

Master LITL - Programmation pour le TAL

Tests unitaires/tests d'intégration

Franck Sajous/CLLE-ERSS



<http://fsajous.free.fr/>

Étapes

- ① description du but/des fonctionnalités du système
(le quoi, pas le comment)
→ test global du système = quelles sorties du système en fonction de quelles entrées
- ② décomposition du système en sous-composants
(fonctionnalités de chaque composant)
→ test unitaire = pour chaque composant, quelles sorties du composant en fonction de quelles entrées
- ③ trouver un scénario d'échec du test d'intégration
(en supposant les tests unitaires validés)

Note : les tests (tests unitaires et tests globaux donnent des éléments de spécification / d'architecture)

Définitions

- Programme qui prend un texte en entrée et fournit pour chaque mot sa catégorie grammaticale, son lemme et les relations de dépendance syntaxique entre ces mots (*sujet de, objet de, etc.*)
- Test : entrée X \Rightarrow sortie Y

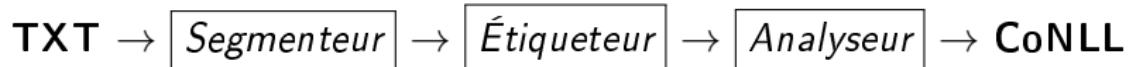
Exemple

- Entrée : « *Les poules du couvent couvent.* »
- Sortie :

1	Les	le	DET	2	det
2	poules	poule	NC	5	suj
3	du	de	P+D	2	dep
4	couvent	couvent	NC	3	obj
5	couvent	couver	V	0	root
6	.	.	PONCT	5	punct

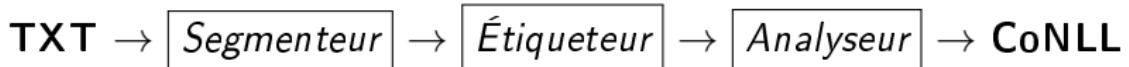
L'analyseur peut être composé de plusieurs modules :

- 1 un module qui segmente le texte d'entrée en « mots »
- 2 un module qui attribue une catégorie syntaxique et un lemme à chaque mot
- 3 un module qui établit les relations entre les mots



L'analyseur peut être composé de plusieurs modules :

- 1 un module qui segmente le texte d'entrée en « mots »
- 2 un module qui attribue une catégorie syntaxique et un lemme à chaque mot
- 3 un module qui établit les relations entre les mots



Formats d'entrée/sortie

1 Segmenteur

- Entrée : texte brut
- Sortie : texte un mot par ligne, numéroté

2 Étiqueteur

- Entrée : texte un mot par ligne numéroté
- Sortie : 3ème colonne = lemme, 4ème colonne = catégorie

3 Analyseur

- Entrée : texte un mot par ligne numéroté avec 3ème et 4ème colonnes = lemme + catégorie
- Sortie : idem avec 5ème colonne = n° du gouverneur syntaxique, 6ème colonne = relation syntaxique



Segmenteur

Entrée :

Les poules du couvent couvent

Sortie :

1	Les
2	poules
3	du
4	couvent
5	couvent
6	.

Segmenteur

Entrée :
Les poules du couvent couvent

Sortie :

1	Les
2	poules
3	du
4	couvent
5	couvent
6	.

Étiqueteur

Entrée :
1 Les
2 poules
3 du
4 couvent
5 couvent
6 .

Sortie :
1 Les le DET
2 poules poule NC
3 du de P+D
4 couvent couvent NC
5 couvent couver V
6 . PONCT

Segmenteur

Entrée :		Sortie :
1	Les	1 Les
2	poules	2 poules
3	du	3 du
4	couvent	4 couvent
5	couvent	5 couvent
6	.	6 .

Étiqueteur

Entrée :		Sortie :
1	Les	1 Les le DET
2	poules	2 poules poule NC
3	du	3 du de P+D
4	couvent	4 couvent couvent NC
5	couvent	5 couvent couver V
6	.	6 . PONCT

Analyseur

Entrée :			Sortie :				
1	Les	le	1	Les	le	DET	2 det
2	poules	poule	2	poules	poule	NC	5 suj
3	du	de	3	du	de	P+D	2 dep
4	couvent	couvent	4	couvent	couvent	NC	3 obj
5	couvent	couver	5	couvent	couver	V	0 root
6	.	.	6	.	.	PONCT	5 punct

Segmenteur

Entrée :		Sortie :	
1	Les	1	Les
2	poules	2	poules
3	du	3	du
4	couvent	4	couvent
5	couvent	5	couvent
6	.	6	.

Étiqueteur

Entrée :		Sortie :	
1	Les	1	Les
2	poules	2	poule
3	du	3	de
4	couvent	4	couvent
5	couvent	5	couver
6	.	6	.

Analyseur

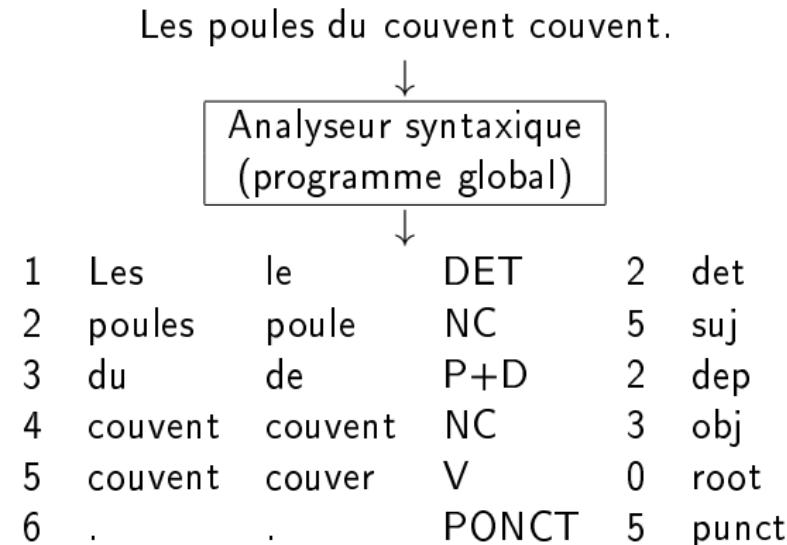
Entrée :			Sortie :				
1	Les	le	1	Les	le	DET	2
2	poules	poule	2	poules	poule	NC	5
3	du	de	3	du	de	P+D	2
4	couvent	couvent	4	couvent	couvent	NC	3
5	couvent	couver	5	couvent	couver	V	0
6	.	PONCT	6	.	.	PONCT	5

Ces 3 couples (entrée, sortie) peuvent constituer :

- une spécification des différents formats
- une partie des tests unitaires de chaque module



Test d'intégration : écriture



Tests unitaires : validation

Les poules du couvent couvent. →

Segmenteur

1 Les
2 poules
3 du
4 couvent
5 couvent
6 .

Tests unitaires : validation

Les poules du couvent couvent.

→ **Segmenteur**

1 Les
2 poules
3 du
4 couvent
5 couvent
6 .

1 Les
2 poules
3 du
4 couvent
5 couvent
6 .

→ **Étiqueteur**

1 Les le DET
2 poules poule NC
3 du de P+D
4 couvent couvent NC
5 couvent couver V
6 . PONCT

Tests unitaires : validation

Les poules du couvent couvent. → **Segmenteur** →

1	Les
2	poules
3	du
4	couvent
5	couvent
6	.

→ **Étiqueteur** →

1	Les	le	DET
2	poules	poule	NC
3	du	de	P+D
4	couvent	couvent	NC
5	couvent	couver	V
6	.		PONCT

1	Les	le	DET
2	poules	poule	NC
3	du	de	P+D
4	couvent	couvent	NC
5	couvent	couver	V
6	.		PONCT

→ **Analyseur** →

1	Les	le	DET	2	det
2	poules	poule	NC	5	suj
3	du	de	P+D	2	dep
4	couvent	couvent	NC	3	obj
5	couvent	couver	V	0	root
6	.		PONCT	5	punct

Tests unitaires : validation

Les poules du couvent couvent. → **Segmenteur** →

1 Les
2 poules
3 du
4 couvent
5 couvent
6 .

1 Les
2 poules
3 du
4 couvent
5 couvent
6 .

→ **Étiqueteur** →

1 Les le DET
2 poules poule NC
3 du de P+D
4 couvent couvent NC
5 couvent couver V
6 . PONCT

1 Les le DET
2 poules poule NC
3 du de P+D
4 couvent couvent NC
5 couvent couver V
6 . PONCT

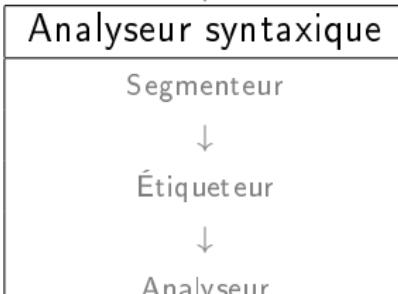
→ **Analyseur** →

1 Les le DET 2 det
2 poules poule NC 5 sujet
3 du de P+D 2 dép
4 couvent couvent NC 3 obj
5 couvent couver V 0 root
6 . PONCT 5 punct

Tests unitaires validés

Test d'intégration

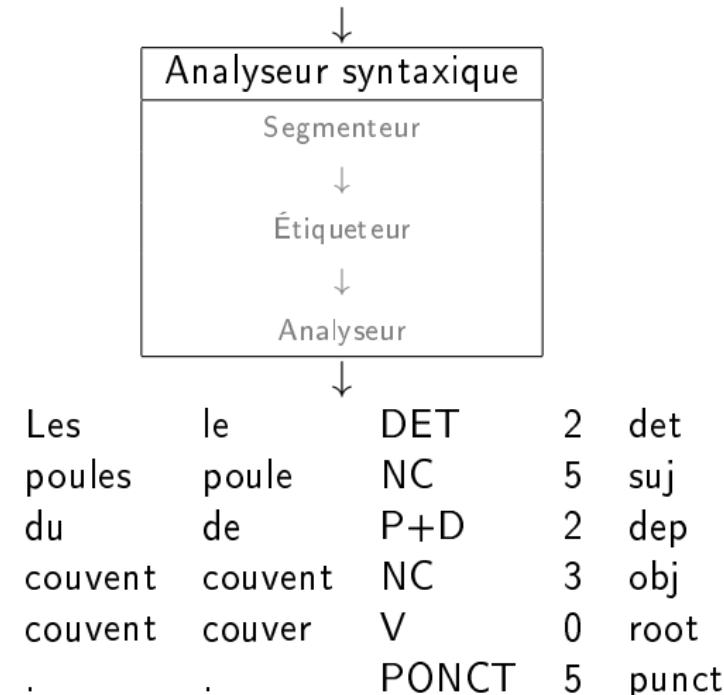
Les poules du couvent couvent.



1	Les	le	DET	2	det
2	poules	poule	NC	5	suj
3	du	de	P+D	2	dep
4	couvent	couvent	NC	3	obj
5	couvent	couver	V	0	root
6	.	.	PONCT	5	punct

Test d'intégration

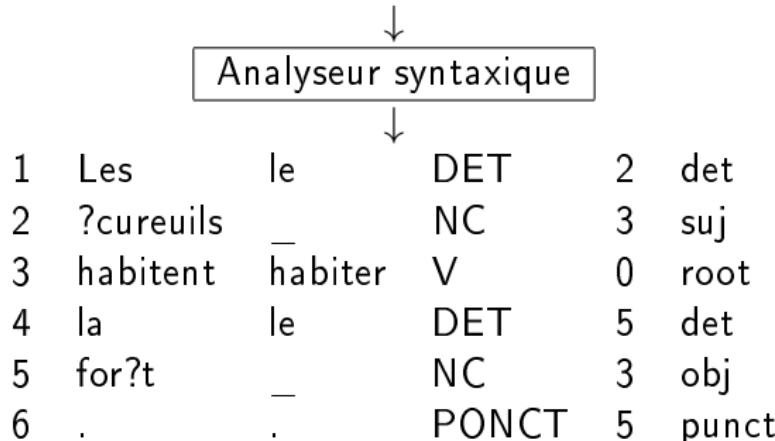
Les poules du couvent couvent.



Test d'intégration validé !

Test d'intégration (2)

Les écureuils habitent la forêt



↓

Analyseur syntaxique

↓

1	Les	le	DET	2	det
2	?cureuils	—	NC	3	suj
3	habitent	habiter	V	0	root
4	la	le	DET	5	det
5	for?t	—	NC	3	obj
6	.	.	PONCT	5	punct

Test d'intégration (2)

Les écureuils habitent la forêt

↓

Analyseur syntaxique					
1	Les	le	DET	2	det
2	?cureuils	—	NC	3	suj
3	habitent	habiter	V	0	root
4	la	le	DET	5	det
5	for?t	—	NC	3	obj
6	.	.	PONCT	5	punct

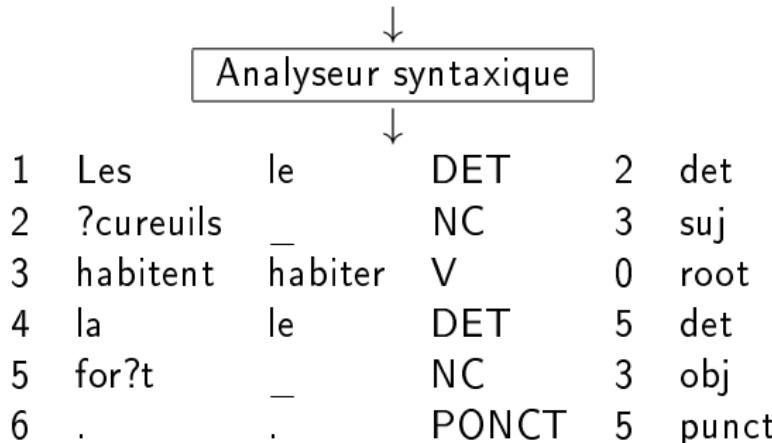
latin1 → **Segmenteur** → latin1

UTF-8 → **Étiqueteur** → UTF-8

UTF-8 → **Analyseur** → UTF-8

Test d'intégration (2)

Les écureuils habitent la forêt



latin1 → Segmenteur → latin1

UTF-8 → Étiqueteur → UTF-8

UTF-8 → Analyseur → UTF-8

latin1 → Segmenteur → Étiqueteur → ???

Description du système

(description des fonctionnalités, des entrées/sorties)

- But : système de sélection de textes de niveau A1 « contenant l'imparfait »
→ contenant au moins un verbe à l'imparfait / N% de verbes à l'imparfait
- Entrée : collection de textes. Sortie : liste de textes sélectionnés

Décomposition

- C1 : Estimation du niveau d'un texte, sélection des textes d'un niveau donné
- C2 : Détection de l'imparfait, sélection des textes contenant N verbes ou X% de verbes à l'imparfait

Pour C1 et C2, entrée = collection de textes,
sortie = sous-collection satisfaisant certains critères



Décomposition

- C1 : Estimation du niveau d'un texte, sélection des textes d'un niveau donné
- C2 : Détection de l'imparfait, sélection des textes contenant N verbes ou X% de verbes à l'imparfait

Pour C1 et C2, entrée = collection de textes,
sortie = sous-collection satisfaisant certains critères

Tests unitaires



Décomposition

- C1 : Estimation du niveau d'un texte, sélection des textes d'un niveau donné
- C2 : Détection de l'imparfait, sélection des textes contenant N verbes ou X% de verbes à l'imparfait

Pour C1 et C2, entrée = collection de textes,
sortie = sous-collection satisfaisant certains critères

Tests unitaires

- pour C1 :
 $\{T1, T2, T3, T4\} \rightarrow \{T3\}$
 $\{T5, T6, T7\} \rightarrow \{T5, T7\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T10, T11\}$



Décomposition

- C1 : Estimation du niveau d'un texte, sélection des textes d'un niveau donné
- C2 : Détection de l'imparfait, sélection des textes contenant N verbes ou X% de verbes à l'imparfait

Pour C1 et C2, entrée = collection de textes,
sortie = sous-collection satisfaisant certains critères

Tests unitaires

- pour C1 :
 $\{T1, T2, T3, T4\} \rightarrow \{T3\}$
 $\{T5, T6, T7\} \rightarrow \{T5, T7\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T10, T11\}$
- pour C2 :
 $\{T1, T2, T3, T4\} \rightarrow \{T2, T3\}$
 $\{T5, T6, T7\} \rightarrow \{\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T9\}$

(T1..T12 = chemins d'accès aux fichiers contenant les textes)



Tests unitaires

- pour C1 :
 $\{T1, T2, T3, T4\} \rightarrow \{T3\}$
 $\{T5, T6, T7\} \rightarrow \{T5, T7\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T10, T11\}$
- pour C2 :
 $\{T1, T2, T3, T4\} \rightarrow \{T2, T3\}$
 $\{T5, T6, T7\} \rightarrow \{\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T9\}$

(T1..T12 = chemins d'accès aux fichiers contenant les textes)

Test d'intégration

Résultat du système = Résultat(C1) \cap Résultat(C2)

Tests unitaires

- pour C1 :
 $\{T1, T2, T3, T4\} \rightarrow \{T3\}$
 $\{T5, T6, T7\} \rightarrow \{T5, T7\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T10, T11\}$
- pour C2 :
 $\{T1, T2, T3, T4\} \rightarrow \{T2, T3\}$
 $\{T5, T6, T7\} \rightarrow \{\}$
 $\{T8, T9, T10, T11, T12\} \rightarrow \{T8, T9\}$

(T1..T12 = chemins d'accès aux fichiers contenant les textes)

Test d'intégration

Résultat du système = Résultat(C1) \cap Résultat(C2)

Test : $\{T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12\}$
 $\rightarrow \{T3, T8\}$

Test d'intégration

Résultat du système = Résultat(C1) \cap Résultat(C2)

Test : {T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12}
 \rightarrow {T3, T8}

Échec possible du test d'intégration

(alors que les tests unitaires sont validés)

- C1 : {/home/m2tal/chemin/T1.txt, /home/m2tal/chemin/T2.txt,
/home/m2tal/chemin/T3.txt, /home/m2tal/chemin/T4.txt}
 \rightarrow {/home/m2tal/chemin/T3.txt}



Test d'intégration

Résultat du système = Résultat(C1) \cap Résultat(C2)

Test : {T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12}
 \rightarrow {T3, T8}

Échec possible du test d'intégration

(alors que les tests unitaires sont validés)

- C1 : {/home/m2tal/chemin/T1.txt, /home/m2tal/chemin/T2.txt, /home/m2tal/chemin/T3.txt, /home/m2tal/chemin/T4.txt}
 \rightarrow {/home/m2tal/chemin/T3.txt}
- C2 : {/home/m2tal/chemin/T1.txt, /home/m2tal/chemin/T2.txt, /home/m2tal/chemin/T3.txt, /home/m2tal/chemin/T4.txt}
 \rightarrow {/home/m2tal/chemin/T2, /home/m2tal/chemin/T3}



Test d'intégration

Résultat du système = Résultat(C1) \cap Résultat(C2)

Test : {T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12}
 \rightarrow {T3, T8}

Échec possible du test d'intégration

(alors que les tests unitaires sont validés)

- C1 : {/home/m2tal/chemin/T1.txt, /home/m2tal/chemin/T2.txt, /home/m2tal/chemin/T3.txt, /home/m2tal/chemin/T4.txt}
 \rightarrow {/home/m2tal/chemin/T3.txt}
- C2 : {/home/m2tal/chemin/T1.txt, /home/m2tal/chemin/T2.txt, /home/m2tal/chemin/T3.txt, /home/m2tal/chemin/T4.txt}
 \rightarrow {/home/m2tal/chemin/T2, /home/m2tal/chemin/T3}

Résultat(C1) \cap Résultat(C2) = {}

pour une raison ou une autre, suppression de l'extension dans les noms de fichier par C2



Système : résumé

- But : système de traduction automatique EN → FR
+ résumé automatique
- Composants : module M1 de traduction EN → FR
+ module M2 de résumé automatique entraîné sur l'anglais
- Échec du test d'intégration car M2 reçoit en entrée un texte en français

Commentaires

Pb de conception initiale, pas de test d'intégration

- Quel test unitaire pour M2 ?
« L'évaluation du composant de résumé automatique des textes anglais est entrepris par un système d'évaluation automatique »
- → c'est le choix du module M2 (résumé EN → EN) qui n'est pas bon.
Et le test unitaire probablement : le test devrait fournir à M2 un texte FR en entrée et s'attendre à trouver un texte FR en sortie.
→ le test devrait invalider le module M2 choisi

Commentaires (suite)

- Syst 1: système de traduction automatique EN → FR
+ résumé automatique

Commentaires (suite)

- Syst 1: système de traduction automatique EN → FR
+ résumé automatique FR → FR

Commentaires (suite)

- Syst 1: système de traduction automatique EN → FR
+ résumé automatique FR → FR
- Syst 2: résumé automatique EN → EN
+ système de traduction automatique EN → FR

Commentaires (suite)

- Syst 1: système de traduction automatique EN → FR
+ résumé automatique FR → FR
- Syst 2: résumé automatique EN → EN
+ système de traduction automatique EN → FR

Tests unitaires (Syst. 2)

- résumé : non déterministe
 - $\{(T1, R11) \rightarrow 1; (T1, R12) \rightarrow 0; (T1, R13) \rightarrow 0; (T1, R14) \rightarrow 1\}$
 - $\{(T2, R21) \rightarrow 0; (T2, R22) \rightarrow 1\}$
 - $\{(T3, R31) \rightarrow 1; (T3, R32) \rightarrow 1; (T3, R33) \rightarrow 0\}$
- « *L'évaluation du composant de résumé automatique des textes anglais est entrepris par un système d'évaluation automatique* »
Attention : si le système d'évaluation change, le résultat des tests unitaires ne doit pas changer !
- traduction : « *par des locuteurs natifs bilingues* »



Commentaires (suite)

- Syst 1: système de traduction automatique EN → FR
+ résumé automatique FR → FR
- Syst 2: résumé automatique EN → EN
+ système de traduction automatique EN → FR

Tests unitaires (Syst. 2)

- résumé : non déterministe
 - $\{(T1, R11) \rightarrow 1; (T1, R12) \rightarrow 0; (T1, R13) \rightarrow 0; (T1, R14) \rightarrow 1\}$
 - $\{(T2, R21) \rightarrow 0; (T2, R22) \rightarrow 1\}$
 - $\{(T3, R31) \rightarrow 1; (T3, R32) \rightarrow 1; (T3, R33) \rightarrow 0\}$

« L'évaluation du composant de résumé automatique des textes anglais est entrepris par un système d'évaluation automatique »

Attention : si le système d'évaluation change, le résultat des tests unitaires ne doit pas changer !
- traduction : *« par des locuteurs natifs bilingues »*

Attention : développer et fixer les textes à l'avance.
Ici, pour une entrée donnée, une infinité de trads possibles.



Tests unitaires (Syst. 2)

- résumé : non déterministe

$\{(T1, R11) \rightarrow 1; (T1, R12) \rightarrow 0; (T1, R13) \rightarrow 0; (T1, R14) \rightarrow 1\}$

$\{(T2, R21) \rightarrow 0; (T2, R22) \rightarrow 1\}$

$\{(T3, R31) \rightarrow 1; (T3, R32) \rightarrow 1; (T3, R33) \rightarrow 0\}$

« *L'évaluation du composant de résumé automatique des textes anglais est entrepris par un système d'évaluation automatique* »

Attention : si le système d'évaluation change, le résultat des tests unitaires ne doit pas changer !

- traduction : « *par des locuteurs natifs bilingues* »

Tests unitaires (Syst. 2)

- résumé : non déterministe

$\{(T1, R11) \rightarrow 1; (T1, R12) \rightarrow 0; (T1, R13) \rightarrow 0; (T1, R14) \rightarrow 1\}$

$\{(T2, R21) \rightarrow 0; (T2, R22) \rightarrow 1\}$

$\{(T3, R31) \rightarrow 1; (T3, R32) \rightarrow 1; (T3, R33) \rightarrow 0\}$

« *L'évaluation du composant de résumé automatique des textes anglais est entrepris par un système d'évaluation automatique* »

Attention : si le système d'évaluation change, le résultat des tests unitaires ne doit pas changer !

- traduction : « *par des locuteurs natifs bilingues* »

Attention : développer et fixer les textes à l'avance.

Ici, pour une entrée donnée, une infinité de trads possibles.

Commentaires

Tests unitaires difficilement réalisables pour la tâche envisagée

- soit viser des phrases à structures simplissimes (et non des textes)
{{("Le petit chat est mort", "The little cat is dead") → 1;
("Le petit chat est mort", "The kitten is dead") → 1;
("Le petit chat est mort", "The kitten died") → 1;
("Le petit chat est mort", "The little cat died") → 1;
("Le petit chat est mort", "The little cat passed away") → 1;
("Le petit chat est mort", "The little cat jumped through the window") → 0;
...
("La petite voiture est morte", "The little car is dead") → 1;
...}}
- soit ne pas faire de tests unitaires
Utiliser effectivement des métriques automatiques *a posteriori*.
Mais métriques toujours existantes ou satisfaisantes.
E.g. BLEU **n'évalue pas** la qualité d'une traduction