

Asignatura Bases Formales de la Computación

Evaluación de Gramáticas Estocásticas

Junio 6 de 2008

1. Indicaciones Generales

Este trabajo debe ser resuelto en forma individual o en grupos de máximo dos personas. Debe entregarse a más tardar el viernes 13 de junio mediante un correo electrónico enviado a la dirección de la profesora: galvarez@puj.edu.co

2. Descripción del Trabajo a Realizar

Considerar la siguiente gramática incontextual:

$O \rightarrow FS FV$
 $FS \rightarrow SUST$
 $FS \rightarrow PRON SUST$
 $FS \rightarrow ART SUS [ADJ]$
 $FS \rightarrow FS FP$
 $FV \rightarrow FV FS$
 $FV \rightarrow VERB [FS]$
 $FV \rightarrow FP V$
 $FP \rightarrow PREP FS$

Convenciones:

O	Oración
FS	Frase sustantivada
FV	Frase verbal
SUST	Sustantivo
PRON	Pronombre
ART	Artículo
ADJ	Adjetivo
FP	Frase preposicional
VERB	Verbo
PREP	Preposición
[]	Lo que está entre corchetes es opcional

2.1. Preguntas

1. Con respecto a la gramática anterior, realizar a mano las siguientes tareas:

- a) Asignar manualmente probabilidades a las producciones de manera que se convierta en una gramática estocástica correcta.
 - b) Calcular la probabilidad de la secuencia de observaciones *El policía detuvo el ladrón con el arma*
 - c) Encontrar el árbol sintáctico más probable para explicar esta secuencia
2. Describir con detalle tres problemas en cuya solución sea pertinente aplicar gramáticas estocásticas. Se debe explicar el problema y luego argumentar qué información contendrían las producciones y qué respuestas se obtendrían del modelo.

Por ejemplo: *El problema de analizar sintácticamente una oración en español para poder detectar el verbo que contiene y así conocer la acción que representa, se puede resolver mediante gramáticas estocásticas, para ello se generan producciones que modelan la estructura sintáctica del idioma, por ejemplo: $O \rightarrow FS FV$ significa que una oración está formada por una frase sustantivada seguida de una frase verbal. Esta gramática, una vez decorada con probabilidades que reflejan la frecuencia de ocurrencia de las diferentes construcciones en un corpus previamente construido de conversaciones entre estudiantes, permite construir la estructura de la oración y por lo tanto establecer cuál es la acción principal y cuáles las frases subordinadas y de ese modo extraer automáticamente la acción principal que contiene*