Apuntes 24 de Mayo

COMPILADORES E INTÉRPRETES SERGIO MOYA VALERIN 2013015682

Contenido

Quiz 11	3
Repaso	3
Inicio de Lección	
Mensajes de Error	4
Ejemplo 0	
Ejemplo 1	
Recuperación de Error	8
Operaciones	9
Ejercicio 1	10
Ejercicio 2	11

Quiz 11

Considere las siguientes 2 gramáticas. Haga todo el proceso necesario para construir la tabla de parsing correspondiente a cada una.

- a) S -> aSdd | A
 - $A \rightarrow bAc \mid bc$
- b) S -> 1 | 1S1 | OSO

Repaso

El algoritmo detecta hileras invalidas - entrada vacía en la tabla LL(1)

Nombre	Tipo	Tamaño	# Parámetros	Tipo
X	Int	2		
γ	Int	2	<u> </u>	
Z	Real	4		
Arregla	Int	200		
Compara	Proced.	Nul	2	int

Entrega simplista:

- "Error de sintaxis aquí"
- No revisar el resto de la hilera

Mejor enfoque:

- Mensajes de error: dar el mejor diagnostico posible de la causa del error
- Recuperación: Seguir analizando para buscar errores posteriores

Distintos compiladores dan diferentes mensajes para el mismo error

Es particular para cada lenguaje

A veces se dan de manera incremental ERRORMSG(17)

Inicio de Lección

Mensajes de Error

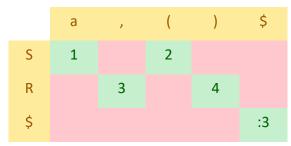
Ejemplo 0

Consideramos la gramática:

Genera S-expressions o sea hileras de la forma:

- a
- (a)
- (a,a,a)
- (a,(a,a))
- ((a,(a,a),(a,a)),a)

Y esta es su tabla de parsing:



Notación: Todos los espacios rojos son errores de sintaxis

Todos los mensajes podrían ser de la forma

"Aparece ___ cuando se esperaba ___"

- Calza bien para el caso del match
- Puede ser confuso para un No-Terminal y un Terminal

Analicemos las entradas vacías:

	а	,	()	\$
S	1	a	2	b	С
R	d	3	е	4	f
\$	g	h	i	j	:3

En los errores a y b se esperaba un "S" (S- expression) y esos inicios son inválidos

Se podría dar un mensaje de error más o menos así:

"Aparece "," cuando se esperaba S-espression"

"Aparece")" cuando se esperaba S-espression"

Las hileras que generan este error son de la forma

- (\$
- (a, \$
- \$

Se podrían dar mensajes como: "Aparece "," cuando se esperaba ")"", pero no hay seguridad de que esto sea así.

Un mensaje como este "Termina la hilera cuando se esperaba un no-terminal", esta correcto sin embargo los programadores no tienes por qué saber de compiladores (Nosotros sí).



Un mensaje que los mortales pueden entender seria "S-expression incompleta" (Si conoce el lenguaje por supuesto)

Ejemplo 1

Los errores d, e y f están asociados al no terminal R

R no tiene un significado evidente para los programadores, pero podríamos mencionar cosas que esperábamos cuando llego al token ofensivo:

- 'a' aparece cuando se esperaba ',' o ')'
- Fin de hilera aparece cuando se esperaba ',' o')'



Las hileras que típicamente causan los errores d y e son de la forma:

- (aa)\$
- ((a)a)\$
- (a(a))\$
- ((a)(a))\$

Un mejor mensaje de error para este caso sería:

• "Falta ','"

Las hileras que típicamente causan el error f son de la forma

- (a,a\$
- ((a),a\$
- (a\$

Un mejor mensaje de error para este caso sería:

• "Falta '('"

Las hileras que típicamente causan los errores g, h, i, j son las que se cree que ya el archivo debe terminar aún quedan símbolos

- (a)a\$
- ((a),a),a\$

Un mejor mensaje de error para estos casos seria

• "__ aparece después de fin de S-Expression"

Se puede construir una tabla de mensajes de error para cada situación, con esto se nos hace más fácil decidir que imprimir según cada error, para este caso podría quedar así.

a	Aparece ',' cuando se esperaba S-Espression
b	Aparece ',' cuando se esperaba S-Espression
С	S-Espression incompleta
d	Falta ','
r	Falta ','
f	Falta ')'
g	a' aparece después de fin de S-Expression
h	',' aparece después de fin de S-Expression
i	(' aparece después de fin de S-Expression
j)' aparece después de fin de S-Expression

Es difícil hacer mensajes de error por:

- Hay que correr muchas hileras para saber cuáles generan un determinado mensaje
- No hay una teoría general, es muy específico para cada lenguaje
- Es difícil saber en qué parte exactamente se encuentra hilera

Todos los mensajes de error deberían de empezar con:

• (Probablemente | Es posible | Tal vez | ...) exista un error en...

En eso el profe nos contó que hizo un sistema para la Unidad que se iba a usar en las votaciones, y se le ocurrió poner su número de teléfono en el programa. El día de las elecciones a Que pasa si una persona las 4am todo mundo lo llamo a preguntar tonteras. Y le quedaron debiendo una plata. Moraleja: No trabajar con políticos



Recuperación de Error

Reestablecer la pila y avanzar la entrada para poder seguir analizando

No es 100% preciso (Heurísticas), Ni gcc lo hace del todo bien

Caso extremo

- Dar un solo mensaje de error
- Baja productividad

Riesgo:

- Cascada de errores
- Dar mensajes equivocados

Programador revisa los 2 o 3 primeros

Para recuperar nos ayudamos con la tabla de errores y la tabla de parsing con los códigos de error y no hay ninguna fórmula o algoritmo, esto lo hacemos específicamente para cada gramática

Operaciones

POP: saca el No terminal del top de la pila

PUSH: inserta los elementos indicados en el TOP de la pila

REPLACE: reemplaza el TOP de la pila por lo elementos indicados

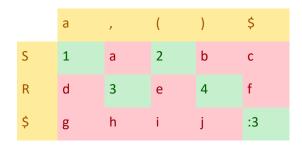
ADVANCE: avanza en la entrada

RETAIN: permanece en el mismo punto de entrada

EXIT: terminar parsing

Vamos a agregar estas operaciones a nuestra tabla de parsing para saber qué acciones tomar cuando nos encontramos un determinado error.

а	Aparece ',' cuando se esperaba S-Espression	POP RETAIN
b	Aparece ',' cuando se esperaba S-Espression	POP RETAIN
С	S-Espression incompleta	EXIT
d	Falta ','	REPLACE (RS) RETAIN
r	Falta ','	REPLACE (RS) RETAIN
f	Falta ')'	EXIT
g	a' aparece después de fin de S-Expression	PUSH (RS) RETAIN
h	',' aparece después de fin de S-Expression	PUSH (R) RETAIN
i	(' aparece después de fin de S-Expression	PUSH (RS) RETAIN
j)' aparece después de fin de S-Expression	ADVANCE



Ejercicio 1 Vamos a parsear la hilera (a,,a\$

Falta ")"

(a,,a\$\$S	(a,,a\$
\$ R S ((a,,a\$
\$ R S	a , , a \$
\$ R a	a , , a \$
\$ R	, , a \$
\$ R S ,	, , a \$
\$ R S	, a \$
Aparece "," cuando se	
esperaba S-expression	POP – RETAIN
\$ R	, a \$
\$ R S ,	, a \$
\$ R S	a \$
\$ R a	a \$
\$ R	\$



EXIT

Ejercicio 2

Parsearla hilera: ((a)a)a a \$	
\$ S	((a)a)aa\$
\$ R S (((a)a)aa\$
\$RS	(a)a)aa\$
\$ R R S ((a)a)aa\$
\$RRS	a)a)aa\$
\$RRa	a)a)aa\$
\$RR)a)aa\$
\$ R))a)aa\$ 📠
\$ R	a)aa\$
Falta ","	REPLACE (R S) RETAIN
\$RS	a)aa\$
\$ R a	a)aa\$
\$ R) a a \$
\$)) a a \$
\$	a a \$
Aparece "a" después de	PUSH (R S)
fin de S-expression	RETAIN
\$RS	a a \$
\$Ra	a a \$
\$ R	a\$
Falta ","	REPLACE (R S)
	RETAIN
\$RS	a\$
\$ R S	a \$
\$ R	\$
Falta ")"	EXIT



Nos sacamos el Bison de nuevo

Bueno muchachos a lo largo de este semestre hemos aprendido algo muy importante, HAY QUE TENER CUIDADO CON LO QUE DICES EN CLASE POR QUE SINO TE GANAS UN MEME O UN APODO.

- Atentamente su amigo Fortran 4chan Forchan