

## Introducción al Procesamiento de Lenguaje Natural Diciembre de 2015

<b>Consideraciones generales</b>
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 [6 puntos]

- i) Escriba la ecuación fundamental de la Traducción Automática Estadística (SMT), suponiendo que se quiere traducir texto del lenguaje F al lenguaje E.
- ii) Para resolver esta ecuación se utilizan tres componentes. Describa brevemente dichos componentes, indicando qué aspecto de la ecuación resuelve cada uno.

### Ejercicio 2 [12 puntos]

i) Identifique los constituyentes de las siguientes oraciones e indique sus funciones sintácticas y roles semánticos (donde corresponda).

- a) Juan come la manzana de mañana.
- b) La manzana es comida por Juan de mañana.

ii) Considere la siguiente gramática con anotaciones semánticas:

$o \rightarrow gn\ gv$	$o.sem = gn.sem(gv.sem)$
$gn \rightarrow nprop$	$gn.sem = nprop.sem$
$gn \rightarrow det\ nom$	$gn.sem = det.sem(nom.sem)$
$nom \rightarrow n$	$nom.sem = n.sem$
$nom \rightarrow n\ adj$	$nom.sem = adj.sem(n.sem)$
$gv \rightarrow v$	$gv.sem = v.sem$
$gv \rightarrow v\ gn$	$gv.sem = v.sem(gn.sem)$
$nprop \rightarrow Juan$	$nprop.sem = \lambda P . P(juan)$
$n \rightarrow perro$	$n.sem = \lambda x . perro(x)$
$n \rightarrow gato$	$n.sem = \lambda x . gato(x)$
$v \rightarrow ladra$	$v.sem = \lambda x . ladra(x)$
$v \rightarrow come$	$v.sem = \lambda P . \lambda y . P(\lambda x . come(y, x))$
$adj \rightarrow negro$	$adj.sem = \lambda P . \lambda x . P(x) \wedge negro(x)$
$det \rightarrow todo$	$det.sem = \lambda P . \lambda Q . \forall P(x) \rightarrow Q(x)$
$det \rightarrow un$	$det.sem = \lambda P . \lambda Q . \exists x P(x) \wedge Q(x)$

Dibuje el árbol sintáctico y derive la expresión lógica asociada a la oración: "Un perro negro ladra".

### Ejercicio 3 [12 puntos]

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

O	→	GV   GN GV
GN	→	Det Nom   Nom   GN GP
GV	→	V   V GN   V GN GP
GP	→	Prep GN
Det	→	una   un   el   la
Nom	→	Juan   sopa   leche   fideos   tomo
V	→	tomo   corro
Prep	→	con

1. Aplique el algoritmo Earley para la entrada “*tomo una sopa con fideos*”, usando la gramática G. ¿Qué salida devuelve el algoritmo?
2. Dibuje los árboles sintácticos posibles para la oración. ¿Es posible obtener esta información del resultado del algoritmo? Justifique brevemente.
3. ¿El algoritmo utilizado expresa de alguna forma preferencia por uno de los árboles? ¿Y si se utilizaran algoritmos de *parsing* probabilísticos? ¿Y si se agregara algún tipo de información semántica a los algoritmos? Justifique brevemente las respuestas.

### Ejercicio 4 [10 puntos]

1. ¿Qué quiere decir que los transductores de estado finito son cerrados bajo composición? ¿Por qué es esto una ventaja para el análisis morfológico? ¿Por qué la morfología de estado finito resuelve al mismo tiempo el problema del análisis y de la generación?
2. Describa el algoritmo Naïve Bayes para clasificación. ¿Cómo se estiman las probabilidades involucradas?
3. ¿Cuál es la diferencia entre un clasificador generativo y uno discriminativo? Nombre un clasificador *secuencial* generativo y otro discriminativo.
4. Explique brevemente en qué consiste el criterio td-idf para la asignación de pesos a los términos de indización en un sistema de recuperación de información