

# Examen de Traducción Automática

MIARFID, Universitat Politècnica de València, 13 de febrero de 2018

Apellidos:

Nombre:

## Cuestiones

Marca cada recuadro con una única opción de entre las dadas. Cada 3 errores restan una respuesta correcta. Las respuestas en blanco ni suman ni restan.

- 1 ☐ C En un modelo de traducción basado en alineamientos estadísticos, indicar qué afirmación es falsa al introducir el concepto de alineamiento ( $I$  es el tamaño de la traducción y  $J$  es el tamaño de la frase a traducir):
- A) Si los alineamientos se definen con una relación  $a \subseteq \{1, \dots, J\} \times \{1, \dots, I\}$ , el número de posibles alineamientos es excesivamente grande.
  - B) Si los alineamientos se definen con una función  $a : \{1, \dots, I\} \rightarrow \{1, \dots, J\}$ , el número de posibles alineamientos es inferior asintóticamente del que se obtiene si se define como una relación.
  - C) Si los alineamientos se definen con una función  $a : \{1, \dots, I\} \rightarrow \{1, \dots, J\}$ , el número de posibles alineamientos es lineal con  $I$ .
  - D) Si los alineamientos se definen con una función  $a : \{1, \dots, I\} \rightarrow \{1, \dots, J\}$ , el número de posibles alineamientos es exponencial con  $I$ .
- 2 ☐ B En la traducción estadística basada en palabras, indicar qué afirmación es incorrecta cuando se define el denominado modelo 1.
- A) La distribución de probabilidad de alinear una posición de la frase fuente con una posición de la traducción es equiprobable.
  - B) La distribución de probabilidad de las palabras de la frase destino sean una traducción de una palabra de la frase fuente dada es equiprobable.
  - C) En la aproximación por Viterbi con el modelo 1 el alineamiento entre palabras se obtiene buscando la palabra que maximice la probabilidad del diccionario estadístico.
  - D) Los únicos parámetros a estimar a partir de un corpus paralelo son los que definen el diccionario estadístico.
- 3 ☐ C En un sistema de traducción estadística basado en “phrases” y en un modelo log-lineal, indicar qué afirmación es correcta:
- A) Solo hay que estimar los pesos del modelo log-lineal
  - B) El número de segmentos en la frase fuente puede ser distinta del número de segmentos en la traducción.
  - C) En la construcción de segmentos bilingües de “phrases” la técnica más utilizada para obtener los alineamientos es la denominada “grow-diagonal”
  - D) La asignación de probabilidades a los segmentos bilingües de “phrases” es arbitraria.
- 4 ☐ A En el proceso de traducción con un sistema de traducción estadística basado en “phrases” y en un modelo log-lineal, indicar qué afirmación es correcta:
- A) La estrategia básica más utilizada está basada en el uso de múltiples colas de prioridad o “stacks”, una por cada talla del conjunto de palabras fuentes ya traducidas.
  - B) La estrategia básica más utilizada está basada en recorrer la frase fuente y generar la palabra de mayor probabilidad de cada palabra fuente.
  - C) La estrategia básica más utilizada está basada en el uso de múltiples colas de prioridad que son lineales en tamaño con la longitud de la frase destino.
  - D) La información almacenada en las colas de prioridad o “stacks” solo contiene las palabras que ya han sido traducidas,
- 5 ☐ A Un traductor neuronal “encoder-decoder” contiene un número elevado de parámetros que hay que estimar, indicar qué afirmación es falsa:
- A) En general, siempre hay que entrenar primero los “word-embeddings” de las palabras del lenguaje fuente y de las palabras del lenguaje destino y luego el resto de parámetros.
  - B) En general, todos los parámetros, incluidos los “word-embeddings” se estiman conjuntamente con el resto de parámetros.
  - C) En algunas partes del traductor hay que utilizar la función “softmax” para obtener distribuciones de probabilidad y su coste es elevado.
  - D) Los algoritmos de optimización más importante utilizados para el entrenamiento son: “Stochastic gradient descent with momentum”, “Adaptive Gradient”, “Adadelta” y “Adam”.
- 6 ☐ B El traductor neuronal “encoder-decoder” que mejores prestaciones está dando es el basado en un modelo de atención, indicar qué afirmación es falsa:

- A) El modelo de atención permite dar más peso a las palabras de la frase fuente para producir una palabra en la fases destino.
- B) El modelo de atención es un modelo de alineamiento semejante al modelo 1 de los traductores estadístico en el que se relaciona una palabra del lenguaje destino con una del lenguaje fuente.
- C) Un modelo bidireccional tienen dos redes recurrentes para analizar la frase fuente, una la analiza en un sentido y la otra en sentido contrario,
- D) La traducción se basa en una búsqueda en árbol con un mecanismo de poda de las hipótesis menos prometedoras.

7 **A** En un sistema de traducción interactiva indicar qué afirmación es correcta:

- A) La interactividad sirve para que el humano corrija un error y el sistema reaccione proporcionando una traducción mejor.
- B) La interactividad sirve para que el humano busque todos los errores y los corrija sin intervención del sistema.
- C) El problema de la búsqueda del mejor sufijo en cada interacción consiste en empezar siempre el proceso de búsqueda desde la primera palabra de la mejor traducción sin restricciones.
- D) Solo existe un protocolo de interacción: el basado en prefijos.

8 **D** La post-edición y la traducción interactiva ofrecen un marco muy adecuado para mejorar los modelos de traducción utilizados, indicar qué afirmación es errónea:

- A) Los modelos se adaptan una vez que el usuario ha acabado de corregir/validar el conjunto de muestras de test.
- B) En un sistema estadístico sería conveniente estimar todos los parámetros, tanto los pesos del modelo log-linear como los modelos que forman parte del modelo log-lineal.
- C) Los modelos se adaptan a las nuevas muestras producidas por el usuario mediante un proceso de aprendizaje “on-line”
- D) El usuario debe modificar los parámetros que han producido los errores.

9 **C** Dada la SITG  $G = \{(p/2, S \rightarrow [SS]), (1 - p, S \rightarrow \langle SS \rangle), (p/2, S \rightarrow a/x)\}$  indica la probabilidad de par  $(aa, xx)$ :

- A) No es posible calcular dicha probabilidad.
- B)  $p^3/8$
- C)  $(1 - p/2) p^2/4$
- D) La probabilidades de dicha SITG no están correctamente definidas.

10 **D** Dada la SITG  $G = \{(p/4, S \rightarrow [SS]), (p/2, S \rightarrow \langle SS \rangle), (p/8, S \rightarrow a/x), (p/8, S \rightarrow b/y)\}$  indica el par de cadenas cuya derivación es de máxima probabilidad:

- A)  $(aba, xyx)$
- B)  $(aba, xxy)$
- C)  $(aba, yyx)$
- D)  $(aba, yxx)$

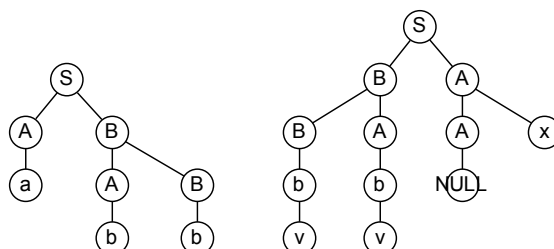
11 **A** La estimación de la probabilidades de una SITG a partir de una muestra (elige la respuesta correcta):

- A) Se puede realizar con el algoritmo inside-outside.
- B) Solo se puede realizar por la aproximación de Viterbi.
- C) Solo se puede realizar con el algoritmo inside-outside.
- D) No se puede realizar por la aproximación de Viterbi.

12 **D** Dadas las siguientes tablas de reordenamiento, inserción y traducción asociadas a un traductor “tree-to-string”:

<i>r</i> -table			<i>n</i> -table				<i>t</i> -table			
original order	reorder	P(reorder)	parent	S	A	B	w	p(ins-w)	a	b
A B	A B	0.7	node	A	A	A	x	0.6	z	0.6
	B A	0.3	P(None)	0.6	0.5	0.4	y	0.4	NULL	0.4
B A	A B	0.6	P(Left)	0.3	0.3	0.4				
	B A	0.4	P(right)	0.1	0.2	0.2			v	0.3
									NULL	0.7

indica la probabilidad de convertir el árbol de la izquierda en el árbol de la derecha:



- A) 0.0052
- B) 0.0004
- C) 0.0032
- D) 0.000324

- 13 ☒ B Dadas las tablas de reordenamiento, inserción y traducción asociadas al traductor “tree-to-string” de la pregunta anterior y el árbol de la izquierda, ¿es posible obtener la cadena “x”?:
- A) No es posible.
  - B) Si, y se puede obtener con diferentes árboles.
  - C) Habría que cambiar las probabilidades.
  - D) Si, y se puede obtener con un solo árbol.
- 14 ☒ D En un sistema de traducción jerárquica la reglas de traducción (elige la respuesta correcta):
- A) no contienen información léxica.
  - B) contienen solo unas pocas reglas y se escriben manualmente.
  - C) Se obtienen manualmente a partir de un corpus gigantesco.
  - D) Se obtienen a partir de los alineamiento producidos por GIZA.
- 15 ☒ B En la traducción jerárquica, el proceso de traducción se realiza (elige la respuesta correcta)
- A) de manera análoga a un decodificador basado en *phrases* como Moses.
  - B) de manera análoga a un análisis con el algoritmo CYK.
  - C) con el algoritmo de entrenamiento de GIZA.
  - D) solo a mano cuando la cadena supera longitud 10 en número de palabras.