Universidad Simón Bolívar
Departamento de Computación y Tecnología de la Información
CI3725 - Traductores e Interpretadores
Septiembre-Diciembre 2011

Examen Parcial II

(25 puntos)

Carnet:

Nombre:

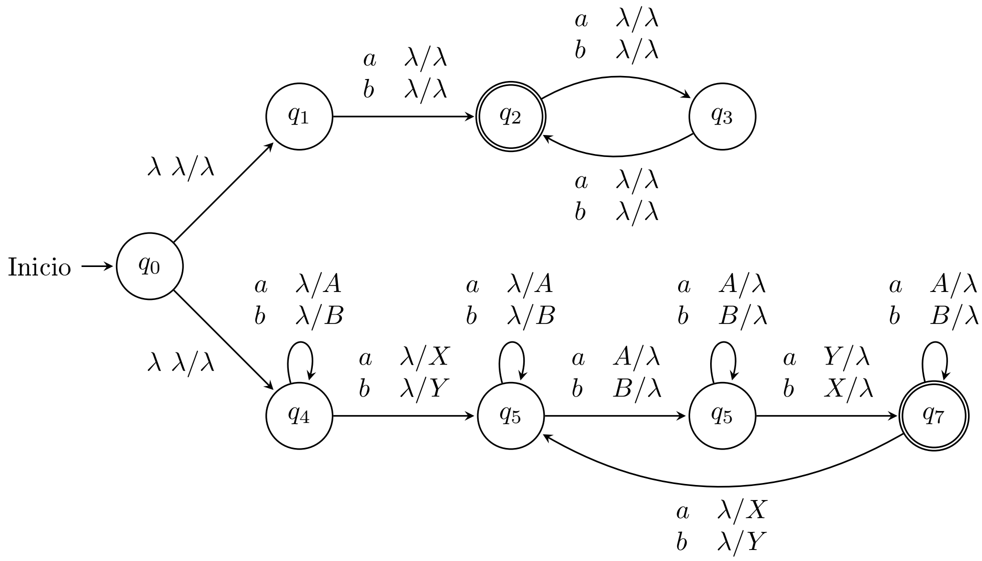
1. Para cada uno de los siguientes lenguajes demuestre si es libre de contexto o no. Para demostrar que un lenguaje es libre de contexto puede construir una Gramática Libre de Contexto que lo genere o construir un PDA que lo acepte (no es necesario explicar la lógica de la gramática ni del PDA, basta 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 más de un método para el mismo lenguaje invalida la pregunta.

   (a) **(4 puntos)** $L = \{a, b\}^* - \{ww^R \mid w \in \{a, b\}^*\}$
   El lenguaje corresponde a cualquier palabra de longitud impar y cualquier palabra de longitud par que *no* sea palíndromo. El lenguaje L puede ser generado con la gramática

$$
\begin{array}{lcl}
S & \rightarrow & A \mid B \\
A & \rightarrow & \mathbf{a}P \mid \mathbf{b}P \\
P & \rightarrow & \mathbf{aa}P \mid \mathbf{ab}P \mid \mathbf{ba}P \mid \mathbf{bb}P \mid \lambda \\
B & \rightarrow & \mathbf{a}B\mathbf{a} \mid \mathbf{b}B\mathbf{b} \mid C \\
C & \rightarrow & \mathbf{a}D\mathbf{b} \mid \mathbf{b}D\mathbf{a} \\
D & \rightarrow & \mathbf{a}D\mathbf{a} \mid \mathbf{b}D\mathbf{b} \mid C \mid \lambda
\end{array}
$$

   En esta gramática el no-terminal A corresponde a generar palabras de longitud impar, para lo cual se garantiza la presencia de al menos un símbolo (A) seguido de una cantidad par de símbolos (P) en todas las posibles combinaciones, incluyendo la vacía. Así mismo, el no-terminal B corresponde a generar palabras con longitud par que no sean palíndromos, para lo cual se utilizan las mismas reglas que uno utilizaría para construir palíndromos, pero incluyendo un nivel intermedio (C) en el cual se "perturban" los palíndromos asegurando al menos una ocurrencia de símbolos que no son simétricos.

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



(b) **(4 puntos)** $L = \{ww^Rw \mid w \in \{a,b\}^*\}$

Asumamos que L es libre de contexto y sea k el especificado en el Lema de Bombeo para Lenguajes Libres de Contexto. Sea $z = a^k b^{2k} a^{2k} b^k$, 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 $uv^iwx^iy \in L$.

Como $|vwx| \leq k$, entonces la subcadena vwx puede tener las siguientes formas:

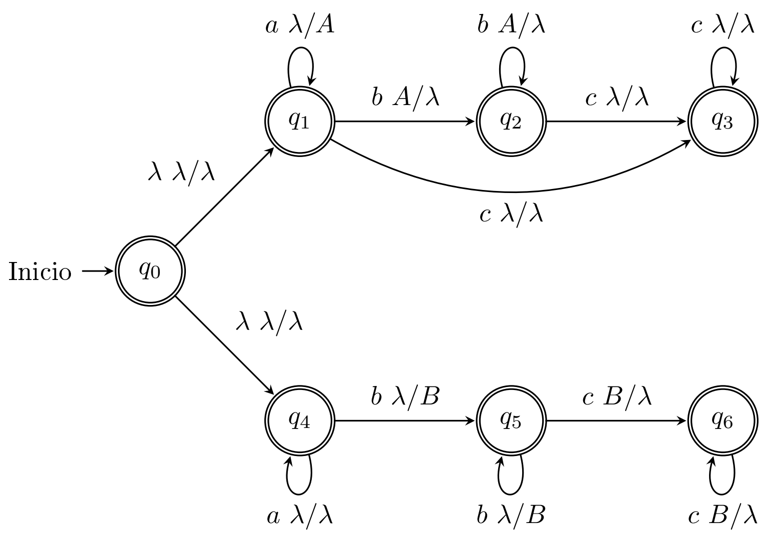
- v o x, e incluso ambas, contienen a^py a^q tomadas del segmento a^k en la palabra. Si se bombea uv^0wx^0y entonces habrá menos a que en el segmento b^k final, por lo que la palabra resultante no pertenece a L.
- v o x, e incluso ambas, contienen b^p y b^q tomadas del segmento b^{2k} en la palabra. Si se bombea uv^0wx^0y entonces habrá menos b que las correspondientes a en el segmento a^{2k}, por lo que la palabra resultante no pertenece a L.
- v o x, e incluso ambas, contienen a^p y a^q tomadas del segmento a^{2k} en la palabra. Si se bombea uv^0wx^0y entonces habrá menos a que las correspondientes b en el segmento b^{2k}, por lo que la palabra resultante no pertenece a L.
- v o x, e incluso ambas, contienen b^p y b^q tomadas del segmento b^k en la palabra. Si se bombea uv^0wx^0y entonces habrá menos b que en el segmento a^k inicial, por lo que la palabra resultante no pertenece a L.
- v contiene a^p del segmento a^k y x contiene b^q del segmento b^{2k}. Si se bombea uv^2wx^2y, si bien p y q podrían aumentar simultáneamente, el segmento b^k no puede aumentar, por lo que la palabra resultante no pertenece a L.
- v contiene b^p del segmento b^{2k} y x contiene a^q del segmento a^k. Si se bombea uv^2wx^2y, si bien p y q podrían aumentar simultáneamente, el segmento a^k no puede aumentar, 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.

(c) **(5 puntos)** $L = \{a^i b^j c^k \mid i = j \lor j = k,\ i \geq 0,\ j \geq 0,\ k \geq 0\}$
El lenguaje L puede ser generado con la gramática

$$\begin{array}{rcl} S & \to & LC \mid AR \\ L & \to & \mathbf{a}L\mathbf{b} \mid \lambda \\ C & \to & \mathbf{c}C \mid \lambda \\ A & \to & \mathbf{a}A \mid \lambda \\ R & \to & \mathbf{b}R\mathbf{c} \mid \lambda \end{array}$$

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



2. Sea la gramática $G = (\{S\}, \{\mathbf{m}, \mathbf{p}, \mathbf{q}, \mathbf{s}, \mathbf{v}, \mathbf{x}\}, P, S)$ cuyo conjunto de producciones P se define como

$$\begin{array}{rcl} S & \to & SS\mathbf{s} \\ S & \to & SS\mathbf{m} \\ S & \to & \mathbf{x} \\ S & \to & \mathbf{v} \\ S & \to & \mathbf{vp}S\mathbf{q} \end{array}$$

(a) **(4 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

$$\begin{array}{rcl} S' & \to & S\$ \\ S & \to & \mathbf{x}T \\ S & \to & \mathbf{v}V \\ V & \to & T \\ V & \to & \mathbf{p}S\mathbf{q}T \\ T & \to & SW \\ T & \to & \lambda \\ W & \to & \mathbf{s}T \\ W & \to & \mathbf{m}T \end{array}$$

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

$$\begin{aligned}
FIRST(S') &= \{\mathbf{v}, \mathbf{x}\} \\
FIRST(S) &= \{\mathbf{v}, \mathbf{x}\} \\
FIRST(V) &= \{\lambda, \mathbf{p}, \mathbf{v}, \mathbf{x}\} \\
FIRST(T) &= \{\lambda, \mathbf{v}, \mathbf{x}\} \\
FIRST(W) &= \{\mathbf{m}, \mathbf{s}\} \\
FOLLOW(S') &= \{\$\} \\
FOLLOW(S) &= \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\} \\
FOLLOW(V) &= \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\} \\
FOLLOW(T) &= \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\} \\
FOLLOW(W) &= \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}
\end{aligned}$$

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

$$\begin{aligned}
LA(S' \to S\$) = FIRST(\{\mathbf{v}, \mathbf{x}\} \cdot \{\$\}) &= \{\mathbf{v}, \mathbf{x}\} \\
LA(S \to \mathbf{x}T) = FIRST(\{\mathbf{x}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\mathbf{x}\} \\
LA(S \to \mathbf{v}V) = FIRST(\{\mathbf{v}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\mathbf{v}\} \\
LA(V \to T) = FIRST(\{\lambda, \mathbf{v}, \mathbf{x}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}, \mathbf{v}, \mathbf{x}\} \\
LA(V \to \mathbf{p}S\mathbf{q}T) = FIRST(\{\mathbf{p}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\mathbf{p}\} \\
LA(T \to SW) = FIRST(\{\mathbf{v}, \mathbf{x}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\mathbf{v}, \mathbf{x}\} \\
LA(T \to \lambda) = FIRST(\{\lambda\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\} \\
LA(W \to \mathbf{s}T) = FIRST(\{\mathbf{s}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\mathbf{s}\} \\
LA(W \to \mathbf{m}T) = FIRST(\{\mathbf{m}\} \cdot \{\$, \mathbf{m}, \mathbf{q}, \mathbf{s}\}) &= \{\mathbf{m}\}
\end{aligned}$$

Puede observarse que $\forall A \to \alpha|\beta \in P, \alpha \neq \beta$ se cumple $LA(A \to \alpha) \cap LA(A \to \beta) = \emptyset$, 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.

	m	**p**	**q**	**s**	**v**	**x**	$
S'					$S' \to S\$$	$S' \to S\$$	
S					$S \to \mathbf{v}V$	$S \to \mathbf{x}T$	
V	$V \to T$	$V \to \mathbf{p}S\mathbf{q}T$	$V \to T$	$V \to T$	$V \to T$	$V \to T$	$V \to T$
T	$T \to \lambda$		$T \to \lambda$	$T \to \lambda$	$T \to SW$	$T \to SW$	$T \to \lambda$
W	$W \to \mathbf{m}T$			$W \to \mathbf{s}T$			

(a) **(3 puntos)** Use el reconocedor para encontrar la derivación más izquierda de la palabra **xvsvpvxmxsqm**.

Pila	Entrada	Acción
$S'\$$	**xvsvpvxmxsqm**\$	$S' \to S\$$
$S\$$	**xvsvpvxmxsqm**\$	$S \to \mathbf{x}T$
$\mathbf{x}T\$$	**xvsvpvxmxsqm**\$	Consumir **x**
$T\$$	**vsvpvxmxsqm**\$	$T \to SW$
$SW\$$	**vsvpvxmxsqm**\$	$S \to \mathbf{v}V$
$\mathbf{v}VW\$$	**vsvpvxmxsqm**\$	Consumir **v**
$VW\$$	**svpvxmxsqm**\$	$V \to T$
$TW\$$	**svpvxmxsqm**\$	$T \to \lambda$
$W\$$	**svpvxmxsqm**\$	$W \to \mathbf{s}T$
$\mathbf{s}T\$$	**svpvxmxsqm**\$	Consumir **s**
$T\$$	**vpvxmxsqm**\$	$T \to SW$
$SW\$$	**vpvxmxsqm**\$	$S \to \mathbf{v}V$
$\mathbf{v}VW\$$	**vpvxmxsqm**\$	Consumir **v**
$VW\$$	**pvxmxsqm**\$	$V \to \mathbf{p}S\mathbf{q}T$
$\mathbf{p}S\mathbf{q}TW\$$	**pvxmxsqm**\$	Consumir **p**
$S\mathbf{q}TW\$$	**vxmxsqm**\$	$S \to \mathbf{v}V$
$\mathbf{v}V\mathbf{q}TW\$$	**vxmxsqm**\$	Consumir **v**
$V\mathbf{q}TW\$$	**xmxsqm**\$	$V \to T$
$T\mathbf{q}TW\$$	**xmxsqm**\$	$T \to SW$
$SW\mathbf{q}TW\$$	**xmxsqm**\$	$S \to \mathbf{x}T$
$\mathbf{x}TW\mathbf{q}TW\$$	**xmxsqm**\$	Consumir **x**
$TW\mathbf{q}TW\$$	**mxsqm**\$	$T \to \lambda$
$W\mathbf{q}TW\$$	**mxsqm**\$	$W \to \mathbf{m}T$
$\mathbf{m}T\mathbf{q}TW\$$	**mxsqm**\$	Consumir **m**
$T\mathbf{q}TW\$$	**xsqm**\$	$T \to SW$
$SW\mathbf{q}TW\$$	**xsqm**\$	$S \to \mathbf{x}T$
$\mathbf{x}TW\mathbf{q}TW\$$	**xsqm**\$	Consumir **x**
$TW\mathbf{q}TW\$$	**sqm**\$	$T \to \lambda$
$W\mathbf{q}TW\$$	**sqm**\$	$W \to \mathbf{s}T$
$\mathbf{s}T\mathbf{q}TW\$$	**sqm**\$	Consumir **s**
$T\mathbf{q}TW\$$	**qm**\$	$T \to \lambda$
$\mathbf{q}TW\$$	**qm**\$	Consumir **q**
$TW\$$	**m**\$	$T \to \lambda$
$W\$$	**m**\$	$W \to \mathbf{m}T$
$\mathbf{m}T\$$	**m**\$	Consumir **m**
$T\$$	\$	$T \to \lambda$
\$	\$	**Acepta**

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

$$
\begin{aligned}
S' &\Rightarrow \underline{S} \\
&\Rightarrow \mathbf{x}\underline{T} \\
&\Rightarrow \mathbf{x}\underline{S}W \\
&\Rightarrow \mathbf{xv}\underline{V}W \\
&\Rightarrow \mathbf{xv}\underline{T}W \\
&\Rightarrow \mathbf{xv}\underline{W} \\
&\Rightarrow \mathbf{xvs}\underline{T} \\
&\Rightarrow \mathbf{xvs}\underline{S}W \\
&\Rightarrow \mathbf{xvsv}\underline{V}W \\
&\Rightarrow \mathbf{xvsvp}\underline{S}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpv}\underline{V}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpv}\underline{T}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpv}\underline{S}W\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvx}\underline{T}W\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvx}\underline{W}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvxm}\underline{T}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvxm}\underline{S}W\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvxmx}\underline{T}W\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvxmx}\underline{W}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvxmxs}\underline{T}\mathbf{q}TW \\
&\Rightarrow \mathbf{xvsvpvxmxsq}\underline{T}W \\
&\Rightarrow \mathbf{xvsvpvxmxsq}\underline{W} \\
&\Rightarrow \mathbf{xvsvpvxmxsqm}\underline{T} \\
&\Rightarrow \mathbf{xvsvpvxmxsqm}
\end{aligned}
$$