

# Examen de Traducción Automática

MIARFID, Universitat Politècnica de València, 12 de febrero de 2021

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. Tiempo de duración del examen: 1 hora.

- 1 ☒ Supongamos que se dispone de un modelo de traducción estadístico basado en el modelo M1 con un vocabulario de 5 palabras fuente  $\mathcal{E} = \{E1, E2, E3, E4, E5\}$  y 4 palabras destino  $\mathcal{F} = \{F1, F2, F3, F4\}$  y que viene dado por la siguiente tabla de probabilidades  $p(y | x)$  donde  $y \in \mathcal{F}$  y  $x \in \mathcal{E}$ :

p	E1	E2	E3	E4	E5
F1	0.4	0.1	0.3	0.0	0.0
F2	0.3	0.2	0.2	0.0	0.7
F3	0.2	0.3	0.0	0.5	0.0
F4	0.1	0.4	0.5	0.5	0.3

Si se utiliza este modelo simplificado para traducir  $\mathbf{x} = (x_1, \dots, x_J)$  de acuerdo a la siguiente regla de inferencia  $y_i = \operatorname{argmax}_{y \in \mathcal{F}} p(y | x_i)$  para  $1 \leq i \leq |\mathbf{x}|$  ¿Qué frase destino  $\mathbf{y}$  se obtendrá al traducir  $\mathbf{x} = (E1, E5, E3)$  según el modelo M1?

- A)  $(F4, F3, F1)$
- B)  $(F4, F3)$
- C)  $(F1, F2, F4)$
- D)  $(F3, F1, F4, F1)$

- 2 ☒ En la traducción estadística basada en palabras, indicar qué afirmación es falsa:

- A) Los modelos M1, M2, HMM y Fast\_align tienen en común el diccionario estadístico.
- B) Los modelos M1, M2 y HMM están basados en tablas.
- C) La probabilidad de alineamiento en Fast\_align se calcula a partir de una función que depende de un parámetro.
- D) Los modelos M1, M2, HMM y Fast\_align son de orden 0. resto de modelos.

- 3 ☒ En la traducción estadística basada en modelos log-lineales indicar qué afirmación es falsa:

- A) Algunas características utilizadas están basadas en el logaritmo de una probabilidad definida sobre la oración de entrada y/o la de salida y algunas sobre el número de segmentos o el tamaño de la traducción.
- B) Las características basadas en modelos de segmentos o “phrases” lexicalizados se basan en alineamientos entre palabras dentro de las “phrases” y en un diccionario estadístico.
- C) En la traducción monótona, la característica basada en un reordenamiento de los segmentos no existe.
- D) Las características basadas en modelos de “phrases” lexicalizados se utilizan cuando no existan las características basadas en modelos de “phrases” inversos.

- 4 ☒ En el entrenamiento de los modelos estadísticos basados en segmentos o “phrases”, indicar qué afirmación es correcta:

- A) El algoritmo MERT permite aprender los pesos del modelo log-lineal optimizando el BLEU sobre un conjunto de validación.
- B) El modelo de lenguaje solo se puede aprender a partir del conjunto de datos bilingües que se utiliza para obtener los segmentos o “phrases”.
- C) Para extraer los segmentos o “phrases” hay que aplicar el algoritmo MERT.
- D) Los pesos del modelo log-lineal se estiman al mismo tiempo que el resto de parámetros.

- 5 ☒ En el proceso de inferencia o traducción con modelos basados en “phrases” mediante el algoritmo “stack decoding” indicar qué afirmación es correcta:

- A) En general, se van expandiendo los prefijos de la entrada.
- B) En cada iteración, se selecciona una hipótesis aleatoria de cada pila.
- C) En cada iteración se van expandiendo los prefijos de la entrada.
- D) En cada iteración se extrae la mejor hipótesis de cada pila y se expande la traducción en curso en base a las “phrases” que den cuenta de palabras de la oración de entrada que aún no han sido traducidas.

6 **A** En general y con respecto a la representación continua de las palabras (“word-embeddings”), indicar qué afirmación es correcta:

- A) Los mejores “word-embeddings” se obtienen a partir de datos utilizando técnicas basadas en redes neuronales.
- B) Los mejores “word-embeddings” se obtienen a partir del conocimiento lingüístico de las palabras.
- C) Los “word-embeddings” tienen un tamaño igual a la talla del vocabulario.
- D) Los “word-embeddings” tienen un tamaño que depende del número de caracteres de la palabra.

7 **A** En la traducción neuronal, qué afirmación es falsa cuando se utilizan LSTMs:

- A) La principal característica de una LSTM dada es que puede procesar las oraciones de izquierda-a-derecha y de derecha-a-izquierda a la vez.
- B) Los parámetros de los LSTMs se aprenden a partir de corpus mediante la técnica de descenso por gradiente.
- C) Las LSTMs permiten seleccionar de forma automática parte de la información pasada.
- D) Para procesar las oraciones de izquierda-a-derecha y de derecha-a-izquierda es necesario utilizar dos LSTMs como mínimo, una en cada sentido.

8 **D** En el entrenamiento de los modelos recurrentes basados en modelos de atención, indicar qué afirmación es falsa:

- A) Hay que fijar los tamaños de los “embeddings” de las palabras del language fuente y del language destino.
- B) Hay que fijar los tamaños de las LSTMs/GRUs del codificador y del decodificador.
- C) Los algoritmo de optimización utilizados pueden ser el Adagrad, Adadelta y Adam entre otros.
- D) El factor de aprendizaje hay que fijarlo desde el principio y no se reajusta nunca.

9 **D** En modelos neuronales basado en LSTMs y modelos de atención, indicar qué afirmación es correcta:

- A) Cada palabra de salida se genera en función solo de la anterior generada y de una palabra de entrada.
- B) Cada palabra de salida se genera en función solo de las palabras traducidas hasta ese momento.
- C) Cada palabra de salida se genera en función solo de la palabras de entrada afectadas por un peso.
- D) Cada palabra de salida se genera a partir de la anterior palabra traducida y de una combinación ponderadas de representaciones continuas de las palabras de entrada.

10 **B** Las traducciones con modelos de traducción neuronal están basadas en la optimización  $\hat{y} = \operatorname{argmax}_{y=y_1 \dots y_I} \prod_{i=1}^I p(y_i | y_1^{i-1}, x, W)$  donde  $y$  es la traducción,  $y_i$  la  $i$ -ésima palabra traducida,  $x$  la frase a traducir y  $W$  el conjunto de parámetros. Indicar qué afirmación es correcta:

- A) El mejor algoritmo de inferencia o búsqueda consiste en  $\hat{y}_i = \operatorname{argmax}_{y_i} p(y_i | y_1^{i-1}, x, W)$
- B) La búsqueda de la mejor traducción  $\hat{y}$  tiene un coste computacionalmente prohibitivo y hay que recurrir a técnicas heurísticas como “beam search” para redir ese coste.
- C) La mejor estrategia de búsqueda de la mejor traducción  $\hat{y}$  está basada en una estrategia voraz.
- D) El mejor algoritmo de inferencia o búsqueda consiste en  $\hat{y}_i = \operatorname{argmax}_{y_i} p(y_i | x, W)$

11 **A** Un sistema de traducción interactiva reacciona ante la realimentación humana generando una nueva traducción. Indicar qué afirmación es falsa:

- A) El sistema de traducción siempre genera una traducción completa de mejor calidad que la que había generado previamente.
- B) En los sistemas basado en prefijos, el sistema de traducción devuelve el sufijo de mayor probabilidad compatible con el prefijo validado por el usuario y por la palabra traducida corregida.
- C) El sistema de traducción puede devolver una traducción de peor calidad que la previamente generada.
- D) El sistema de traducción puede devolver una traducción de mayor TER que la previamente generada.

12 **B** La traducción interactiva basada en prefijos se define mediante el proceso de optimización  $\hat{y}_s = \operatorname{argmax}_{y_s} p(y_s | x, y_p)$ , donde  $y_s, y_p$  son el sufijo y el prefijo de la traducción y  $x$  la oración a traducir, indicar que afirmación es la correcta.

- A) El problema de optimización es una búsqueda en el espacio de las traducciones completas.
- B) El problema de optimización se reduce a una búsqueda en el espacio de los sufijos  $y_s$  compatibles con el prefijo  $y_p$ .
- C) El problema de optimización se reduce a una búsqueda en el espacio de los sufijos  $y_s$  compatibles o no con el prefijo  $y_p$ .
- D) El problema de optimización se reduce a una búsqueda en el espacio de los prefijos  $y_p$ .

13 ☒ **A** La acción humana en un proceso de traducción consistente en realizar post-edición o traducción interactiva permite obtener traducciones de alta calidad. Indicar qué afirmación de las siguientes es falsa:

- A) La post-edición y la traducción interactiva generan datos que permiten re-entrenar los modelos de traducción.
- B) En la post-edición el sistema de traducción solo genera la traducción inicial para cada oración de entrada mientras que la traducción interactiva genera una nueva traducción cada vez que el humano introduce una corrección.
- C) La traducción interactiva permite obtener muestras bilingües para re-entrenar los modelos de traducción, pero la post-edición no.
- D) Tanto los modelos de traducción neuronal como los puramente estadísticos pueden ser utilizados tanto en la post-edición como en la traducción interactiva.

14 ☒ **D** Dada la SITG  $G = \{(p/3, S \rightarrow [SS]), (p/3, S \rightarrow \langle SS \rangle), (p/3, S \rightarrow a/x)\}$  indica la probabilidad del par  $(aaa, xxx)$ :

- A)  $p^3/8$
- B)  $30p/3^5$
- C)  $40p/3$
- D)  $8p^5/3^5$

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

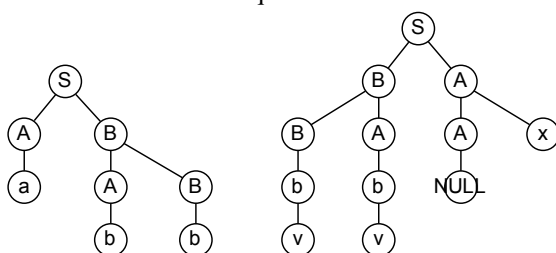
<i>r</i> -table		
original order	reorder	P(reorder)
A B	A B	0.7
	B A	0.3
B A	A B	0.6
	B A	0.4

<i>n</i> -table			
parent	S	A	B
node	A	A	A
P(None)	0.6	0.5	0.4
P(Left)	0.3	0.3	0.4
P(right)	0.1	0.2	0.2

w	p(ins-w)
x	0.6
y	0.4

<i>t</i> -table			
a		b	
z	0.6	v	0.3
NULL	0.4	NULL	0.7

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



- A)
- B)
- C)
- D)

**¡ANULADA!**

16 ☒ **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 “zzzz”?:

- A) Habría que cambiar la probabilidades.
- B) No es posible.
- C) Si es posible, y se puede obtener con varios árboles.
- D) Si, y se puede obtener con un solo árbol.