



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Máster en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital
Universitat Politècnica de València

Boletín de ejercicios (II)

Predicción Estructurada Estadística

Autor: Juan Antonio López Ramírez

Curso 2019-2020

Teoría

Ejercicio 3

En este ejercicio se pide computar, con el algoritmo Inside-Outside, la estimación de la regla:

$$Suj \rightarrow Art Nom Adj \quad (1)$$

Para el siguiente conjunto de muestras de entrenamiento:

$$D = \{(la\ vieja)(demanda\ ayuda), la\ mujer\ oculta\ pelea, la\ vieja\ ayuda\} \quad (2)$$

En cuanto a la gramática empleada, las reglas y sus probabilidades asociadas son las siguientes:

1,0	$S \rightarrow Suj Pre$
0,5	$Suj \rightarrow Art Nom$
0,3	$Suj \rightarrow Art Adj Nom$
0,2	$Suj \rightarrow Art Nom Adj$
0,3	$Pre \rightarrow Verb Nom$
0,7	$Pre \rightarrow Verb$
1,0	$Art \rightarrow 'la'$
0,1	$Nom \rightarrow 'vieja'$
0,2	$Nom \rightarrow 'ayuda'$
0,3	$Nom \rightarrow 'mujer'$
0,2	$Nom \rightarrow 'pelea'$
0,2	$Nom \rightarrow 'demanda'$
0,3	$Verb \rightarrow 'demanda'$
0,2	$Verb \rightarrow 'ayuda'$
0,1	$Verb \rightarrow 'oculta'$
0,4	$Verb \rightarrow 'pelea'$
0,3	$Adj \rightarrow 'vieja'$
0,7	$Adj \rightarrow 'oculta'$

Los árboles de derivación (ponderados) de cada muestra de entrenamiento se pueden observar en la figura 1.

Para calcular la estimación de la regla que hemos comentado al principio del ejercicio, hacemos uso de la fórmula del algoritmo Inside-Outside, que es:

$$p(Suj \rightarrow Art Nom Adj) = \frac{\sum_{x \in D} \frac{1}{P_\theta(x)} \sum_{t_x} N(Suj \rightarrow Art Nom Adj, t_x) P_\theta(x, t_x)}{\sum_{x \in D} \frac{1}{P_\theta(x)} \sum_{t_x} N(Suj, t_x) P_\theta(x, t_x)} \quad (3)$$

Por tanto, el resultado final de aplicar dicha fórmula es **0.3096**, tal y como se puede apreciar en la figura 2.

Ejercicio 5

De forma similar a como hemos realizado la estimación en el ejercicio anterior, aquí se pide hacer lo mismo, pero teniendo en cuenta el siguiente conjunto de muestras:

$$D = \{la\ vieja\ demanda\ ayuda, la\ mujer\ oculta\ pelea, la\ vieja\ mujer\ oculta\ demanda\ ayuda\} \quad (4)$$

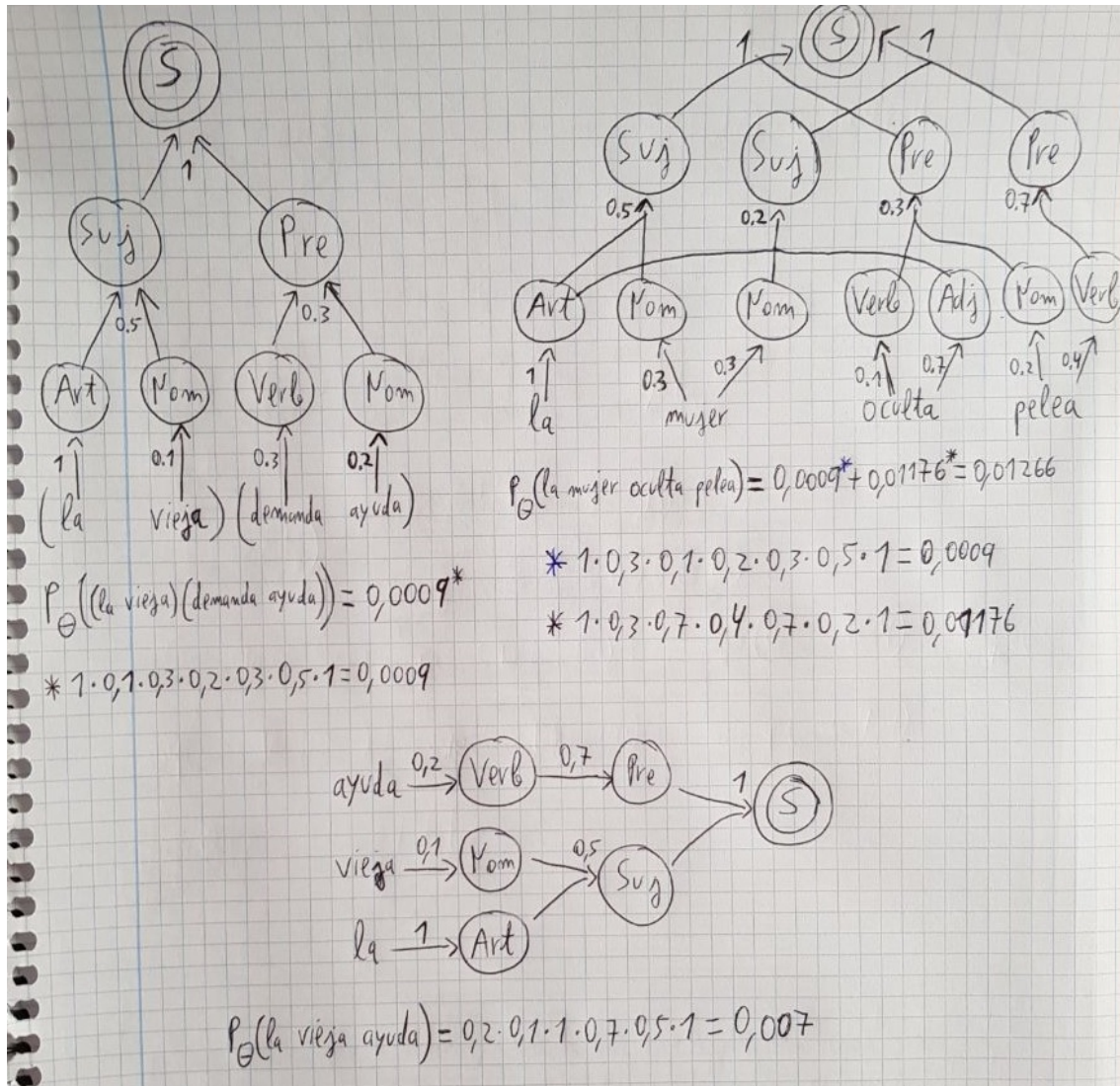


Figura 1: Árboles de cada muestra de entrenamiento del ejercicio 3.

$$\bar{P}(Suj \rightarrow Art\ Nom\ Adj) = \frac{\frac{1}{0,0009} \cdot (0,0009) + \frac{1}{0,01266} \cdot (0,0009 + 1 \cdot 0,01176) + \frac{1}{0,007} \cdot (0,007)}{\frac{1}{0,0009} \cdot (1 \cdot 0,0009) + \frac{1}{0,01266} \cdot (1 \cdot 0,0009 + 1 \cdot 0,01176) + \frac{1}{0,007} \cdot (1 \cdot 0,007)}$$

$$= \frac{\frac{1}{0,01266} \cdot 0,01176}{3} = 0,3096$$

Figura 2: Aplicación de la estimación del algoritmo Inside-Outside en el ejercicio 3 para la regla $Suj \rightarrow Art\ Nom\ Adj$

Como se puede observar, la última muestra no pertenece al lenguaje generado por nuestra gramática. Por tanto, la hemos modificado, cambiando aquellas reglas que tienen el no terminal Suj a la izquierda y reasignando las probabilidades, de forma que quede tal que así:

- 0,4 Suj \rightarrow Art Nom
 0,3 Suj \rightarrow Art Adj Nom
 0,2 Suj \rightarrow Art Nom Adj
 0,1 Suj \rightarrow Art Adj Nom Adj

El resto de reglas de la gramática son las mismas que en el ejercicio 3. Los árboles de derivación (ponderados) de cada muestra de entrenamiento se pueden observar en la figura 3.

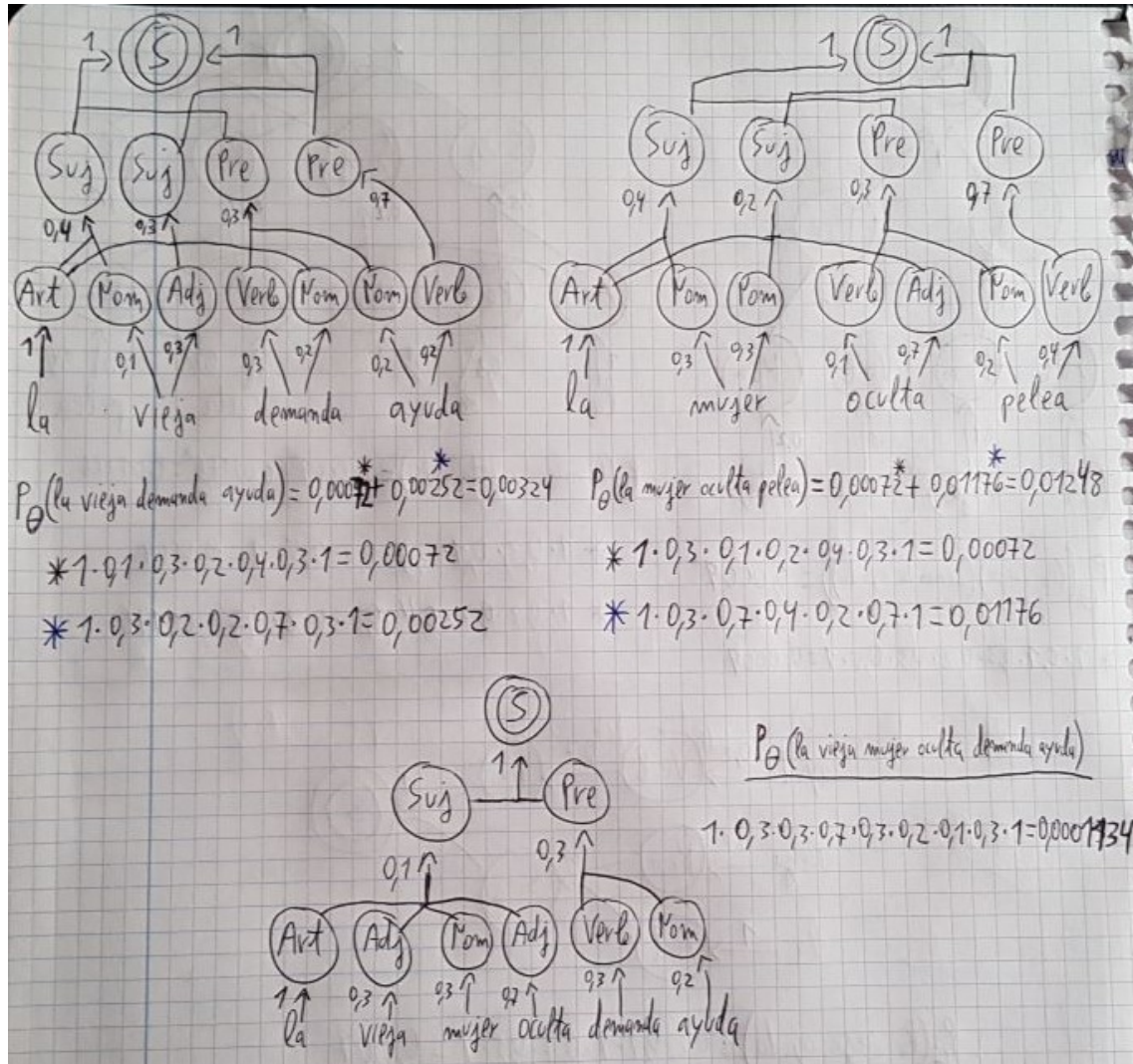


Figura 3: Árboles de cada muestra de entrenamiento del ejercicio 5.

Una vez hecho esto, calculamos la estimación de la regla con la ecuación del algoritmo de Inside-Outside explicada en el ejercicio 3, y obtenemos como resultado **0.3141**, tal y como se puede apreciar en la figura 4.

Ejercicio 6

Debido a que en el ejercicio 5 cada árbol que representa cada frase del conjunto de muestras tiene, como máximo, dos subárboles de derivación, el resultado de aplicar el algoritmo de estimación k -best con $k=2$ sería el mismo que en el ejercicio 5, es decir, **0.3141**.

$$\begin{aligned} \bar{P}(\text{Suj} \rightarrow \text{Art Nom Adj}) &= \frac{\frac{1}{0,00324}(0,00072 + 0,00252) + \frac{1}{0,01248}(0,00072 + 1,0,01176) + \frac{1}{0,0001134}(0,0001134)}{\frac{1}{0,00324}(1,0,00072 + 1,0,00252) + \frac{1}{0,01248}(1,0,00072 + 1,0,01176) + \frac{1}{0,0001134}(1,0,0001134)} \\ &= \frac{\frac{1}{0,01248} \cdot 0,01176}{3} = 0,3141 // \end{aligned}$$

Figura 4: Aplicación de la estimación del algoritmo Inside-Outside en el ejercicio 5 para la regla $\text{Suj} \rightarrow \text{Art Nom Adj}$

Prácticas

Ejercicio 8

Hemos cogido el conjunto de datos *SampleTriangle-10K*, que representa los triángulos rectangulares. Posteriormente, hemos entrenado y probado el PCFG con las siguientes instrucciones:

```
1 scfg-toolkit/scfg_cgr -g MODELS/G-triangle -n -f MODELS/Gt-k
2 scfg-toolkit/scfg_learn -g MODELS/Gt-k -f MODELS/Gmk-new-700 -i 700 -m DATA/
  SampleTriangle-10K
3 scfg-toolkit/scfg_gstr -g MODELS/Gmk-new-700 -c 1000 > tri-test-k
4 awk -f scfg-toolkit/checkTriangle tri-test-k | grep Y | wc -l
```

Donde *G-triangle-n* especifica cuántos nodos terminales y no terminales va a tener la gramática y sus nombres. Por tanto, tendremos 4 archivos, cada uno con *n* símbolos no terminales, para *n* = 5, 10, 15 y 20. *Gt-k* es la gramática que se genera con cada archivo, de forma que *Gt-1* es el modelo para 5 símbolos no terminales, *Gt-2* para 10, *Gt-3* para 15 y *Gt-4* para 20.

En las siguientes líneas entrenamos dichos gramáticas y obtenemos cuatro modelos, que son *Gm1-new-700*, *Gm2-new-700*, *Gm3-new-700*, *Gm4-new-700*. Luego entrenamos dichos modelos y los resultados (archivos *tri-test-1*, *tri-test-2*, *tri-test-3*, *tri-test-4*) se pueden apreciar en la siguiente tabla:

Símbolos no terminales	Triángulos rectángulos
5	29
10	63
15	61
20	84

Ahora trabajaremos con los tres tipos de triángulos: triángulos rectos, equiláteros e isósceles. Estos tipos de triángulos exhiben algunas relaciones que PCFG puede capturar. Los comandos utilizados en la figura 5 realizan un experimento simple para entrenamiento y pruebas, obteniendo un error del **41.13 %**, tal y como se explica en el enunciado del ejercicio.

Ejercicio 9

Al final del anterior ejercicio, hemos calculado el error usando Inside-Outside para muestras con brackets. En este ejercicio tendremos en cuenta muestras con y sin brackets,

```

jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_learn -g MODELS/g-0 -f MODELS/right -m DATA/Tr-right -i 100
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right -m DATA/Tr-right > r
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil -m DATA/Tr-right > e
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc -m DATA/Tr-right > i

jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($2>m) {m=$2;argm="equil";}}
-bash: syntax error near unexpected token `{'
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($2>m) {m=$2;argm="equil";}} if ($3>m) {m=$3;argm="isosc";}} pri
ntf("right %s\n",argm);}' > results
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right -m DATA/Tr-right > r
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil -m DATA/Tr-right > e
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc -m DATA/Tr-right > i
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$2;argm="equil"; if ($1>m) {m=$1;argm="right";}} if ($3>m) {m=$3;argm="isosc";}} pr
intf("equil %s\n",argm);}' > results
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right -m DATA/Tr-right > r
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil -m DATA/Tr-right > e
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc -m DATA/Tr-right > i
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$3;argm="isosc"; if ($1>m) {m=$1;argm="right";}} if ($2>m) {m=$2;argm="equil";}} pr
intf("isosc %s\n",argm);}' > results
jalopec@DESKTOP-HPIB76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ cat results | scfg-toolkit/confus
      equil isosc righ Err Err%
equil 794 206      0 206 20.6
isosc 531 225 244 775 77.5
righ 108 145 747 253 25.3

Error: 1234/3000 = 41.13%

```

Figura 5: Instrucciones empleadas para el entrenamiento y prueba de los modelos; y para el cálculo del error al clasificar mediante *confus*.

además de utilizar el algoritmo de Viterbi. De esta forma hemos calculado tres errores alternativos al 41.13 % del problema anterior.

Para empezar, modificamos los archivos Tr que se encuentran en la carpeta DATA, y le quitamos los brackets a las muestras, guardando los archivos en nuevos ficheros que hemos nombrado *Tr-right-nobrackets*, *Tr-equil-nobrackets*, *Tr-isosc-nobrackets*.

Posteriormente, ejecutamos las instrucciones del ejercicio anterior, solo que ahora los modelos están renombrados para que hagan referencia al algoritmo empleado (IO o Viterbi) y a si las muestras usadas en entrenamiento contenían brackets o no. Por ejemplo, para el modelo de equiláteros con Viterbi y sin brackets, su nombre es *equil-V-nobrackets*.

Los errores obtenidos se pueden observar en la siguiente tabla:

Algoritmo usado	Brackets	Error
Inside-Outside	Sí	41,13 %
Inside-Outside	No	36,20 %
Viterbi	Sí	58,70 %
Viterbi	No	58,27 %

A raíz de los resultados anteriores, podemos apreciar que el mejor algoritmo es el Inside-Outside, mientras que es mejor entrenar con las muestras sin brackets, debido a que la gramática genera un número más elevado de árboles de derivación que si las muestras tuviesen brackets. El problema es que, de esta forma, el entrenamiento de los modelos es más costoso desde el punto de vista temporal.

Las ejecuciones se pueden observar en las figuras 6, 7 y 8. Si en alguna de las imágenes falta alguna instrucción *scfg_learn* es debido a que se ejecutaron distintas versiones en terminales diferentes para optimizar el entrenamiento de los modelos, especialmente para los que utilizaban muestras sin brackets.

```

jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-IO-nobrackets -m DATA/Ts-right > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-IO-nobrackets -m DATA/Ts-right > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-IO-nobrackets -m DATA/Ts-right > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($2m) (m=$2;argm="equil"); if ($3m) (m=$3;argm="isosc");printf("right %s\n",argm);}' > results-IO-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-IO-nobrackets -m DATA/Ts-equil > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-IO-nobrackets -m DATA/Ts-equil > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-IO-nobrackets -m DATA/Ts-equil > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="equil"; if ($2m) (m=$2;argm="right"); if ($3m) (m=$3;argm="isosc"); printf("equil %s\n",argm);}' > results-IO-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-IO-nobrackets -m DATA/Ts-isosc > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-IO-nobrackets -m DATA/Ts-isosc > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-IO-nobrackets -m DATA/Ts-isosc > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="isosc"; if ($1m) (m=$1;argm="right"); if ($2m) (m=$2;argm="equil"); printf("isosc %s\n",argm);}' > results-IO-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ cat results-IO-nobrackets | scfg-toolkit/confus
equi 150 0 0 0 0
equi 783 217 0 217 21.7
isos 483 366 151 634 63.4
righ 48 187 765 235 23.5
Error: 1086/3000 = 36.20%

```

Figura 6: Instrucciones empleadas para el cálculo del error de muestras sin brackets mediante Inside-Outside.

```

jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_learn -v -g MODELS/G-0 -f MODELS/right-V-brackets -m DATA/Tr-right -i 100
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_learn -v -g MODELS/G-0 -f MODELS/equil-V-brackets -m DATA/Tr-right -i 100
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_learn -v -g MODELS/G-0 -f MODELS/isosc-V-brackets -m DATA/Tr-right -i 100
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-brackets -m DATA/Ts-right > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-brackets -m DATA/Ts-right > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-V-brackets -m DATA/Ts-right > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($2m) (m=$2;argm="equil"); if ($3m) (m=$3;argm="isosc");printf("right %s\n",argm);}' > results-5-V-brackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-brackets -m DATA/Ts-equil > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-brackets -m DATA/Ts-equil > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-V-brackets -m DATA/Ts-equil > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($2m) (m=$2;argm="equil"); if ($3m) (m=$3;argm="isosc"); printf("equil %s\n",argm);}' > results-5-V-brackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-brackets -m DATA/Ts-isosc > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-brackets -m DATA/Ts-isosc > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-V-brackets -m DATA/Ts-isosc > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="isosc"; if ($1m) (m=$1;argm="right"); if ($2m) (m=$2;argm="equil"); printf("isosc %s\n",argm);}' > results-5-V-brackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ cat results-V-brackets | scfg-toolkit/confus
equi 150 0 0 0 0
equi 77 843 88 923 92.3
isos 70 850 80 150 15.0
righ 12 676 312 688 68.8
Error: 1761/3000 = 58.70%

```

Figura 7: Instrucciones empleadas para el cálculo del error de muestras con brackets mediante Viterbi.

```

jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_learn -v -g MODELS/G-0 -f MODELS/right-V-nobrackets -m DATA/Tr-right-nobrackets -i 100
Error: File not found: DATA/Ts-right-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-nobrackets -m DATA/Ts-right-nobrackets > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-nobrackets -m DATA/Ts-right > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-nobrackets -m DATA/Ts-right > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-nobrackets -m DATA/Ts-right > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-nobrackets -m DATA/Ts-right > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-V-nobrackets -m DATA/Ts-right > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($2m) (m=$2;argm="equil"); if ($3m) (m=$3;argm="isosc");printf("right %s\n",argm);}' > results-V-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-nobrackets -m DATA/Ts-equil > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-nobrackets -m DATA/Ts-equil > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-V-nobrackets -m DATA/Ts-equil > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="right"; if ($1m) (m=$1;argm="right"); if ($2m) (m=$2;argm="equil"); if ($3m) (m=$3;argm="isosc"); printf("equil %s\n",argm);}' > results-V-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/right-V-nobrackets -m DATA/Ts-isosc > r
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/equil-V-nobrackets -m DATA/Ts-isosc > e
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ scfg-toolkit/scfg_prob -g MODELS/isosc-V-nobrackets -m DATA/Ts-isosc > i
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ paste r e i | awk '{m=$1;argm="isosc"; if ($1m) (m=$1;argm="right"); if ($2m) (m=$2;argm="equil"); printf("isosc %s\n",argm);}' > results-V-nobrackets
jalopez@DESKTOP-HIP1B76:/mnt/d/Cosas_de_Juan/PEE/Boletin II/PCFGs-19-20$ cat results-V-nobrackets | scfg-toolkit/confus
equi 150 0 0 0 0
equi 67 933 0 933 93.3
isos 171 612 217 388 38.8
righ 55 372 573 427 42.7
Error: 1748/3000 = 58.27%

```

Figura 8: Instrucciones empleadas para el cálculo del error de muestras sin brackets mediante Viterbi (no está el *scfg_learn* de los otros dos tipos de triángulo debido a que se ejecutó en máquinas distintas para paralelizar las ejecuciones).