

La Ambigüedad en el Parsing

Definición y Ejemplos

Universidad de Cantabria

Outline

- 1 El Problema
- 2 Conceptos
- 3 Ambigüedad

El Problema

En nuestra búsqueda por encontrar la estructura exploraremos como elegir una derivación entre varias opciones. No solo queremos que la derivación sea “única” por sus propiedades, queremos que refleje “algo” del árbol de derivación.

El Problema

Nótese que hay dos elementos indeterministas, para dar la solución al problema de palabra:

- La producción que elegimos.
- La elección de sobre qué variable actuamos.

El Problema

La idea de introducir derivaciones “más a la izquierda” o “más a la derecha” consiste en tratar de reducir el ingrediente indeterminístico a la hora de seleccionar qué variable es la variable sobre la que vamos a actuar con nuestras producciones.

Derivaciones a la Izquierda y la Derecha

Definición (Derivaciones Leftmost)

Sea $G = (V, \Sigma, Q_0, P)$ una gramática libre de contexto. Sean $c, c' \in (V \cup \Sigma)^*$ dos formas sentenciales.

Diremos que c' se obtiene mediante derivación “más a la izquierda” (o leftmost) de c , si existen

$\omega \in \Sigma^*$, $A \in V$, $\alpha \in (V \cup \Sigma)^*$, y existe una producción $A \mapsto \beta$, con $\beta \in (V \cup \Sigma)^*$ tales que

$$c = \omega A \alpha, \quad c' = \omega \beta \alpha.$$

Denotaremos mediante $c \rightarrow_{lm}^G c'$.

Derivaciones a la Izquierda y la Derecha

Definición (Derivación Rightmost)

Sea $G = (V, \Sigma, Q_0, P)$ una gramática libre de contexto. Sean $c, c' \in (V \cup \Sigma)^*$ dos formas sentenciales.

Diremos que c' se obtiene mediante derivación “más a la derecha” (o rightmost) de c , si existen

$\omega \in (V \cup \Sigma)^*$, $A \in V$, $\alpha \in \Sigma^*$, y existe una producción $A \mapsto \beta$, con $\beta \in (V \cup \Sigma)^*$ tales que

$$c = \omega A \alpha, \quad c' = \omega \beta \alpha.$$

Denotaremos mediante $c \rightarrow_{rm}^G c'$.

Observaciones

Usualmente, y si no hay confusión, omitiremos el super-índice G .

Las derivaciones a la izquierda suelen ser las más naturales, pero las derivaciones más a la derecha tienen más interés práctico.

Cadenas de Derivaciones

Definición

Diremos que c' es deducible de c mediante derivaciones más a la izquierda (y lo denotaremos mediante $c \vdash_{lm}^G c'$) si existe una cadena finita de derivaciones más a la izquierda que va de c a c' . Esto es, si existen:

$$c = c_0 \xrightarrow{G}_{lm} c_1 \xrightarrow{G}_{lm} \cdots c_{k-1} \xrightarrow{G}_{lm} c_k = c'.$$

Cadenas de Derivaciones

Definición

Diremos que c' es deducible de c mediante derivaciones más a la derecha (y lo denotaremos mediante $c \vdash_{rm}^G c'$) si existe una cadena finita de derivaciones más a la derecha que va de c a c' . Esto es, si existen:

$$c = c_0 \rightarrow_{rm}^G c_1 \rightarrow_{rm}^G \cdots c_{k-1} \rightarrow_{rm}^G c_k = c'.$$

Ejemplos

Ejemplo

Tomemos la gramática cuyas producciones son:

$$P := \{Q_0 \mapsto AB \mid CA \mid AQ_0 \mid 0, \\ A \mapsto BA \mid 0A0 \mid 1, B \mapsto Q_0A, C \mapsto 1\}.$$

Una cadena de derivaciones leftmost (más la izquierda) sería la siguiente:

$$Q_0 \rightarrow AB \rightarrow CAB \rightarrow 1AB \rightarrow 11B \rightarrow 11Q_0A \rightarrow 110A \rightarrow 1101.$$

Una cadena de derivaciones rightmost (más a la derecha) sería la siguiente:

$$Q_0 \rightarrow AB \rightarrow AQ_0A \rightarrow AQ_01 \rightarrow A01 \rightarrow 0A001 \rightarrow 01001.$$

Grámaticas Ambiguas

Hemos dicho que la información de un programa esta contenida en la estructura. ¿Qué pasa si una palabra en el lenguaje tiene más de una estructura?

Gramáticas Ambiguas

Definición (Gramáticas Ambiguas)

Una gramática se dice ambigua si existe una forma sentencial $\omega \in (V \cup \Sigma)^$ alcanzable desde el símbolo inicial (i.e. $Q_0 \vdash_G \omega$) tal que existen al menos dos computaciones (derivaciones) más a la izquierda (o más a la derecha) distintas que permiten generar ω .*

Ejemplo

Ejemplo

Tomemos la gramática $P := \{E \mapsto E + E \mid E * E \mid a\}$. Ahora disponemos de dos cadenas de derivación para $a + a * a$ distintas:

$$E \rightarrow_{lm} E + E \rightarrow_{lm} a + E \rightarrow_{lm} a + E * E \rightarrow_{lm} a + a * E \rightarrow_{lm} a + a * a.$$

Y también

$$E \rightarrow_{lm} E * E \rightarrow_{lm} E + E * E \rightarrow_{lm} a + E * E \rightarrow_{lm} a + a * E \rightarrow_{lm} a + a * a.$$

Por lo que la anterior gramática es ambigua.

La Ambigüedad es Indecidable

Teorema

Decidir si una gramática libre de contexto es ambigua es indecible (i.e. no existe algoritmo que permita decidir la cualidad de ser ambigua).

Observaciones

- Hay muchas gramáticas que se generan por medio de máquinas (serialization)
- Hay lenguajes que solo admiten gramáticas ambiguas.
- Hay técnicas para eliminar algunas causas de la ambigüedad.