



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Máster en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital
Universitat Politècnica de València

Traducción basada en frases con MOSES

Traducción Automática

Autor: Juan Antonio López Ramírez

Curso 2019-2020

Introducción

El objetivo de esta práctica era construir sistemas de traducción a partir de conjuntos de pares de frases bilingües. Para ello hemos usado la herramienta Moses, un motor empleado para entrenar modelos estadísticos de traducción de textos desde un idioma de entrada a otro de salida.

Experimentos realizados y resultados obtenidos

Ejercicio 1

Primero, se nos pedía entrenar y probar el modelo de forma similar a como lo habíamos hecho en la tarea que realizamos en la sesión de prácticas, pero sin el ajuste de pesos log lineal. El BLEU obtenido en este caso ha sido de **87.96**. Para ello, hemos creado una carpeta en el directorio *TA* llamada *Ejer1* y las líneas de ejecución utilizadas han sido las mismas que en el boletín de prácticas, a excepción del apartado 7, y para el proceso de traducción hemos modificado la instrucción siguiente para utilizar el *moses.ini* en lugar de *moses.turista.ini*:

```
1 ~/TA/Practical1/Ejer1> clean-corpus-n.perl test es en test.clean 1 60
2 ~/TA/Practical1/Ejer1/test> $MOSES/bin/moses -f ../train/work/model/moses.ini <
  test.clean.es > test.hyp
```

Ejercicio 2

Para este ejercicio, se ha analizado la evolución del BLEU cuando aumentamos el número máximo de iteraciones del MERT cuando hacemos un ajuste de pesos log lineal.

Los resultados obtenidos se pueden apreciar en la siguiente tabla:

Nº máximo de iteraciones	BLEU obtenido
5	92.08
6	91.83
7	92.17
8	91.75

Tal y como se puede apreciar en la tabla, realizar un ajuste de los pesos permite obtener un incremento en la precisión del modelo de traducción estadística. Cuando usamos 7 iteraciones, para este corpus, es cuando obtenemos el BLEU más alto, de **92.17**, aunque el resto de resultados no varían mucho.

Ejercicio 3

La prioridad del ejercicio es determinar si, al variar los n-gramas en el modelo de lenguaje, mejoramos los resultados obtenidos por el modelo de traducción estadística. De este modo, los resultados obtenidos han sido los que se pueden observar en la tabla **1**.

A la vista de los resultados obtenidos en la tabla anterior, se puede apreciar que el modelo de lenguaje es vital a la hora de entrenar el modelo de traducción estadística, pasando de un valor de mínimo de 91.32 con bigramas a uno máximo de **92.57** con cuatrigramas. Que estos últimos sean los que obtienen mejores resultados puede deberse

n-grama	BLEU obtenido
2	91.32
3	92.08
4	92.57
5	91.50

Tabla 1: Resultados al variar n-gramas.

a que el corpus de entrenamiento es lo suficientemente grande para entrenar hasta n-gramas de orden 4, ya que para pentagramas, como se puede observar en la tabla, los resultados empeoran.

Ejercicio 4

En esta parte se pedía probar MIRA para el ajuste de pesos del modelo log-lineal.

Para ello, hay que ejecutar la siguiente instrucción durante el entrenamiento de los pesos:

```
$MOSES/scripts/training/mert-moses.pl --batch-mira train/development.clean.es
train/development.clean.en $MOSES/bin/moses train/work/model/moses.ini --
maximum-iterations=5 --mertdir $MOSES/bin/
```

De esta forma, conseguimos un BLEU de **90.77**, un valor bastante más bajo que usando MERT, pero converge de manera más rápida, obteniendo una solución en menos tiempo.

Ejercicio 5

El objetivo de este ejercicio es comprobar resultados que se obtienen al utilizar distintos tipos de suavizado y ver como varía el BLEU al aplicarlos sobre el modelo de lenguaje. Los resultados obtenidos han sido:

	Backoff	Interpolación
Witten-bell	91.96	91.54
Kneser-Ney	92.05	92.08

A la vista de los resultados que muestra la tabla superior, podemos observar que en nuestro caso el mejor BLEU se consigue con el suavizado de interpolación y descuento Witten-bell (es decir, la configuración original), aunque es necesario destacar que la variación en el BLEU no es significativa.

Ejercicio 6

El último ejercicio consiste en experimentar con moses monótono. El BLEU obtenido con esta configuración ha sido de **91.8**. Aunque este modelo ni siquiera alcanza a su homólogo en MOSES normal, que era de 92.08 para trigramas, 5 iteraciones máximas en el entrenamiento de los pesos del modelo log-lineal, suavizado por interpolación y descuento Kneser-Ney, se sitúa cercano a este.