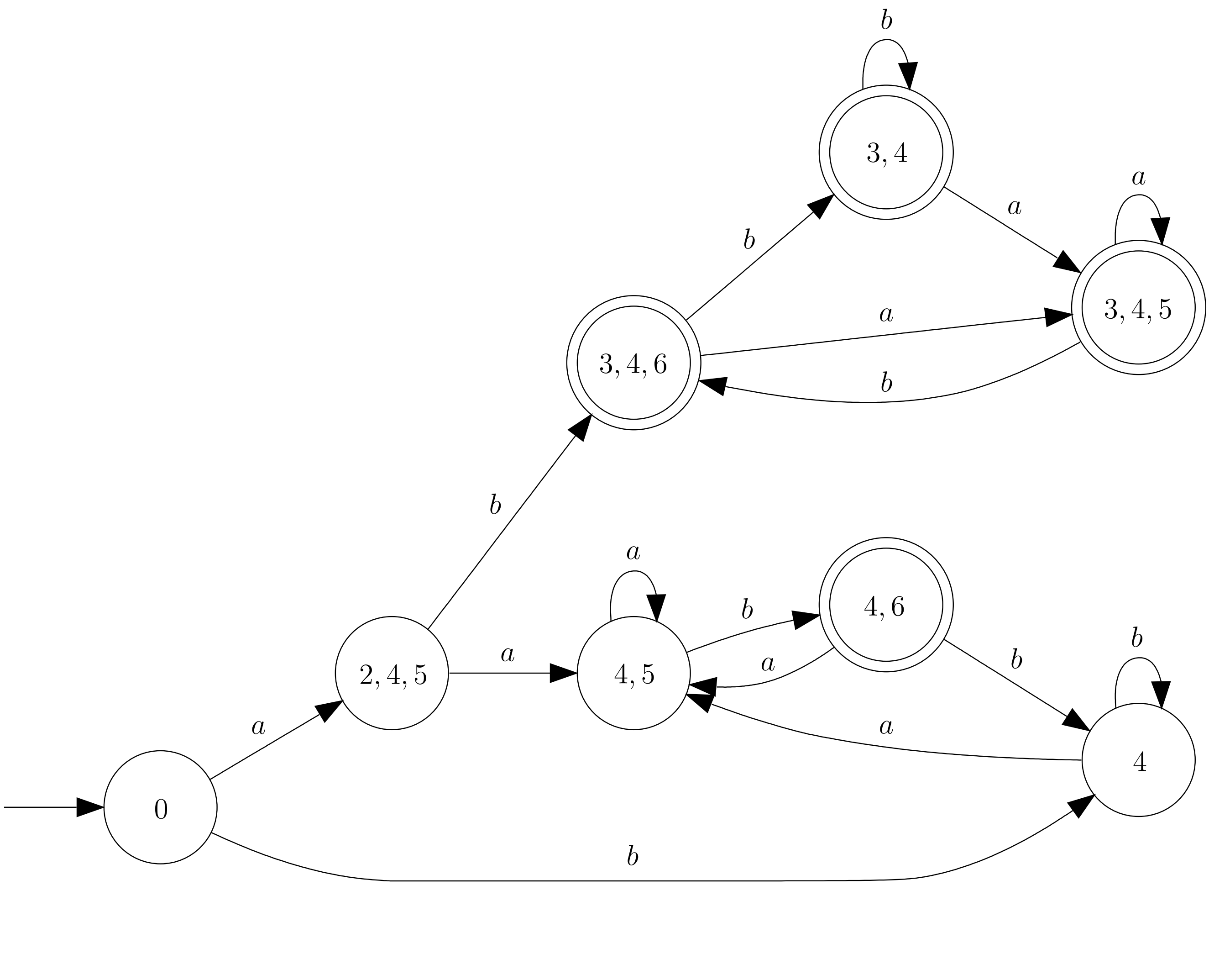
    El AFN-ϵ podría ser el siguiente:



Su correspondiente AFN es:



Y su correspondiente AFD es:



27. Para todo autómata finito M (con uno o más estados de aceptación), ¿existe un autómata finito M' con un solo estado de aceptación tal que $L(M) = L(M')$?

    a) Sí, para todo M.

    b) Si y sólo si los estados de aceptación de M no son origen de ninguna transición.

    c) Depende del alfabeto considerado.

Solución:

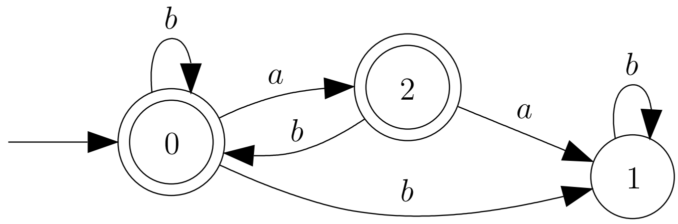
a) Cierta. Se añade un nuevo estado que será el único estado final, y los antiguos estados finales se unen a él mediante arcos ϵ. Eso sí, obtendremos un AFN-ϵ y si lo convertimos en AFN volverán a aparecer varios estados finales.

b) También puede ser cierta. Si los estados finales no tienen arcos salientes, se pueden colapsar en uno solo.

c) No es cierta.

Autómatas finitos y expresiones regulares

28. Obtenga la expresión regular del lenguaje aceptado por el siguiente autómata finito:



Solución:

$$
\begin{aligned}
A_1 &= b\,A_1 = b^*\,\emptyset = \emptyset \\
A_2 &= a\,A_1 \cup b\,A_0 \ \cup \epsilon = a\,\emptyset \cup b\,A_0 \cup \epsilon = \emptyset \cup b\,A_0 \cup \epsilon = b\,A_0 \cup \epsilon \\
A_0 &= a\,A_2 \cup b\,A_0 \cup \epsilon = a\,(b\,A_0 \cup \epsilon) \cup b\,A_0 \cup \epsilon = a\,b\,A_0 \cup a \cup b\,A_0 \cup \epsilon \\
&= (b \cup ab)\,A_0 \cup a \cup \epsilon = \underline{(b \cup ab)^*\,(a \cup \epsilon)}
\end{aligned}
$$

Propiedades de bombeo de los lenguajes regulares

29. El lema del bombeo para lenguajes regulares puede utilizarse para demostrar que un determinado lenguaje L no es regular. Uno de los pasos de la aplicación de dicho lema consiste en suponer que el lenguaje L es regular, y que por tanto existe una cierta constante k asociada a L. ¿Qué representa realmente esa constante k?

a) Es el número de símbolos del alfabeto sobre el que está definido L.

b) Es el número de estados del AFD que aceptaría L, en el caso de que L fuese regular.

c) Es la longitud máxima de cualquiera de las cadenas de L.

Solución:

La respuesta correcta es la b).

30. El siguiente razonamiento intenta demostrar mediante el lema del bombeo que el lenguaje $L = 0^*1^*$ no es regular:

Supongamos que L es regular. Entonces existe una constante k asociada a L tal que cualquier cadena $z \in L$, $|z| \geq k$, puede descomponerse en $z = uvx$, donde $|v| \geq 1$ y $|uv| \leq k$, y de forma que $uv^ix \in L$, $\forall i \geq 0$. Consideremos la cadena $z = 0^k1^k$ y todas sus posibles descomposiciones:

- Si v está formada sólo por ceros, cualquier bombeo $i \geq 2$ produce más ceros que unos.
- Si v está formada sólo por unos, cualquier bombeo $i \geq 2$ produce más unos que ceros.
- Si v está formada por ceros y por unos, cualquier bombeo $i \geq 2$ mezcla los ceros y los unos produciendo cadenas que tampoco pertenecen a L.

De todo esto se concluye que L no es regular.

Indique de forma clara si este razonamiento es correcto o no.

Solución:

Este razonamiento no es correcto por las siguientes razones:

- Si v está formada sólo por ceros, cualquier bombeo produce cadenas pertenecientes al lenguaje L.
- El caso de que v esté formada sólo por unos no es posible.
- El caso de que v esté formada sólo por ceros y unos tampoco es posible.

31. Demuestre mediante el lema del bombeo que el lenguaje $L = \{wcw^I \mid w \in \{a, b\}^*\}$ no es regular.

Solución:

Suponemos que L es regular y sea k su constante asociada según el lema del bombeo. Elegimos la cadena $w = a^kb^kc\,b^ka^k$. Dado que $w \in L$ y $|w| = 4k + 1 \geq k$, w está en las condiciones del lema.

Por tanto, w debería poder descomponerse de la forma $w = uvx$, donde $|v| \geq 1$ y $|uv| \leq k$, y cualquier bombeo uv^ix debería estar en L, $\forall i \geq 0$.

Si $|uv| \leq k$, entonces las subcadenas u y v están formadas sólo por $a$es, es decir:

$$u = a^r \qquad v = a^s\ (s \geq 1) \qquad x = a^{k-(r+s)}b^kc\,b^ka^k$$

Y si elegimos, por ejemplo, el bombeo $i = 2$, obtenemos lo siguiente:

$$uv^2x = a^ra^{2s}a^{k-(r+s)}b^kc\,b^ka^k = a^{k+s}b^kc\,b^ka^k$$

Pero, dado que $s \geq 1$, la cadena $a^{k+s}b^kc\,b^ka^k$ no tiene el mismo número de $a$es al principio que al final. Es decir, el bombeo $i = 2$ produce una cadena que no está en L. Por lo tanto, L no es regular.

32. Demuestre mediante el lema del bombeo que el lenguaje $L = \{a^nb^m \,|\, n \geq m\}$ no es regular.

Solución:

Suponemos que lo es y tomamos $w = a^kb^k$, donde k es la constante asociada a L. Dado que $w \in L$ y $|w| = 2k \geq k$, w está en las condiciones del lema.

Por tanto, w debería poder descomponerse de la forma $w = uvx$, donde $|v| \geq 1$ y $|uv| \leq k$, y cualquier bombeo uv^ix debería estar en L, $\forall i \geq 0$.

Si $|uv| \leq k$, entonces las subcadenas u y v están formadas sólo por $a$es, es decir:

$$u = a^r \qquad v = a^s\ (s \geq 1) \qquad x = a^{k-(r+s)}b^k$$

En este caso, cualquier bombeo $i \geq 2$ producirá cadenas con más $a$es que $b$es. Pero todas esas cadenas estarían dentro del lenguaje L.

Así pues, esta vez el único bombeo que va a producir cadenas fuera de L va a ser el bombeo $i = 0$, tal y como vemos a continuación:

$$uv^0x = a^r a^{k-(r+s)}b^k = a^{k-s}b^k$$

Dado que $s \geq 1$, la cadena $a^{k-s}b^k$ tiene menos $a$es que $b$es. Es decir, el bombeo $i = 0$ produce una cadena no perteneciente a L. Por lo tanto, L no es regular.

Otras propiedades de cierre de los lenguajes regulares

33. Sea L un lenguaje regular sobre el alfabeto $\Sigma = \{a, b\}$, y sea M un autómata finito determinista tal que $L(M) = L$. Indique cómo, a partir de M, se puede construir otro autómata finito M' tal que $L(M') = \overline{L}$.

    Solución:

    $F' = Q - F$. Es decir, los estados finales de M' son los estados que antes no eran finales en M. Y viceversa, es decir, los estados no finales de M' son los estados que antes eran finales en M. De esta forma, M' acepta las cadenas que M rechaza, y rechaza las que M acepta. Por tanto, $L(M') = \overline{L}$.

34. Demuestre que si L es un lenguaje regular, su inverso L^I también lo es. Sugerencia: considere un autómata finito M tal que $L(M) = L$, y explique detalladamente cómo a partir de M se puede construir otro autómata finito M' tal que $L(M') = L^I$.

    Solución:

    A partir de un autómata finito M tal que $L(M) = L$, se construye un nuevo autómata finito M' como sigue:

    - Se cambia el sentido de todos los arcos.
    - Los estados finales dejan de serlo, y el estado inicial pasa a ser el único final.
    - Se crea un nuevo estado inicial y se conecta mediante arcos ϵ con los antiguos estados finales.

    De esta forma, $L(M') = L^I$ y por tanto L^I es un lenguaje regular.

35. Encuentre un algoritmo que determine si dos lenguajes regulares L_1 y L_2 tienen por lo menos una cadena en común. Sugerencia: indique cómo se podría demostrar que $L = L_1 \cap L_2$ es regular, y cómo se podría comprobar si $L = \emptyset$ o no.

Solución:

Dados dos autómatas finitos M_1 y M_2, tales que $L(M_1) = L_1$ y $L(M_2) = L_2$, construimos el "autómata producto cartesianso" $M_1 \times M_2$, que aceptará las cadenas que cumplen las propiedades de L_1 y de L_2 al mismo tiempo. Es decir, $L(M_1 \times M_2) = L_1 \cap L_2$. Si en dicho autómata existe algún estado final conectado con el inicial, entonces $L_1 \cap L_2 \neq \emptyset$. Es decir, L_1 y L_2 tendrán alguna cadena en común.

36. Sean L_1 y L_2 dos lenguajes regulares. Mediante el uso de operaciones sobre lenguajes (unión, intersección, complementario, etc.), demuestre que $L_1 - L_2$ es regular.

    Solución:

    $L_1 - L_2 = L_1 \cap \overline{L_2}$

37. Sea L un lenguaje regular. Demuestre que el conjunto $P = \{u|uv \in L\}$ de los prefijos de L es también un lenguaje regular.

    Solución:

    A partir de un autómata finito M tal que $L(M) = L$, se construye un nuevo autómata finito M' poniendo todos los estados como estados finales. De esta forma, $L(M') = P$ y por tanto P es un lenguaje regular.

38. Sea L un lenguaje regular. Demuestre que el conjunto $S = \{v|uv \in L\}$ de los sufijos de L es también un lenguaje regular.

    Solución:

    A partir de un autómata finito M tal que $L(M) = L$, se construye un nuevo autómata finito M' añadiendo arcos ϵ desde el estado inicial a todos los demás estados. De esta forma, $L(M') = S$ y por tanto S es un lenguaje regular.

39. Sea L un lenguaje regular. Demuestre que la colección de cadenas cuyas inversas pertenecen a L es también un lenguaje regular.

    Solución:

    Esa colección de cadenas es un lenguaje que puede denotarse como $L \cap L^I$ y por tanto es un lenguaje regular.

    Por ejemplo:

    - Supongamos que $L = \{\underline{a}, \underline{aba}, \underline{bca}, \underline{acb}, ab, accb, ca, \ldots\}$.
    - Entonces $L^I = \{\underline{a}, \underline{aba}, \underline{acb}, \underline{bca}, ba, bcca, ac, \ldots\}$.
    - Y efectivamente las cadenas que nos interesan para nuestro lenguaje (las que aparecen subrayadas) son las que están en L y en L^I a la vez. Y como los lenguajes regulares son cerrados para las operaciones de intersección e inverso, $L \cap L^I$ es también un lenguaje regular.

Aplicaciones prácticas de expresiones regulares y AF,s

40. Sobre el alfabeto $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, -, ., e, E\}$, escriba una expresión regular que denote el lenguaje de todos los números reales escritos en notación científica.

    Solución:

    A continuación mostramos el formato de los números reales en notación científica, donde, como suele ser habitual, `<` y `>` indican componentes obligatorias, `[` y `]` indican componentes opcionales, y `|` indica componentes alternativas:

    ```
    [signo] <parte-entera> [<punto-decimal> [decimales]] [<E|e> [signo] <exponente>]
    ```

    Si denotamos entonces con $c = 0 \cup 1 \cup 2 \cup \ldots \cup 9$ las cifras, y con $p = .$ el punto decimal, la expresión regular que se pide sería:

    $$(+ \cup - \cup \epsilon) \cdot c^+ \cdot (p \cdot c^* \;\cup\; \epsilon) \cdot ((E \cup e) \cdot (+ \cup - \cup \epsilon) \cdot c^+ \;\cup\; \epsilon)$$