

Trabajo Traducción Automáticas

Marcos Esteve Casademunt

Febrero 2020

Contents

1	Primer ejercicio	2
2	Segundo Ejercicio	3

Se ha optado por realizar el ejercicio 8 del tema Stochastic inversion transduction grammars y el ejercicio 1 del tema Tree to string transducers

1 Primer ejercicio

(***) Write a SITG and a two pairs of different source and target strings and use the idea described in [Sánchez 06] to obtain the set of possible word phrases.

Para el desarrollo de este ejercicio se ha optado por los pares de oraciones:

- El compró un ferrari rojo / He bought a red ferrari
- Tengo un ferrari nuevo / I have a new ferrari

Gramática:

$S \rightarrow Suj \quad Pred$

$Suj \rightarrow El/He$

$Suj \rightarrow \$/I$

$Pred \rightarrow SV \quad SN$

$SV \rightarrow Compró/Bought$

$SV \rightarrow Tengo/Have$

$SN \rightarrow Art \quad C$

$Art \rightarrow un/a$

$C \rightarrow Sust \quad Adj$

$Sust \rightarrow Ferrari/Ferrari$

$Adj \rightarrow nuevo/new$

Para el primer par de oraciones tenemos el siguiente árbol de derivación:

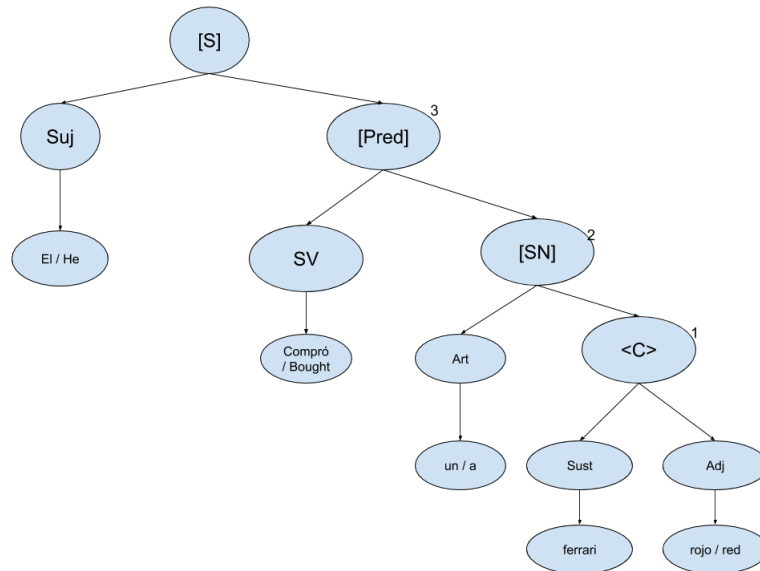


Figure 1: árbol de derivación para el primer par de frases

Siguiendo el procedimiento descrito en el paper obtenemos las siguientes word-phrases:

- 1 - rojo / red
- 1 - ferrari / ferrari
- 2 - ferrari rojo / red ferrari
- 2 - un / a
- 3 - un ferrari rojo / a red ferrari
- 3 - compró / bough

Siguiendo el mismo procedimiento para la segunda muestra obtenemos el siguiente árbol:

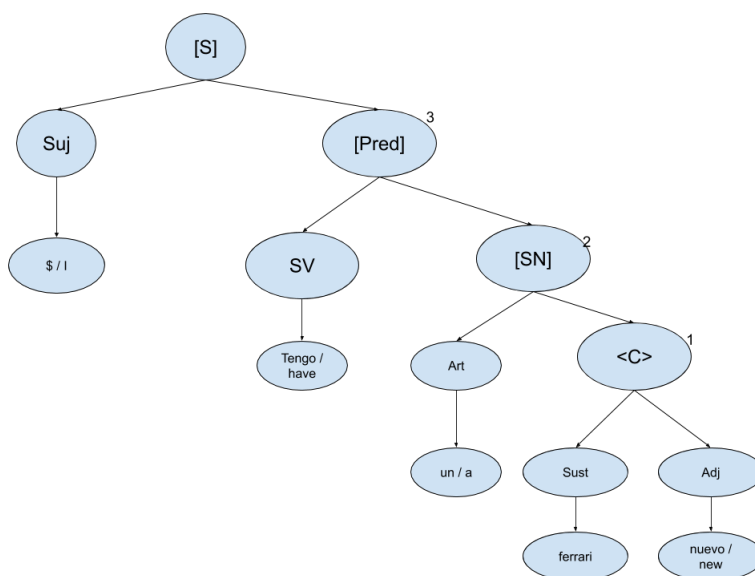


Figure 2: árbol de derivación para segundo par de frases

Como podemos observar el sujeto está omitido en el castellano por esta razón será necesario realizar una inserción. Siguiendo el mismo procedimiento para extraer las word-phrases obtenemos:

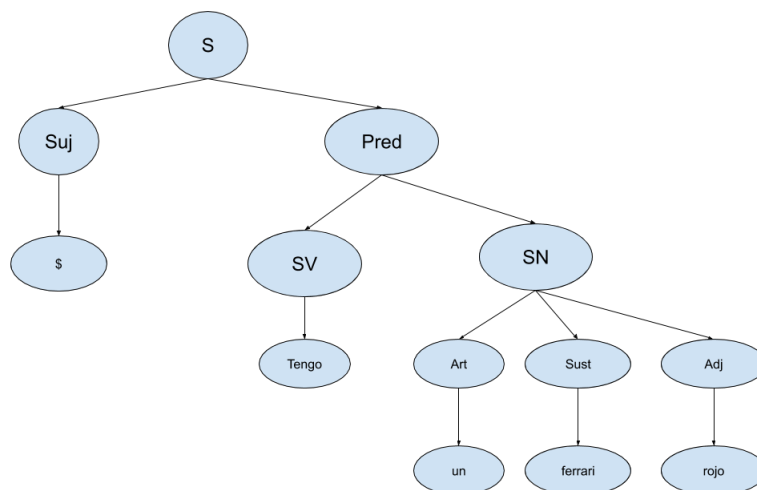
- 1 - nuevo / new
- 1 - ferrari / ferrari
- 2 - ferrari nuevo / new ferrari
- 2 - un / a
- 3 - un ferrari nuevo / a new ferrari
- 3 - tengo / have

2 Segundo Ejercicio

(*) Write an example similar to the slides in the pages 4–7 with the pair of languages that you prefer. Define the probabilities as you like.

Para la realización de este ejercicio se ha optado por el par de oraciones
Tengo un ferrari rojo / I have a red ferrari.

Partimos del árbol de derivación:



La tabla de reordenación asociada a este árbol es:

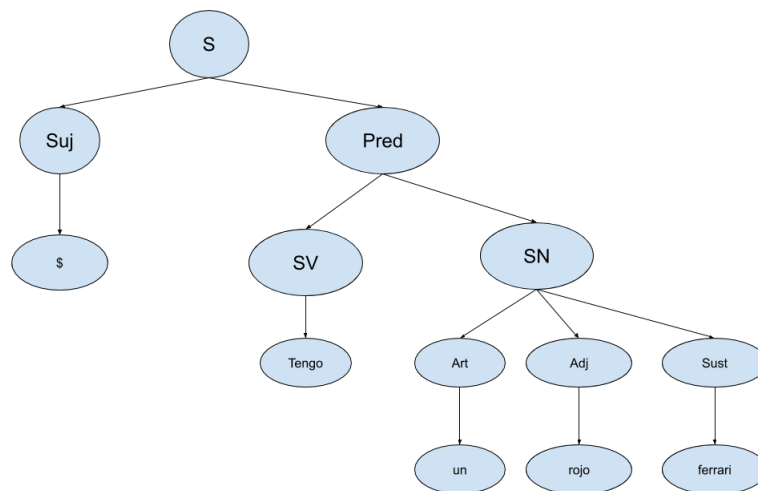
Origen	reordering	P(reorder)
Suj Pred	Suj Pred	0.99
	Pred Suj	0.01
SV SN	SV SN	0.66
	SN SV	0.33
ART SUST ADJ	ART SUST ADJ	0.31
	ART ADJ SUST	0.52

	SUST ART ADJ	0.01

La probabilidad asociada al reordenamiento es:

$$P = 0.99 * 0.66 * 0.52$$

Una vez definido el mejor reordenamiento posible el árbol se nos quedaría como:



Las tablas de inserciones quedarían como:

Padre	Top	S	S	Pred	Pred	SN	SN	SN
Nodo	S	Suj	Pred	SV	SN	Art	Adj	Sust
P(None)	0.8	0.02	0.7	0.8	0.8	0.9	0.9	0.9
P(Left)	0.1	0.49	0.2	0.15	0.1	0.05	0.05	0.05
P(Right)	0.1	0.49	0.1	0.05	0.1	0.05	0.05	0.05

w	P(w)
I	0.3
have	0.05
...	...
ferrari	0.05

$$P_{\text{insercion}} = 0.8 * (0.49 * 0.3) * 0.7 * 0.8 * 0.8 * 0.9 * 0.9 * 0.9$$

Quedando el árbol como:

Tengo	un	ferrari	rojo
have 0.9	a 0.8	ferrari 1.0	red 0.9
hold 0.1	one 0.2				...

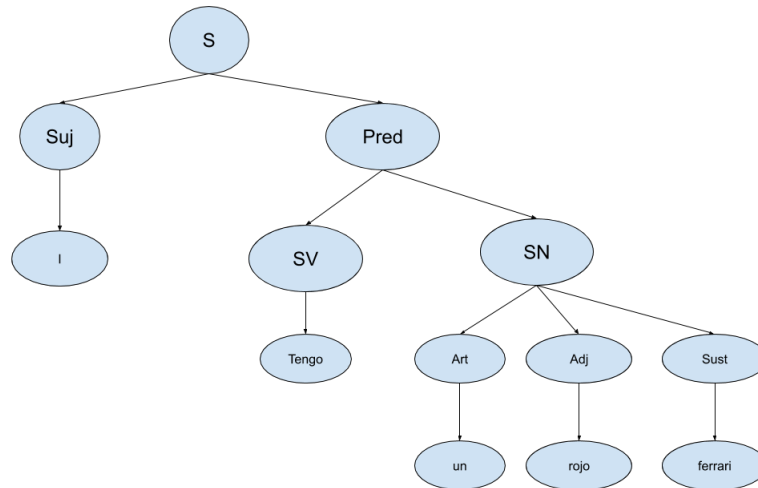


Figure 3: árbol de derivación para segundo par de frases

Por último tendríamos la tabla de traducción: La probabilidad de traducción asociada quedaría como:

$$P(traduccion) = 0.9 * 0.8 * 1 * 0.9$$

y el árbol como:

