

Processament del Llenguatge Humà - PLH

Examen final

13 de Juny del 2024

Ejercicio 1. (3 puntos)

Dada la siguiente oración morfológicamente analizada,

I	saw	bats	yesterday
PRP	VBD	NNS	NN
NN	NN	JJ	ADV
		VBZ	

y el modelo HMM parcialmente representado por las siguientes matrices,

A	PRP	JJ	NN	NNS	VBZ	VBD	ADV
*	0.4	0.1		0.3			0.1
PRP					0.2	0.3	0.1
JJ			0.8	0.2			
NN			0.2		0.3	0.4	0.1
NNS			0.1			0.5	
VBZ		0.2	0.2	0.3			0.2
VBD		0.1	0.4	0.2			0.2
ADV		0.1	0.1	0.1	0.2	0.3	

B	I	saw	bats	yesterday
PRP	1			
JJ			0.2	
NN	0.1	0.4		0.1
NNS			0.5	
VBZ			0.3	
VBD		0.5		
ADV				1

- a) Aplica el algoritmo de Viterbi para obtener la mejor secuencia de etiquetas PoS. Aporta la tabla dinámica resultante con toda la información requerida para conseguir dicha secuencia.
- b) ¿Cuál es la secuencia de etiquetas PoS resultante y su probabilidad? La respuesta debe estar justificada mediante la información contenida en la tabla dinámica aportada en el apartado anterior. En caso contrario, la respuesta será considerada incorrecta.
- c) ¿La secuencia de etiquetas PoS resultante es la correcta? Justifica brevemente tu respuesta.

Ejercicio 2. (3 puntos)

Dada la siguiente oración y la secuencia de etiquetas PoS asociada a dicha oración:

John saw your brother playing with his glasses
 NNP VBD PRP\$ NN VBG IN PRP\$ NNS

NNP: nombre propio; VBD: verbo en pasado; PRP\$: pronombre posesivo; NN/NNS, nombre singular/plural; VBG: gerundio; IN: preposición

- a) Queremos aprender un modelo bigrama de CRFs para el reconocimiento de frases nominales. Diseña una plantilla de atributos correcta y útil para reconocer más de una frase nominal de la oración. Deriva, de ella, una función de atributos correcta y útil.
- b) Dibuja los árboles de constituyentes derivados de la siguiente PCFG para la oración. ¿Cuál es el mejor árbol? ¿Cuál es el resultado de aplicar el algoritmo CKY con esta misma gramática? Justifica brevemente tus respuestas.

S	→	NP VP	(1.0)	NNP	→	John	(1.0)
NP	→	PRP\$ NN	(0.5)	NN	→	brother	(1.0)
NP	→	PRP\$ NNS	(0.3)	NNS	→	glasses	(1.0)
NP	→	NNP	(0.1)	PRP\$	→	your	(0.6)
NP	→	NP AP	(0.1)	PRP\$	→	his	(0.4)
PP	→	IN NP	(1.0)	IN	→	with	(1.0)
AP	→	VBG PP	(1.0)	VBD	→	saw	(1.0)
VP	→	VBD NP	(0.4)	VBG	→	playing	(1.0)
VP	→	VP AP	(0.6)				

- c) Supón que el significado correcto de la oración es aquel en el que John es el propietario de las gafas. Supón, también, que queremos aprender un modelo de par de menciones para la resolución de correferencias de frases nominales, y que la oración es parte del corpus de entrenamiento. Enumera (1) las menciones, (2) las cadenas de correferentes, y (3) los *singletons* que deben ser anotados a mano en la oración, así como (4) los ejemplos positivos y negativos que se obtienen según la estrategia *best-first*.

Ejercicio 3. (3 puntos)

Teniendo en cuenta la siguiente oración analizada morfológicamente:

<i>John</i>	<i>saw</i>	<i>bats</i>	<i>ready</i>	<i>to</i>	<i>eat</i>	<i>food</i>
John+NNP	saw+NN	bat+NNS	ready+JJ	to+TO	eat+VB	food+NN
	saw+VB	bat+VB				
	see+VBD	bat+JJ				

NNP: nombre propio; VBD: verb en pasado; NN/NNS, nombre singular/plural; JJ: adjetivo; TO: partícula *to*; VB: verbo

y los siguientes *synsets* de WordNet:

Bat	S1: (n) bat, chiropteran (nocturnal mouselike mammal with forelimbs modified to form membranous wings) S2: (n) bat, at-bat ((baseball) a turn trying to get a hit) "he was at bat when it happened"; "he got four hits in four at-bats" S3: (v) bat (strike with, or as if with a baseball bat) "bat the ball" S4: (adj) balmy, bats (informal or slang terms for mentally irregular) "it used to drive my husband balmy"
Saw	S1: (n) proverb, adage, saw, byword (memorable saying embodying important fact of experience considered true by many people) S2: (n) saw (hand tool having a toothed blade for cutting) S3: (n) power saw, saw, sawing machine (a power tool for cutting wood) S4: (v) saw (cut with a saw) "saw wood for the fireplace"
See	S1: (n) see (the seat within a bishop's diocese where his cathedral is located) S2: (v) see (perceive by sight or have the power to perceive by sight) "He is blind—he cannot see"
Ready	S1: (adj) ready (made suitable and available for immediate use) "dinner is ready"
Eat	S1: (v) eat (take in solid food) "She was eating a banana"; "What did you eat for dinner last night?" S2: (v) corrode, eat, rust (cause to deteriorate due to the action of water, air, or an acid) "The acid corroded the metal"
Food	S1: (n) food, nutrient (any substance that can be metabolized by an animal to give energy and build tissue)

- a) Entre SVMs, CRFs y UKB, ¿cuáles de estos métodos son apropiados para aprender modelos para desambiguación de sentidos de palabras? Justifica brevemente tu respuesta.
- b) Considerando los métodos apropiados del apartado (a), ¿Cuántos modelos y qué conjuntos de categorías deben ser aprendidos para desambiguar los sentidos de las palabras de la oración? Justifica brevemente tus respuestas.
- c) Aplica el algoritmo original de Lesk para desambiguar el sentido de la palabra *bats* dentro de la oración. ¿Cuál es el sentido resultante? En caso necesario, ¿qué información extra sería útil para conseguir el sentido correcto? Justifica tus respuestas brevemente.

Ejercicio 4. (1 punto)

1. ¿A qué nos referimos cuando hablamos de embeddings con información de sub-palabras (subword embeddings)?
2. ¿Qué utilidades tienen este tipo de embeddings?