

- f une phrase source (ex : français)
- e une phrase cible (ex : anglais)
- traduire \Leftrightarrow résoudre :

$$\operatorname{argmax}_e P(e|f) = \operatorname{argmax}_e P(f|e)P(e)$$

Deux modèles (au moins)

- $p(f|e)$ modèle de transfert
- $p(e)$ modèle de langue

Un décodeur

- NP-complet [***Knight, 1999***]

○ un **corpus parallèle**

<ul style="list-style-type: none">● The Legislative Assembly convened at 3.30 pm.● Mr. Quirke (Clerk-Designate) :● THURSDAY, APRIL 1, 1999	<ul style="list-style-type: none">● sitamiq, ipuru 1, 1999● maligaliurvik matuiqtau-lauqtuq 3 :30mi unnusakkut● mista kuak (titiraqti - tik-kuaqtausimajuq) :
--	---

- un **corpus parallèle** + un *aligneur de phrases* = **bitexte**

<ul style="list-style-type: none">● The Legislative Assembly convened at 3.30 pm.● Mr. Quirke (Clerk-Designate) :● THURSDAY, APRIL 1, 1999	<ul style="list-style-type: none">● sitamiq, ipuru 1, 1999● maligaliurvik matuiqtau-lauqtuq 3 :30mi unnusakkut● mista kuak (titiraqti - tik-kuaqtausimajuq) :
---	--

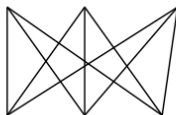
- un **corpus parallèle** + un *aligneur de phrases* = **bitexte**

<ul style="list-style-type: none">● The Legislative Assembly convened at 3.30 pm.● Mr. Quirke (Clerk-Designate) :● THURSDAY, APRIL 1, 1999	<ul style="list-style-type: none">● sitamiq, ipuru 1, 1999● maligaliurvik matuiqtalauqtuq 3 :30mi unnusakkut● mista kuak (titiraqti - tik-kuaqtausimajuq) :
---	---

- des aligneurs (de phrases) de bonne qualité
- des corpus parallèles pour plusieurs paires de langues (débat parlementaires, best sellers, documentations techniques, etc.)

- $P(f|e) = \sum_a P(a, f|e)$
où a est un *alignement* entre e and f
- plusieurs modèles de la jointe $P(a, f|e)$
- une procédure d'estimation itérative [**Brown et al., 1993**]

... la maison ... la maison blue ... la fleur ...



... the house ... the blue house ... the flower ...

alignements équiprobables

- $P(f|e) = \sum_a P(a, f|e)$
où a est un *alignement* entre e and f
- plusieurs modèles de la jointe $P(a, f|e)$
- une procédure d'estimation itérative [**Brown et al., 1993**]

... la maison ... la maison blue ... la fleur ...

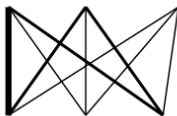


... the house ... the blue house ... the flower ...

la/the, maison/house émergent ...

- $P(f|e) = \sum_a P(a, f|e)$
où a est un *alignement* entre e and f
- plusieurs modèles de la jointe $P(a, f|e)$
- une procédure d'estimation itérative [**Brown et al., 1993**]

... la maison ... la maison blue ... la fleur ...



... the house ... the blue house ... the flower ...

la/the, maison/house se renforcent

- $P(f|e) = \sum_a P(a, f|e)$
où a est un *alignement* entre e and f
- plusieurs modèles de la jointe $P(a, f|e)$
- une procédure d'estimation itérative [**Brown et al., 1993**]

... la maison ... la maison bleu ... la fleur ...



... the house ... the blue house ... the flower ...

bleue/blue, fleur/flower apparaît

phrase based model

have not been democratically |||
 n' ont été démocratiquement ||| 0.5 9.81223e-08
 particularly in baby products . |||
 en particulier dans les produits pour bébés . ||| 0.5 4.17267e-12
 there is no cruelty |||
 il n' est nullement question de cruauté ||| 1 3.15409e-10
 intergovernmental conference and reform of the treaty |||
 cig et la révision du traité ||| 1 1.03844e-17
 is an absolute must . |||
 c' est un nécessité absolue . ||| 1 1.35065e-12
 parliament has already done a considerable amount |||
 le parlement a accompli un travail énorme ||| 1 8.38686e-21
 presidency 's programme |||
 du programme de la présidence ||| 1 9.33702e-07

- des millions de paramètres pour un bitexte d'un million de phrases

THIS BEAUTIFUL PLANT IS UNIQUE

1

1

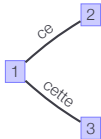
transfer table

this	↔	ce
	↔	cette
beautiful	↔	belle
	↔	beau
plant	↔	plante
	↔	usine
is	↔	est
unique	↔	seule
	↔	unique
beautiful plant		
↓		
belle plante		
plante magnifique		

language model

ce beau plante	:-(cette belle usine	:- belle usine est	:-) ...
----------------	--------------------------	------------------------	------------

THIS BEAUTIFUL PLANT IS UNIQUE

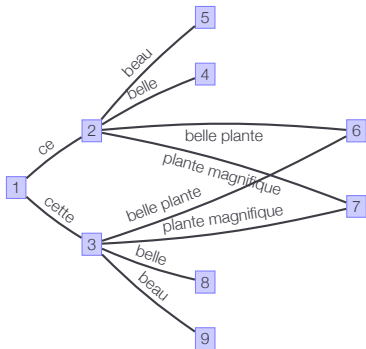


transfer table		
this	↔	ce
	↔	cette
beautiful	↔	belle
	↔	beau
plant	↔	plante
	↔	usine
is	↔	est
unique	↔	seule
	↔	unique
beautiful plant		
↓		
belle plante		
plante magnifique		

language model	
ce beau plante	:-(cette belle usine : belle usine est : ... :-)

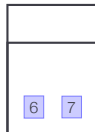


THIS BEAUTIFUL PLANT IS UNIQUE

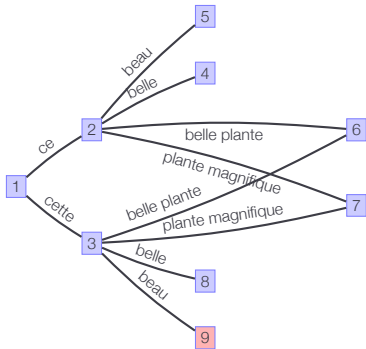


transfer table		
this	↔	ce
	↔	cette
beautiful	↔	belle
	↔	beau
plant	↔	plante
	↔	usine
is	↔	est
unique	↔	seule
	↔	unique
beautiful plant ↓ belle plante plante magnifique		

language model	
ce beau plante	:-
cette belle usine	:-
belle usine est	:-)
...	

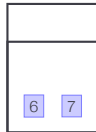


THIS BEAUTIFUL PLANT IS UNIQUE

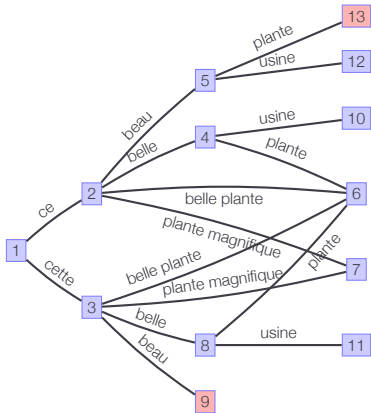


transfer table		
this	↔	ce
	↔	cette
beautiful	↔	belle
	↔	beau
plant	↔	plante
	↔	usine
is	↔	est
unique	↔	seule
	↔	unique
beautiful plant ↓ belle plante plante magnifique		

language model	
ce beau plante	:- (
cette belle usine	:-
belle usine est	:-)
...	

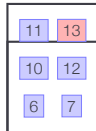


THIS BEAUTIFUL PLANT IS UNIQUE

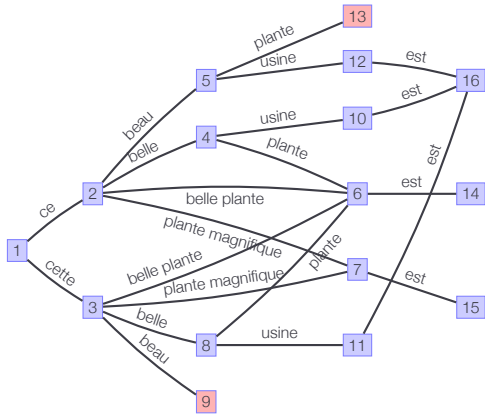


transfer table	
this	↔ ce
	↔ cette
beautiful	↔ belle
	↔ beau
plant	↔ plante
	↔ usine
is	↔ est
unique	↔ seule
	↔ unique
beautiful plant	
↓	
belle plante	
plante magnifique	

language model	
ce beau plante	:-
cette belle usine	:-
belle usine est	:-)
...	

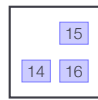
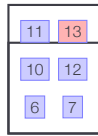
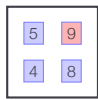
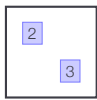


THIS BEAUTIFUL PLANT IS UNIQUE

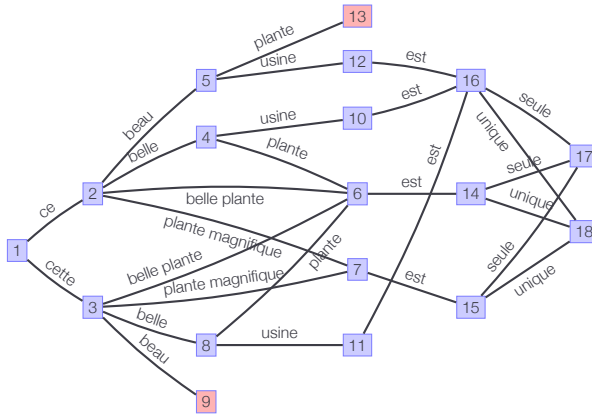


transfer table	
this	↔ ce
	↔ cette
beautiful	↔ belle
	↔ beau
plant	↔ plante
	↔ usine
is	↔ est
unique	↔ seule
	↔ unique
beautiful plant	
↓	
belle plante	
plante magnifique	

language model	
ce beau plante	:-
cette belle usine	:-
belle usine est	:-)
...	

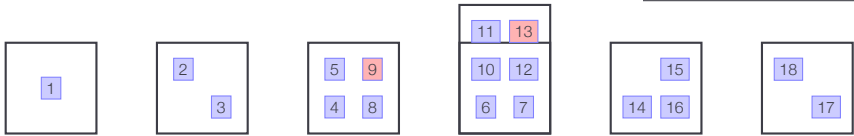


THIS BEAUTIFUL PLANT IS UNIQUE



transfer table	
this	↔ ce
	↔ cette
beautiful	↔ belle
	↔ beau
plant	↔ plante
	↔ usine
is	↔ est
unique	↔ seule
	↔ unique
beautiful plant ↓ belle plante plante magnifique	

language model	
ce beau plante	:-
cette belle usine	:-
belle usine est	:-)
...	



source

- Barack Obama becomes the fourth American president to receive the Nobel Peace Prize
- The American president Barack Obama will fly into Oslo, Norway for 26 hours to receive the Nobel Peace Prize, the fourth American president in history to do so.

traduction

- Barack Obama se convierte en el cuarto presidente estadounidense para recibir el Premio Nobel de la Paz
- El presidente estadounidense Barack Obama va a volar en Oslo, Noruega, por 26 horas para recibir el Premio Nobel de la Paz, el cuarto presidente estadounidense en la historia de hacerlo