Universidad Simón Bolívar
Departamento de Computación y Tecnología de la Información
CI3725 - Traductores e Interpretadores
Septiembre-Diciembre 2012

Examen Parcial II

(25 puntos)

Carnet:

Nombre:

1. Para cada uno de los siguientes lenguajes demuestre si es libre de contexto o no. Para demostrar que un lenguaje es libre de contexto puede construir una Gramática Libre de Contexto que lo genere o construir un PDA que lo acepte (no es necesario explicar la lógica de la gramática ni del PDA, basta que funcione). Para demostrar que un lenguaje no es libre de contexto, aplique el Lema de Bombeo para Lenguajes Libres de Contexto con toda la formalidad del caso. Debe usar un método diferente para cada lenguaje; usar el mismo método más de una vez invalida la pregunta, así como usar los dos métodos para un mismo lenguaje.

   a) **(5 puntos)** $L = \{a^n b a^n b a^n \mid n > 0\}$

   Asumamos que L es libre de contexto y sea k el especificado en el Lema de Bombeo para Lenguajes Libres de Contexto. Sea $z = a^k b a^k b a^k$, como $z \in L$ y $|z| > k$ entonces de acuerdo con el Lema de Bombeo para Lenguajes Libres de Contexto, z tiene al menos una forma de escribirse $uvwxy$ que satisface $|vwx| \leq k$, $|v| + |x| > 0$ y luego $\forall i \geq 0$ se cumple $uv^i wx^i y \in L$.

   Como $|vwx| \leq k$, entonces la subcadena vwx:

   - Puede estar compuesta exclusivamente de a, estando enteramente en el primer, segundo o tercer segmento a^k de la palabra. Si se bombea uv^2wx^2y entonces habrá más a en ese segmento que en los otros dos, por lo que la palabra resultante no pertenece a L.
   - Puede tener exactamente una b que *no* está ni en v ni en y, de modo que puede estar entre el primer y segundo segmento de a^k, o entre el segundo y tercer segmento de a^k. Si se bombea uv^2wx^2y entonces de seguro habrá menos a en aquel segmento que no está siendo cubierto por vwx, por lo que la palabra resultante no pertenece a L.
   - Puede tener exactamente una b que *si* está en v o en y. Si se bombea uv^2wx^2y entonces habrá más de dos b en la palabra, por lo que la palabra resultante no pertenece a L.

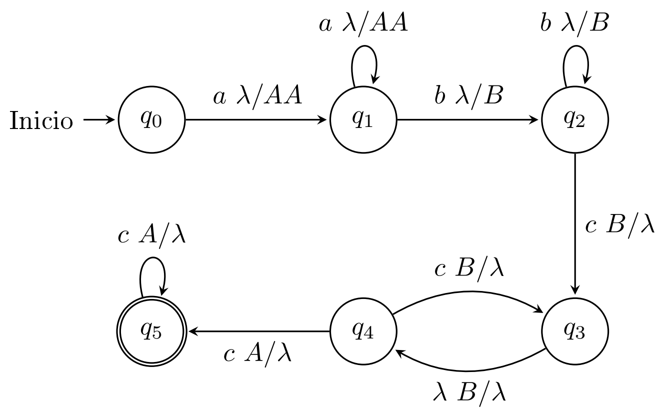
   Como todas las particiones posibles para la palabra contradicen el Lema de Bombeo para Lenguajes Libres de Contexto, se concluye que L no puede ser Lenguaje Libre de Contexto.

b) **(5 puntos)** $L = \{a^n b^{2m} c^{2n+m} \mid n, m > 0\}$

El lenguaje L puede ser generado con la gramática

$$
\begin{array}{rcl}
S & \rightarrow & \mathbf{a}S\mathbf{cc} \mid \mathbf{a}X\mathbf{cc} \\
X & \rightarrow & \mathbf{bb}X\mathbf{c} \mid \mathbf{bbc}
\end{array}
$$

El lenguaje L también puede ser generado con el PDA de la figura



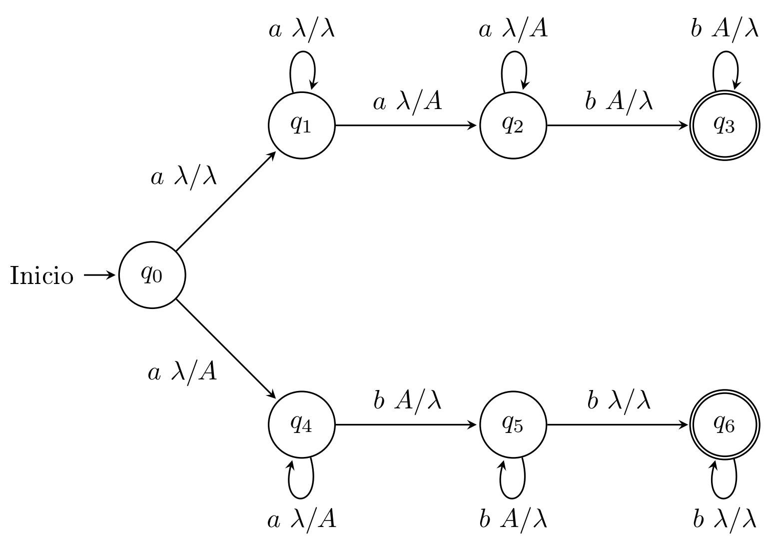
c) **(5 puntos)** $L = \{a^i b^j \mid i \neq j\}$

El lenguaje L corresponde a las palabras que comienzan con **a** y terminan con **b**, pero la cantidad de símbolos de cada tipo es diferente. Puesto que solamente hay dos símbolos involucrados, las palabras son aquellas en las cuales "sobran" **a** o "sobran" **b**. El lenguaje L puede ser generado con la gramática

$$
\begin{array}{rcl}
S & \rightarrow & AX \mid XB \\
X & \rightarrow & \mathbf{a}X\mathbf{b} \mid \lambda \\
A & \rightarrow & \mathbf{a} \mid \mathbf{a}A \\
B & \rightarrow & \mathbf{b} \mid \mathbf{b}B
\end{array}
$$

El no-terminal X genera una sub-palabra, posiblemente vacía, con la misma cantidad de símbolos **a** y **b**. El no-terminal S considera los dos casos posibles para una palabra en L: tiene un excedente no vacío de **a** al comienzo, o tiene un excedente no vacío de **b** al final.

El lenguaje L también puede ser generado con el PDA de la figura



2. Sea la gramática $G = (\{S\}, \{\mathbf{e}, \mathbf{w}, \mathbf{y}, !\}, P, S)$ cuyo conjunto de producciones P se define como

$$\begin{array}{lcl} S & \rightarrow & S! \\ S & \rightarrow & SS\mathbf{w} \\ S & \rightarrow & SS\mathbf{y} \\ S & \rightarrow & \mathbf{e} \end{array}$$

   a) **(5 puntos)** Modifique la gramática hasta que sea *fuertemente* $LL(1)$ y demuestre formalmente tal propiedad en la gramática resultante.

   La gramática no es $LL(1)$ porque el símbolo inicial es recursivo y además tiene recursión izquierda. Introducimos un nuevo símbolo inicial S', agregamos la marca de final de entrada y, después de eliminar la recursión izquierda, y factorizar, la gramática queda

$$\begin{array}{rcl} S' & \rightarrow & S\$ \\ S & \rightarrow & \mathbf{e}T \\ T & \rightarrow & !T \\ T & \rightarrow & SW \\ T & \rightarrow & \lambda \\ W & \rightarrow & \mathbf{w}T \\ W & \rightarrow & \mathbf{y}T \end{array}$$

   Calculamos el $FIRST$ y el $FOLLOW$ para cada símbolo no terminal

$$\begin{array}{rcl} FIRST(S') & = & \{\mathbf{e}\} \\ FIRST(S) & = & \{\mathbf{e}\} \\ FIRST(T) & = & \{\lambda, !, \mathbf{e}\} \\ FIRST(W) & = & \{\mathbf{w}, \mathbf{y}\} \\ FOLLOW(S') & = & \{\$\} \\ FOLLOW(S) & = & \{\$, \mathbf{w}, \mathbf{y}\} \\ FOLLOW(T) & = & \{\$, \mathbf{w}, \mathbf{y}\} \\ FOLLOW(W) & = & \{\$, \mathbf{w}, \mathbf{y}\} \end{array}$$

   Calculamos los Conjuntos de Lookahead $LA(A \rightarrow \alpha) = FIRST(FIRST(\alpha) \cdot FOLLOW(A))$ para las producciones

$$\begin{array}{rcl} LA(S' \rightarrow S\$) = FIRST(\{\mathbf{e}\} \cdot \{\$\}) & = & \{\mathbf{e}\} \\ LA(S \rightarrow \mathbf{e}T) = FIRST(\{\mathbf{e}\} \cdot \{\$\}) & = & \{\mathbf{e}\} \\ LA(T \rightarrow !T) = FIRST(\{!\} \cdot \{\$, \mathbf{w}, \mathbf{y}\}) & = & \{!\} \\ LA(T \rightarrow SW) = FIRST(\{\mathbf{e}\} \cdot \{\$, \mathbf{w}, \mathbf{y}\}) & = & \{\mathbf{e}\} \\ LA(T \rightarrow \lambda) = FIRST(\{\lambda\} \cdot \{\$, \mathbf{w}, \mathbf{y}\}) & = & \{\$, \mathbf{w}, \mathbf{y}\} \\ LA(W \rightarrow \mathbf{w}T) = FIRST(\{\mathbf{w}\} \cdot \{\$, \mathbf{w}, \mathbf{y}\}) & = & \{\mathbf{w}\} \\ LA(W \rightarrow \mathbf{y}T) = FIRST(\{\mathbf{y}\} \cdot \{\$, \mathbf{w}, \mathbf{y}\}) & = & \{\mathbf{y}\} \end{array}$$

   Puede observarse que $\forall A \rightarrow \alpha|\beta \in P, \alpha \neq \beta$ se cumple $LA(A \rightarrow \alpha) \cap LA(A \rightarrow \beta) = \emptyset$, por lo tanto la gramática es fuertemente $LL(1)$.

b) **(2 puntos)** Construya la tabla de análisis para un reconocedor predictivo no recursivo utilizando el algoritmo presentado en clase.

	e	**w**	**y**	!	\$
S'	$S' \to S\$$				
S	$S \to \mathbf{e}T$				
T	$T \to SW$	$T \to \lambda$	$T \to \lambda$	$T \to !T$	$T \to \lambda$
W		$W \to \mathbf{w}T$		$W \to \mathbf{y}T$	

c) **(3 puntos)** Use el reconocedor para encontrar la derivación más izquierda de la palabra **eey!eew!eww!**

Pila	Entrada	Acción
$S'\$$	**eey!eew!eww!**\$	$S' \to S\$$
$S\$$	**eey!eew!eww!**\$	$S \to \mathbf{e}T$
$\mathbf{e}T\$$	**eey!eew!eww!**\$	Consumir **e**
$T\$$	**ey!eew!eww!**\$	$T \to SW$
$SW\$$	**ey!eew!eww!**\$	$S \to \mathbf{e}T$
$\mathbf{e}TW\$$	**ey!eew!eww!**\$	Consumir **e**
$TW\$$	**y!eew!eww!**\$	$T \to \lambda$
$W\$$	**y!eew!eww!**\$	$W \to \mathbf{y}T$
$\mathbf{y}T\$$	**y!eew!eww!**\$	Consumir **y**
$T\$$	**!eew!eww!**\$	$T \to !T$
$!T\$$	**!eew!eww!**\$	Consumir !
$T\$$	**eew!eww!**\$	$T \to SW$
$SW\$$	**eew!eww!**\$	$S \to \mathbf{e}T$
$\mathbf{e}TW\$$	**eew!eww!**\$	Consumir **e**
$TW\$$	**ew!eww!**\$	$T \to SW$
$SWW\$$	**ew!eww!**\$	$S \to \mathbf{e}T$
$\mathbf{e}TWW\$$	**ew!eww!**\$	Consumir **e**
$TWW\$$	**w!eww!**\$	$T \to \lambda$
$WW\$$	**w!eww!**\$	$W \to \mathbf{w}T$
$\mathbf{w}TW\$$	**w!eww!**\$	Consumir **w**
$TW\$$	**!eww!**\$	$T \to !T$
$!TW\$$	**!eww!**\$	Consumir !
$TW\$$	**eww!**\$	$T \to SW$
$SWW\$$	**eww!**\$	$S \to \mathbf{e}T$
$\mathbf{e}TWW\$$	**eww!**\$	Consumir **e**
$TWW\$$	**ww!**\$	$T \to \lambda$
$WW\$$	**ww!**\$	$W \to \mathbf{w}T$
$\mathbf{w}TW\$$	**ww!**\$	Consume **w**
$TW\$$	**w!**\$	$T \to \lambda$
$W\$$	**w!**\$	$W \to \mathbf{w}T$
$\mathbf{w}T\$$	**w!**\$	Consumir **w**
$T\$$	**!**\$	$T \to !T$
$!T\$$	**!**\$	Consumir !
$T\$$	\$	$T \to \lambda$
\$	\$	**Acepta**

Por lo tanto, la derivación más izquierda para la palabra será

$$\begin{aligned}
S' &\Rightarrow \underline{S} \\
&\Rightarrow \mathbf{e}\underline{T} \\
&\Rightarrow \mathbf{e}\underline{S}W \\
&\Rightarrow \mathbf{ee}\underline{T}W \\
&\Rightarrow \mathbf{ee}\underline{W} \\
&\Rightarrow \mathbf{eey}\underline{T} \\
&\Rightarrow \mathbf{eey!}\underline{T} \\
&\Rightarrow \mathbf{eey!}\underline{S}W \\
&\Rightarrow \mathbf{eey!e}\underline{T}W \\
&\Rightarrow \mathbf{eey!e}\underline{S}WW \\
&\Rightarrow \mathbf{eey!ee}\underline{T}WW \\
&\Rightarrow \mathbf{eey!ee}\underline{W}W \\
&\Rightarrow \mathbf{eey!eew}\underline{T}W \\
&\Rightarrow \mathbf{eey!eew!}\underline{T}W \\
&\Rightarrow \mathbf{eey!eew!}\underline{S}WW \\
&\Rightarrow \mathbf{eey!eew!e}\underline{T}WW \\
&\Rightarrow \mathbf{eey!eew!e}\underline{W}W \\
&\Rightarrow \mathbf{eey!eew!ew}\underline{T}W \\
&\Rightarrow \mathbf{eey!eew!eww}\underline{W} \\
&\Rightarrow \mathbf{eey!eew!eww}\underline{T} \\
&\Rightarrow \mathbf{eey!eew!eww!}\underline{T} \\
&\Rightarrow \mathbf{eey!eew!eww!}
\end{aligned}$$