Hoja 4 de ejercicios. Autómatas y Computabilidad.
Facultad de Matemáticas. UCM

1. Diseña una gramática incontextual para cada uno de los siguientes lenguajes:
   - $L_1 = PAL_{par}$
   - $L_2 = PAL_{impar}$
   - $L_3 = PAL$
   - $L_4 = \{(ab)^n(ab)^n | n \geq 0\}$
   - $L_5 = \{(ab)^n b(ab)^n | n \geq 0\}$
   - $L_6 = \{(ab)^n(ba)^n | n \geq 0\}$
   - $L_7 = \{a^n ba^m | n \geq m\}$
   - $L_8 = L_4\ L_3$
   - $L_9 = \{a^n b^{n+m} c^m | n, m \geq 0\}$

2. ¿Puedes diseñar alguna gramática regular para alguno de los lenguajes del apartado anterior? Hazlo para todos los que sea posible. ¿Por qué no es posible para el resto de lenguajes?

3. Dada la gramática no ambigua de las expresiones aritméticas vista en clase, construye el árbol de derivación para $a + (b0 - a1 * a) - bb11$. Repite el ejercicio para la cadena $(10 * b/1) + (b - a - a - b)$

4. Dadas las siguientes reglas de una gramática incontextual para definir expresiones aritméticas

$$E \to E + E | E - E | E * E | E/E | (E) | I | N$$

   donde N e I representan naturales e identificadores:
   - Completa la gramática definiendo las reglas adecuadas para que N recoja todos los posibles naturales construidos con dígitos de 0 a 9, así como para que I recoja todos los posibles identificadores válidos. Un identificador será válido si sólo está compuesto por dígitos y letras, y además empieza por una letra (a, b) y los dos últimos caracteres también son letras. Fíjate en que la longitud debe ser al menos de dos caracteres.
   - Transforma la gramática de modo que se elimine el ambigüedad, y de forma que el $*$ tenga prioridad sobre $+$ y $-$, y que estos dos operadores tengan prioridad sobre $/$. Todos los operadores deberán asociar a la derecha, es decir, $2 - 3 - 7$ se interpretará como $2 - (3 - 7)$.

5. Dadas las siguientes reglas de una gramática incontextual para definir expresiones lógicas

$$E \to E \oplus E \mid E \vee E \mid E \wedge E \mid (E) \mid I \mid \texttt{true} \mid \texttt{false}$$

   donde I representa identificadores:
   - Completa la gramática definiendo las reglas adecuadas para que I recoja todos los posibles identificadores válidos. Un identificador será válido si sólo está compuesto por dígitos y letras, empieza por una letra (a, b), y tiene al menos tantos dígitos como letras.
   - Transforma la gramática de modo que se elimine el no determinismo, y de forma que el $\wedge$ tenga prioridad sobre $\vee$ y $\oplus$. Todos los operadores deberán asociar a la derecha, es decir, $a \wedge b \wedge a$ se interpretará como $a \wedge (b \wedge a)$.

6. Convierte a FNC la gramática de las expresiones aritméticas.

7. Convierte a FNC la siguiente gramática:

   $I \rightarrow aAa|bBb|\epsilon$
   $A \rightarrow C|a$
   $B \rightarrow C|b$
   $C \rightarrow CDB|\epsilon$
   $D \rightarrow A|B|ab$

8. ¿Son ciertas las siguientes afirmaciones? Justifica tu respuesta.

   - Después de eliminar producciones ϵ pueden aparecer nuevas producciones unitarias.
   - Después de eliminar producciones ϵ pueden aparecer nuevos símbolos inútiles.
   - Después de eliminar producciones unitarias pueden aparecer nuevos símbolos inútiles.
   - Después de eliminar producciones unitarias pueden aparecer nuevas producciones ϵ.
   - Después de eliminar símbolos inútiles pueden aparecer nuevas producciones unitarias.
   - Después de eliminar símbolos inútiles pueden aparecer nuevas producciones ϵ.
   - Después de eliminar símbolos inalcanzables pueden aparecer nuevos símbolos no generadores.
   - Después de eliminar símbolos no generadores pueden aparecer nuevos símbolos inalcanzables.

9. Define un autómata con pila para cada uno de los lenguajes del ejercicio 1.

10. ¿Has conseguido que alguno de los autómatas del ejercicio anterior se comporte de forma completamente determinista? Si no es así, inténtalo con el lenguaje L_7.

11. Para cada uno de los lenguajes siguientes, indica si son incontextuales o no. En caso de que lo sean, proporciona una GI o un AP, y si no lo son demuéstralo mediante el lema del bombeo. En caso de que alguno de los lenguajes sea regular, escribe su expresión regular.

    - $L_1 = \{a^n b^m a^n b^m | n > 0, m > 1\}$
    - $L_2 = \{a^n b^n a^m b^m | n > 0, m > 1\}$
    - $L_3 = \{a^n b^n a^m | n > 0, m \leq n\}$
    - $L_4 = \{a^i b^j a^k | i > j > k\}$
    - $L_5 = \{a^i b^j a^k | (i \geq j) or (i \geq k)\}$
    - $L_6 = \{a^i b^j a^k | (i \neq j) or (i \neq k)\}$
    - $L_7 = \{(ab)^n (ba)^n (bba)^n | n > 0\}$
    - $L_8 = \{w a^n w | w \in \Sigma^*, n > 0\}$
    - $L_9 = \{wvw | w, v \in \Sigma^*\}$

12. ¿Son ciertas las siguientes afirmaciones? Justifica tu respuesta.

    a) Un LI puede tener una GI no ambigua, y no tener ningún APD.
    b) Un LI puede tener una GI ambigua, y tener un APD.
    c) Un LR puede tener una GI ambigua.
    d) Un LR puede tener una GR ambigua.
    e) Un LR puede no tener ninguna GR no ambigua.

f) Todo LR también es un LI.

g) Todo LI tiene alguna GI ambigua.

h) Todo LI tiene alguna GR ambigua.

i) Todo LI tiene alguna GR no ambigua.

j) La intersección de dos LI nunca es un LI.

k) La intersección de dos LI siempre es un LI.

l) La diferencia entre un LR y un LI siempre es un LI.

m) La diferencia entre un LR y un LI siempre es un LR.

n) La diferencia entre un LR y un LI puede ser un LR.

ñ) La diferencia entre un LI y un LR siempre es un LI.

o) La diferencia entre un LI y un LR siempre es un LR.

p) La diferencia entre un LI y un LR puede ser un LR.

13. Define una GI que determine si una frase escrita en español es correcta o no. Puedes suponer que todas las frases son afirmativas, que no hay concordancia en sujeto y predicado ni entre sustantivos y adjetivos (e.g. "Trump lanzan misiles humanitario" sería correcto), y que el orden siempre es sujeto seguido de verbo seguido de complemento directo (suponemos que todos los verbos son transitivos) y posiblemente seguido de un complemento circunstancial. Supón que un grupo nominal puede estar compuesto por un único sustantivo, o por un sustantivo seguido de varios adjetivos. Fíjate en que no nos preocupamos de la veracidad de las frases, sino sólo de la corrección sintáctica.

14. Modifica la gramática anterior para introducir concordancia en número entre el sujeto y el predicado. Es decir, la frase "Trump lanzan misiles humanitarios" no será aceptada. Modifica también la gramática para que dentro de cada grupo nominal deba haber concordancia en número entre sustantivo y adjetivo (es decir, "misiles humanitario" será incorrecto). Indicación: para conseguir concordancia entre sujeto y predicado sólo hace falta tener dos categorías distintas para los sujetos y otras dos para los predicados.

15. Dado $\Sigma = \{a, b, c\}$, diseña un autómata con pila determinista que reconozca el lenguaje $\{a^n b^m c^p | n > m + p \geq 0\}$.

16. Dado $\Sigma = \{a, b\}$, diseña un autómata con pila determinista que reconozca el lenguaje $\{a^n b^m | m \geq n > 0\}$.

17. Para cada uno de los lenguajes siguientes, indica si son incontextuales o no. Si lo son, proporciona su gramática incontextual, y si no lo son demuéstralo:

    - $L_1 = \{a^n b^n b^n a^m | n, m \geq 0\}$
    - $L_2 = \{(ab)^n (ba)^m (ab)^n (ba)^m | n, m \geq 0\}$
    - $L_3 = \{a^n | n \text{ es el cubo de un número primo}\}$
    - $L_4 = \{a^n b^m b^n a^m | n, m \geq 0\}$
    - $L_5 = \{(ab)^n (ba)^m (ab)^m (ba)^n | n, m \geq 0\}$
    - $L_6 = \{a^n | n \text{ es un número primo}\}$

18. Para cada una de las siguientes afirmaciones, indica si es cierta o no, justificando tu respuesta:

    a) Si L_1 y L_2 son lenguajes incontextuales, su intersección no puede ser incontextual. Sin embargo, si L_2 fuera regular, tendríamos que la intersección entre L_1 y L_2 siempre es incontextual.

b) El complementario de un lenguaje incontextual puede ser incontextual, pero el complementario de uno no incontextual no puede ser incontextual.

c) Los autómatas con pila deterministas que aceptan por estado final tienen menos potencia expresiva que las gramáticas incontextuales no ambiguas, pero tienen la misma potencia expresiva que los autómatas con pila deterministas que aceptan por pila vacía.

d) Sean L_1 y L_2 lenguajes incontextuales. Si L_2 fuera regular, tendríamos que la intersección entre L_1 y L_2 *puede* ser incontextual.

e) El complementario de un lenguaje incontextual siempre es incontextual.

f) En una gramática incontextual, después de eliminar símbolos inalcanzables es posible que aparezcan nuevos símbolos no generadores. Además, después de eliminar los símbolos no generadores tampoco pueden aparecer nuevos símbolos inalcanzables.