

Introducción al Procesamiento de Lenguaje Natural Diciembre de 2011

Consideraciones generales

- i) La prueba es sin material escrito.
- ii) Escriba nombre y C.I. en todas las hojas.
- iii) Numere todas las hojas.
- iv) En la primera hoja, indique el total de hojas.
- v) Comience cada ejercicio en una hoja nueva.
- vi) Utilice las hojas de un solo lado.
- vii) Entregue los ejercicios en orden
- viii) El total de puntos es 40

Ejercicio 1 [4 puntos]

Defina Precision, Recall, Medida F y medida E y explique brevemente para que sirven. ¿Qué función cumple la variable β en la fórmula de la medida E ?

Ejercicio 2 [6 puntos]

Suponga que se desea traducir una oración f , de un lenguaje foráneo F , al lenguaje E .

1. Escriba la ecuación fundamental de la traducción automática estadística.
2. Para resolver esta ecuación, el método visto en el curso utiliza tres componentes. Enumere y describa brevemente dichos componentes, indicando qué aspecto de la ecuación resuelven.

Ejercicio 3 [9 puntos]

1. Se desea aplicar un modelo HMM para resolver la asignación de categorías léxicas a un texto. Del análisis del corpus de entrenamiento, se estiman las siguientes probabilidades:

Probabilidades de transición $P(\text{tag}_j | \text{tag}_i)$

	j:	0	1	2	3	4	5
i		</s>	PRON	NOM	DET	V	ADJ
0	<s>	0	0,05	0,2	0,4	0,1	0,1
1	PRON	0,1	0	0,01	0	0,06	0,01
2	NOM	0,1	0,02	0,1	0	0,4	0,2
3	DET	0	0	0,5	0	0	0,3
4	V	0,1	0,1	0,1	0,1	0,01	0
5	ADJ	0,1	0	0,6	0,1	0	0

Probabilidades de emisión diferentes de 0:

$P(a \mid \text{PRON}) = 0,1$
 $P(la \mid \text{NOM}) = 0,0002$
 $P(la \mid \text{DET}) = 0,4$
 $P(nota \mid \text{V}) = 0,2$
 $P(nota \mid \text{NOM}) = 0,02$
 $P(pálida \mid \text{ADJ}) = 0,03$

Aplique el algoritmo *forward* para calcular la probabilidad de la secuencia "la nota pálida"

2. Mencione una forma de estimar la probabilidad de ocurrencia de un n-grama, si no fue visto nunca en el corpus de entrenamiento.

Ejercicio 3 [11 puntos]

Considere la siguiente gramática, con componente semántica, similar a la vista en el curso:

$o \rightarrow gn \ gv$	$o.sem = gn.sem \ (gv.sem)$
$gv \rightarrow v \ gn$	$gv.sem = v.sem(gn.sem)$
$gv \rightarrow v$	$gv.sem = v.sem$
$gn \rightarrow det \ n'$	$gn.sem = det.sem(n'.sem)$
$gn \rightarrow npropio$	$gn.sem = npropio.sem$
$npropio \rightarrow Juan$	$npropio.sem = \lambda P.P(j)$
$npropio \rightarrow María$	$npropio.sem = \lambda P.P(m)$
$n' \rightarrow n \ adj$	$n'.sem = adj.sem(n.sem)$
$n' \rightarrow n$	$n'.sem = n.sem$
$n \rightarrow niño$	$n.sem = \lambda x.niño(x)$
$n \rightarrow mujer$	$n.sem = \lambda x.mujer(x)$
$adj \rightarrow joven$	$adj.sem = \lambda P.\lambda x.(P(x) \wedge joven(x))$
$v \rightarrow salta$	$v.sem = \lambda x.salta(x)$
$v \rightarrow quiere_a$	$v.sem = \lambda P.\lambda x.P(\lambda y.quiere(x,y))$
$det \rightarrow todos_los$	$det.sem = \lambda P.\lambda Q.\forall x(P(x) \rightarrow Q(x))$
$det \rightarrow un$	$det.sem = \lambda P.\lambda Q.\exists x(P(x) \wedge Q(x))$

Se desea incorporar construcciones simples con frases relativas. Una frase relativa es un complemento de nombre, con funcionamiento similar al de un adjetivo, pero con estructura interna de tipo oración (verbo conjugado y argumentos).

Se desea cubrir ejemplos como los siguientes:

1. *Un niño que salta quiere a María.* (frase relativa : *que salta*)

2. *Una mujer que comió todos los bombones desapareció.* (frase relativa : *que comió todos los bombones*)

Notar que en los ejemplos anteriores la frase relativa es “casi” una oración; lo sería si reemplazamos el pronombre relativo **que** por su antecedente (*un niño* en 1, *una mujer* en 2). En ambos casos **que** ocupa la posición sujeto.

Se pide:

1. Indicar un tipo para el pronombre relativo *que*.
2. Completar la gramática anterior con reglas que permitan dar cuenta de los ejemplos anteriores.
3. Realizar la derivación de la representación semántica del primer ejemplo anterior.
4. Considere el siguiente ejemplo:

El libro que leyó Juan desapareció.

¿Qué diferencia encuentra con la función del pronombre relativo en este caso y en el de los 2 ejemplos anteriores? Describa brevemente los problemas a resolver para incorporar a la gramática esta clase de frases relativas.