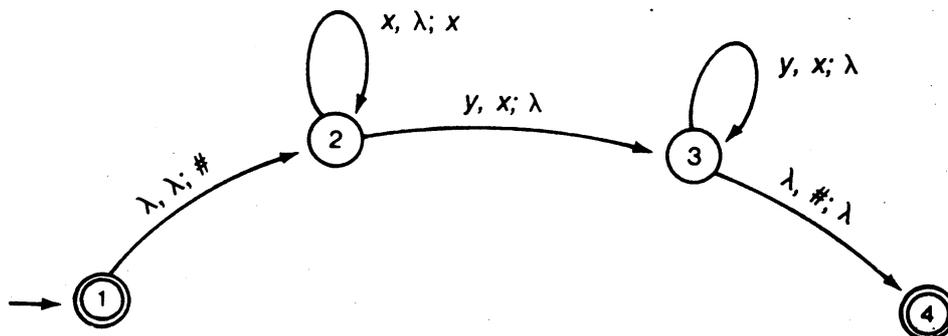


AUTÓMATAS DE PILA Y LENGUAJES INDEPENDIENTES DEL CONTEXTO

AUTÓMATAS DE PILA

- Son autómatas finitos con una memoria en forma de pila.
- Símbolos de pila:
 - Terminales
 - Alfabeto de la pila
 - #: símbolo de fondo de pila (indicador de pila vacía).
- Transiciones: $(p,x,s;q,y)$ donde:
 - p,q : estados
 - x : símbolo de entrada (puede ser λ)
 - s,y : símbolos de pila (uno de los dos o ambos pueden ser λ)
- Ejemplo de diagrama de transiciones



- Proceso de reconocimiento de una cadena:
 - Se parte del estado inicial y la pila vacía.
 - Se lee la cadena símbolo a símbolo de izquierda a derecha.
 - Por cada símbolo leído se produce una transición desde el estado actual a otro a través de la flecha cuyo símbolo de entrada coincide con el símbolo leído, siempre y cuando la cabecera de la pila coincida con el símbolo de pila que figura a la izquierda del punto y coma.
 - La cabecera de la pila es sustituida en cada transición por el símbolo de pila a la derecha del punto y coma en la etiqueta de la flecha.
 - El autómata puede realizar transiciones entre estados sin consumir símbolos de entrada o símbolos de pila a través de flechas en las que en el lugar correspondiente de la etiqueta aparezca el símbolo λ .
 - La cadena es reconocida si es posible que el autómata alcance un estado de aceptación, estando completamente consumida la cadena de entrada.

- AP: $M(S, \Sigma, \Gamma, T, i, F)$

- S : conjunto finito de estados.
- Σ : alfabeto de entrada.
- Γ : alfabeto de símbolos de pila.
- T : colección finita de transiciones: $S \times (\Sigma \cup \lambda) \times (\Gamma \cup \lambda) \rightarrow S \times (\Gamma \cup \lambda)$.
- i : estado inicial.
- F : conjunto de estados de aceptación: $F \subseteq S$

- En general, los autómatas de pila son no deterministas: pueden existir una, varias o ninguna transición aplicable en un momento dado.

- Autómatas deterministas: no se encuentran con alternativas.

- Se puede admitir que un autómata de pila sea capaz de insertar en la cima de la pila más de un símbolo en una sola transición.

- Aceptación de una cadena: el autómata debe alcanzar un estado de aceptación después de haber consumido toda la cadena de entrada.

- Puede exigirse también que la pila esté vacía al final del proceso de reconocimiento.

Teorema 2.1

Cada autómata de pila que acepta cadenas sin vaciar su pila tiene un autómata equivalente que lo hace vaciando su pila.

DEMOSTRACIÓN:

Sea $M(S, \Sigma, \Gamma, T, i, F)$ un autómata de pila que acepta cadenas sin vaciar la pila. Se puede modificar M de la forma siguiente:

- Incluir un nuevo estado inicial i' , y una transición $(i', \lambda, \lambda; i, \#)$.
- Convertir todos los estados de F en estados de no aceptación, y añadir un estado nuevo p , y de cada antiguo estado s de F añadir la transición $(s, \lambda, \lambda; p, \lambda)$.
- Añadir $(p, \lambda, x; p, \lambda) \forall x \in \Gamma$.
- Añadir un nuevo estado de aceptación q , al que llegará la transición $(p, \lambda, \#; q, \lambda)$.

El nuevo autómata aceptará las mismas cadenas que el original, pero dejando la pila vacía.

GRAMÁTICAS INDEPENDIENTES DEL CONTEXTO

- En sus reglas aparece a la izquierda un único símbolo no terminal, y a la derecha cualquier combinación de símbolos terminales y no terminales, o la palabra vacía.

- Derivaciones: árboles que se pueden leer de diferentes formas.

- Derivación por la derecha: siempre se aplican las reglas de reescritura al símbolo no terminal más a la derecha de la cadena de derivación.
- Derivación por la izquierda: siempre se aplican las reglas de reescritura al símbolo no terminal más a la izquierda de la cadena de derivación.
- Ejemplo:

1. $S \rightarrow AB$
2. $S \rightarrow A$
3. $A \rightarrow aAa$
4. $A \rightarrow \lambda$
5. $B \rightarrow Bb$
6. $B \rightarrow b$

Derivación por la izquierda:

$S \Rightarrow AB \Rightarrow aAaB \Rightarrow aaB \Rightarrow aaBb \Rightarrow aaBbb \Rightarrow aabbb$

Derivación por la derecha:

$S \Rightarrow AB \Rightarrow ABb \Rightarrow ABbb \Rightarrow Abbb \Rightarrow aAabbb \Rightarrow aabbb$

Teorema 2.2

Para cada gramática independiente del contexto $G(V, \Sigma, A_x, R)$ existe un autómata de pila $M(S, \Sigma, \Gamma, T, i, F)$ tal que $L(G) = L(M)$.

DEMOSTRACIÓN

Dada una gramática $G(V, \Sigma, A_x, R)$ independiente del contexto, se crea un autómata de pila $M(S, \Sigma, \Gamma, T, i, F)$ de la manera siguiente:

- El conjunto de símbolos de entrada de M coincide con los símbolos terminales de G , mientras que $\Gamma = \Sigma \cup V \cup \{\#\}$
- El conjunto de estados de M será $S = \{i, p, q, f\}$
- $F = \{f\}$
- Se introduce la transición $(i, \lambda, \lambda; p, \#)$
- Se introduce la transición $(p, \lambda, \lambda; q, A_x)$, siendo A_x el axioma de la gramática.
- Para cada regla de reescritura $N \rightarrow w$ de la gramática se introduce una transición de la forma $(q, \lambda, N; q, w)$.
- Para cada símbolo terminal de G se introduce una transición $(q, x, x; q, \lambda)$.
- Introducir la transición $(q, \lambda, \#; f, \lambda)$.

Teorema 2.3

Para cada autómata de pila $M(S, \Sigma, \Gamma, T, i, F)$ existe una gramática independiente del contexto $G(V, \Sigma, A_x, R)$ tal que $L(G) = L(M)$.

FORMA NORMAL DE CHOMSKY

- Una gramática independiente del contexto se encuentra en forma normal de Chomsky si el lado derecho de cada una de sus reglas está constituido por un único símbolo terminal o por dos no terminales.

Eliminación de reglas λ

- Encadenamiento λ de longitud n : secuencia de reglas de la forma

$$N_n \rightarrow N_{n-1}, N_{n-1} \rightarrow N_{n-2}, \dots, N_1 \rightarrow N_0, N_0 \rightarrow \lambda.$$

- Cálculo de encadenamientos λ de una gramática:

- Se define $U_0 = \{N_i \mid N_i \rightarrow \lambda \in R\}$ encadenamiento de longitud 0.
- Se calcula U_1 como $U_1 = U_0 \cup \{N_i \mid N_i \rightarrow N_j \in R, N_j \in U_0\}$
- Se prosigue calculando los encadenamientos de sucesivas longitudes, hasta que $U_n = U_{n-1} = U$

- Para cada gramática G puede obtenerse otra G' que no tiene reglas λ si se eliminan todas ellas, y para cada regla $N \rightarrow w$ se añade otra regla $N \rightarrow w'$ donde w' es cualquier cadena no vacía obtenida al eliminar de w una o más ocurrencias de símbolos no terminales que pertenezcan a U .

Teorema 2.4

Si L es un lenguaje independiente del contexto sin la palabra vacía, existe una gramática en forma normal de Chomsky que genera dicho lenguaje.

DEMOSTRACIÓN

Sea la gramática $G(V, \Sigma, S, R)$, que no contiene reglas λ . Se realiza sobre ella el siguiente proceso:

1. Para cada símbolo terminal x se introduce la regla $X \rightarrow x$, donde X es un símbolo no terminal nuevo que no pertenecía a V , y se sustituye x por X en todas las demás reglas de G .
2. Cada regla de la forma $N \rightarrow N_1 N_{n-2} \dots N_n$, siendo $n \geq 2$, se sustituye por la secuencia de reglas $N \rightarrow N_1 R_1, R_1 \rightarrow N_2 R_2, \dots, R_{n-2} \rightarrow N_{n-1} N_n$, donde los R_i son símbolos no terminales nuevos que no pertenecían a V .

Al llegar a este punto, la gramática contiene sólo reglas que en la parte derecha tienen un único terminal, dos no terminales o un único no terminal. Estas últimas reglas se pueden eliminar efectuando el siguiente paso:

3. Para cada secuencia de reglas $N_n \rightarrow N_{n-1}, N_{n-1} \rightarrow N_{n-2}, \dots, N_2 \rightarrow N_1$ se añade la regla $N_n \rightarrow x$ si $N_1 \rightarrow x$ estaba en G , y $N_n \rightarrow AB$ si $N_1 \rightarrow AB$ estaba en G , y se eliminan entonces las reglas que tienen un único no terminal en su parte derecha.

- Si el lenguaje contiene la palabra vacía, se añade un nuevo símbolo S' , que será el inicial, con una regla $S' \rightarrow \lambda$, y para cada regla $S \rightarrow w$ en G , siendo w un terminal o dos no terminales, se añade la regla $S' \rightarrow w$.

LÍMITES DE LOS AUTÓMATAS DE PILA

Lema de bombeo para lenguajes independientes del contexto

- Si L es un lenguaje independiente del contexto con un número infinito de cadenas, existirá en L una cadena con la forma $svuwt$, siendo v o w no vacías, y también existirán en L cadenas de la forma $sv^n uw^n t$ para todo $n > 0$.

Autómatas de pila deterministas

- En cada momento tienen una y sólo una opción posible.

Teorema 2.6

Existen lenguajes independientes del contexto que no pueden ser aceptados por ningún autómata de pila determinista.

- Lenguajes independientes del contexto deterministas: aceptados por algún autómata de pila determinista.
- Existen lenguajes independientes del contexto que no pueden ser aceptados por ningún autómata de pila determinista que vacíe su pila.

Principio de preanálisis

- Técnica que permite a los autómatas de pila observar uno o varios símbolos más allá de donde se encuentra la cabeza lectora del autómata, pero sin leerlos realmente.
- Esta técnica permite superar el no determinismo de algunos autómatas de pila.

ANALIZADORES SINTÁCTICOS $LL(k)$

- Leen la cadena de entrada de izquierda a derecha, y producen una derivación por la izquierda.
- Se basan en el proceso de reconocimiento de los autómatas de pila construidos según el teorema 2.2
- Se les denomina:
 - Analizadores sintácticos predictivos, pues deben predecir cuál es la siguiente regla de reescritura que debe aplicarse en la derivación.
 - Analizadores sintácticos descendentes, pues insertan en la pila el símbolo inicial de la gramática y terminan generando en ella una cadena de símbolos equivalente a la cadena de entrada.
 - * Para ello emplean una técnica de preanálisis con k símbolos.
- Existen lenguajes independientes del contexto deterministas que no pueden ser reconocidos por analizadores $LL(k)$, cualquiera que sea k .

Proceso de análisis LL

- Previo: marcar el fondo de la pila (#) e insertar el símbolo inicial de la gramática.
- Proceso: repetición de los siguientes pasos
 1. Si en la cima de la pila hay un símbolo no terminal, reemplazarlo de acuerdo con una regla de la gramática.
 2. Si en la cima de la pila hay un símbolo terminal que coincide con el símbolo actual de entrada, sacarlo de la pila y leerlo de la entrada. Si dichos símbolos no coinciden, la cadena es rechazada.
 3. Si en la cabecera de la pila aparece #, la cadena es aceptada.
- El principio de preanálisis permite elegir entre posibles alternativas en el paso 1.

Tablas de análisis sintáctico $LL(k)$

- Matriz bidimensional cuyos elementos indican la parte derecha de la regla que se debe aplicar en el paso 1 del proceso de análisis cuando aparece un cierto símbolo no terminal en la cabecera de la pila, dada una cierta combinación de símbolos de preanálisis.
 - Tantas filas como símbolos no terminales.
 - Tantas columnas como combinaciones posibles de símbolos de preanálisis (incluyendo el fin de cadena).
- Error: cuando no se puede aplicar ninguna regla.

ANALIZADORES SINTÁCTICOS LR(k)

- Leen la cadena de entrada de izquierda a derecha, y producen una derivación por la derecha.
- Se les denomina analizadores sintácticos ascendentes, pues el proceso de reconocimiento que emplean conduce a la inserción en la pila del símbolo inicial de la gramática.
 - Para ello emplean una técnica de preanálisis con k símbolos.
- Los lenguajes aceptados por los analizadores LR(k) son los lenguajes independientes del contexto deterministas.
- Se basan en el proceso de reconocimiento de los autómatas de pila construidos según el siguiente procedimiento:
 - Establecer 4 estados: i (inicial), f (final), p y q .
 - Introducir las transiciones $(i, \lambda, \lambda; p, \#)$ y $(q, \lambda, \#; f, \lambda)$.
 - Para cada símbolo terminal de la gramática establecer la transición $(p, x, \lambda; p, x)$ (operación de desplazamiento).
 - Para cada regla de reescritura $N \rightarrow w$ establecer la transición $(p, \lambda, w; p, N)$ (operación de reducción).
 - Introducir la transición $(p, \lambda, S; q, \lambda)$.

Proceso de análisis LR

- Previo: marcar el fondo de la pila ($\#$).
- Proceso:
 1. Ir insertando en la pila los símbolos terminales que aparecen a la entrada (operación de desplazamiento).
 2. Cuando los símbolos en la cabecera de la pila coincidan con la parte derecha de alguna regla, sustituirlos por el no terminal en la izquierda de dicha regla (operación de reducción).
 3. Al final, en la pila sólo quedan el axioma y $\#$.
- El principio de preanálisis permite seleccionar entre posibles alternativas a la hora de elegir si se realiza una operación de desplazamiento o de reducción, y, en este último caso, también sirve para identificar cuál de las reglas de la gramática va a ser reducida.
- Las reglas involucradas en las operaciones de reducción leídas de abajo a arriba dan una derivación por la derecha.

Tablas de análisis sintáctico LR(k)

- Se utilizan para evitar tener que consultar más de un símbolo de la pila al realizar operaciones de reducción en el paso 2 del proceso de análisis LR.
- Las tablas de análisis sintáctico LR(k) se basan en la existencia de un autómata finito que es capaz de aceptar las cadenas de símbolos de la gramática (terminales y no terminales) que conducen a una operación de reducción.
 - Estados del autómata finito: símbolos especiales.
- Cada vez que se inserta en la pila un símbolo terminal (desplazamiento) o no terminal (reducción), también se inserta un símbolo especial.
 - El símbolo especial insertado es un patrón que representa lo que hay por debajo de él en la pila.
- Una tabla de análisis LR(1) es una matriz bidimensional cuyos elementos indican la acción que debe realizar el analizador en cada momento (desplazar, reducir, aceptar, error) y los símbolos que debe insertar en la pila.
 - Columnas: símbolos terminales, fin de cadena y no terminales.
 - * Columnas con terminales o fin de cadena: operaciones de desplazamiento, reducción, aceptación o error.
 - * Columnas con no terminales: indican el símbolo especial insertado en la cima de la pila tras una reducción que tuviese en la parte izquierda de la regla el no terminal de la columna.
 - Filas: símbolos especiales (estados del autómata finito).
- Proceso detallado de análisis:
 - Previo: se inserta el símbolo especial 1 en la pila.
 - Proceso: acceder a la casilla (X,a) , siendo X el símbolo especial en la cabecera de la pila y a el símbolo terminal de preanálisis, y realizar la operación indicada en dicha casilla:
 - * Desplazar Y : operación de desplazamiento. Se inserta el símbolo terminal a en la pila, y a continuación el símbolo especial Y .
 - * $N \rightarrow w$: operación de reducción. Se extraen de la pila $2*|w|$ símbolos. En la cabecera de la pila quedará un símbolo especial Z , sobre el cual se insertará el no terminal N y encima de él el símbolo especial contenido en la casilla (Z,N) .
 - * Aceptar: la cadena es aceptada.
 - * Casilla en blanco: la cadena es rechazada.
 - El proceso termina al alcanzar una casilla de aceptación o de rechazo.

COMENTARIOS FINALES

- La unión de LIC es un LIC.
- La intersección de dos LIC no tiene por qué ser un LIC.
- La complementación de un LIC no tiene por qué ser un LIC.