

Approches en corpus pour la traduction Le cas MÉTÉO

Philippe Langlais, Thomas Leplus, Simona Gandrabur
and
Guy Lapalme

RALI
Département d'informatique et de recherche opérationnelle
Université de Montréal

TALN – 6-10 Juin 2005

Il était une fois ...

- Dans les années 70, le groupe TAUM de l'université de Montréal développe le système TAUM-MÉTÉO
 - MÉTÉO 1 & 2 ([Grimaila and Chandioux, 1992](#))
- Les systèmes MÉTÉO sont utilisés de manière continue depuis 1984
 - lire ([Hutchins and Somers, 1992](#)) pour plus d'information
 - traduction de 45 000 mots par jour
 - toutes les traductions automatiques sont révisées par des traducteurs professionnels

Il était une fois ...

- Dans les années 70, le groupe TAUM de l'université de Montréal développe le système TAUM-MÉTÉO
 - MÉTÉO 1 & 2 ([Grimaila and Chandioux, 1992](#))
- Les systèmes MÉTÉO sont utilisés de manière continue depuis 1984
 - lire ([Hutchins and Somers, 1992](#)) pour plus d'information
 - traduction de 45 000 mots par jour
 - toutes les traductions automatiques sont révisées par des traducteurs professionnels

Il était une fois ...

- Dans les années 70, le groupe TAUM de l'université de Montréal développe le système TAUM-MÉTÉO
 - MÉTÉO 1 & 2 ([Grimaila and Chandioux, 1992](#))
- Les systèmes MÉTÉO sont utilisés de manière continue depuis 1984
 - lire ([Hutchins and Somers, 1992](#)) pour plus d'information
 - traduction de 45 000 mots par jour
 - toutes les traductions automatiques sont révisées par des traducteurs professionnels

Il était une fois ...

- Dans les années 70, le groupe TAUM de l'université de Montréal développe le système TAUM-MÉTÉO
 - MÉTÉO 1 & 2 ([Grimaila and Chandioux, 1992](#))
- Les systèmes MÉTÉO sont utilisés de manière continue depuis 1984
 - lire ([Hutchins and Somers, 1992](#)) pour plus d'information
 - traduction de 45 000 mots par jour
 - toutes les traductions automatiques sont révisées par des traducteurs professionnels

Corpus parallèle

- Prévisions en français et en anglais produites pendant 2002-2003
- Quelques centaines de lignes Perl + un peu de sueur + de l'alignement de phrases (JAPA) =
le bitexte MÉTÉO

corpus	pairs	sent _{e ≠}	English toks	types	French toks	types
TRAIN	4 187 041	488 391	30 446 549	10 429	37 284 810	11 141
TRAIN _M	4 187 041	301 459	30 290 318	3 352	37 284 810	4 416
DEV	122 357	21 923	891 641	3 022	1 092 208	3 252
DEV _M	122 357	15 454	887 499	1 681	1 092 208	1 908
TEST	36 228	7 878	269 927	1 874	333 370	1 989
TEST _M	36 228	5 994	268 820	1 378	333 370	1 495

(*M* désigne les corpus "méta-tokénisés")

Corpus parallèle

- Prévisions en français et en anglais produites pendant 2002-2003
- Quelques centaines de lignes Perl + un peu de sueur + de l'alignement de phrases (JAPA) =

le bitexte MÉTÉO

corpus	pairs	sent _{e ≠}	English toks	types	French toks	types
TRAIN	4 187 041	488 391	30 446 549	10 429	37 284 810	11 141
TRAIN _M	4 187 041	301 459	30 290 318	3 352	37 284 810	4 416
DEV	122 357	21 923	891 641	3 022	1 092 208	3 252
DEV _M	122 357	15 454	887 499	1 681	1 092 208	1 908
TEST	36 228	7 878	269 927	1 874	333 370	1 989
TEST _M	36 228	5 994	268 820	1 378	333 370	1 495

(M désigne les corpus "méta-tokénisés")

Exemple de bulletin dans les deux langues

FPCN18 CWUL 312130

SUMMARY FORECAST FOR
WESTERN QUEBEC ISSUED BY
ENVIRONMENT CANADA

MONTREAL AT 4.30 PM EST
MONDAY 31 DECEMBER 2001
FOR TUESDAY 01 JANUARY
2002. VARIABLE CLOUDINESS
WITH FLURRIES. HIGH NEAR
MINUS 7.

END/LT

FPCN78 CWUL 312130

RESUME DES PREVISIONS POUR
L'OUEST DU QUEBEC EMISES
PAR ENVIRONNEMENT CANADA

MONTREAL 16H30 HNE LE
LUNDI 31 DECEMBRE 2001
POUR MARDI LE 01 JANVIER
2002. CIEL VARIABLE AVEC
AVERSES DE NEIGE. MAX PRES
DE MOINS 7.

FIN/TR

Motivations

Contexte :

- TIDES, WPT'05,
- Verbmobil ([Wahlster,2000](#)), IWSLT ([Akiba et al., 2004](#))

La tâche MÉTÉO :

- ↪ une tâche réelle
- ↪ domaine restreint
- ↪ une quantité importante de données
- ↪ un système de référence opérationnel

Performance d'approches en corpus sur la tâche MÉTÉO ?

Motivations

Contexte :

- TIDES, WPT'05,
- Verbmobil ([Wahlster,2000](#)), IWSLT ([Akiba et al., 2004](#))

La tâche MÉTÉO :

- ↪ une tâche réelle
- ↪ domaine restreint
- ↪ une quantité importante de données
- ↪ un système de référence opérationnel

Performance d'approches en corpus sur la tâche MÉTÉO ?

Motivations

Contexte :

- TIDES, WPT'05,
- Verbmobil ([Wahlster,2000](#)), IWSLT ([Akiba et al., 2004](#))

La tâche MÉTÉO :

- ↪ une tâche réelle
- ↪ domaine restreint
- ↪ une quantité importante de données
- ↪ un système de référence opérationnel

Performance d'approches en corpus sur la tâche MÉTÉO ?

Motivations

Contexte :

- TIDES, WPT'05,
- Verbmobil ([Wahlster,2000](#)), IWSLT ([Akiba et al., 2004](#))

La tâche MÉTÉO :

- ↪ une tâche réelle
- ↪ domaine restreint
- ↪ une quantité importante de données
- ↪ un système de référence opérationnel

Performance d'approches en corpus sur la tâche MÉTÉO ?

Motivations

Contexte :

- TIDES, WPT'05,
- Verbmobil ([Wahlster,2000](#)), IWSLT ([Akiba et al., 2004](#))

La tâche MÉTÉO :

- ↪ une tâche réelle
- ↪ domaine restreint
- ↪ une quantité importante de données
- ↪ un système de référence opérationnel

Performance d'approches en corpus sur la tâche MÉTÉO ?

Motivations

Contexte :

- TIDES, WPT'05,
- Verbmobil ([Wahlster,2000](#)), IWSLT ([Akiba et al., 2004](#))

La tâche MÉTÉO :

- ↪ une tâche réelle
- ↪ domaine restreint
- ↪ une quantité importante de données
- ↪ un système de référence opérationnel

Performance d'approches en corpus sur la tâche MÉTÉO ?

Plan

1 Un peu d'histoire

2 Motivations

3 Approches

- Mémoire de traduction
- Traduction statistique par segments continus
- Traduction par consensus
- Ré-ordonnancement par réseau de neurones
- Combinaison

4 Bon ou mauvais ?

5 Conclusions & travaux futurs

Pré/Post- traitements

MONDAY .. CLOUDY PERIODS IN THE MORNING WITH **30** PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **pré-traitement**

DAY1 .. CLOUDY PERIODS IN THE MORNING WITH
INT1 PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **traduction**

DAY1 .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE INT1 POUR CENT D AVERSES DE NEIGE EN MATINEE .

↓ **post-traitement**

LUNDI .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE 30 POUR CENT D AVERSES DE NEIGE EN MATINEE .

Pré/Post- traitements

MONDAY .. CLOUDY PERIODS IN THE MORNING WITH 30 PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **pré-traitement**

DAY1 .. CLOUDY PERIODS IN THE MORNING WITH
INT1 PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **traduction**

DAY1 .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE INT1 POUR CENT D AVERSES DE NEIGE EN MATINEE .

↓ **post-traitement**

LUNDI .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE 30 POUR CENT D AVERSES DE NEIGE EN MATINEE .

Pré/Post- traitements

MONDAY .. CLOUDY PERIODS IN THE MORNING WITH 30 PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **pré-traitement**

DAY1 .. CLOUDY PERIODS IN THE MORNING WITH
INT1 PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **traduction**

DAY1 .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE INT1 POUR CENT D AVERSES DE NEIGE EN MATINEE .

↓ **post-traitement**

LUNDI .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE 30 POUR CENT D AVERSES DE NEIGE EN MATINEE .

Pré/Post- traitements

MONDAY .. CLOUDY PERIODS IN THE MORNING WITH 30 PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **pré-traitement**

__DAY1__ .. CLOUDY PERIODS IN THE MORNING WITH
__INT1__ PERCENT CHANCE OF FLURRIES EARLY IN THE MORNING .

↓ **traduction**

__DAY1__ .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE __INT1__ POUR CENT D AVERSES DE NEIGE EN MATINEE .

↓ **post-traitement**

LUNDI .. DEVENANT NUAGEUX TOT EN MATINEE AVEC POSSIBILITE DE **30** POUR CENT D AVERSES DE NEIGE EN MATINEE .

- Peupler la mémoire de toutes les phrases sources de TRAIN.
 - ≈ 85% des phrases de DEV sont présentes dans TRAIN
 - ≈ même couverture avec un dixième seulement des phrases source de TRAIN
- Garder les 5 traductions les plus fréquentes
 - ≈ 90% des phrases de TRAIN n'ont qu'une traduction
- Traduire une nouvelle phrase en allant chercher dans la mémoire les phrases les plus proches — similarité via distance d'édition —
 - 57% des différentes phrases de TEST présentes dans TRAIN
 - 75% ont une phrase proche d'au plus une opération d'édition
 - 84% ont une phrase proche d'au plus deux opérations d'édition

- Peupler la mémoire de toutes les phrases sources de TRAIN.
 - ≈ 85% des phrases de DEV sont présentes dans TRAIN
 - ≈ même couverture avec un dixième seulement des phrases source de TRAIN
- Garder les 5 traductions les plus fréquentes
 - ≈ 90% des phrases de TRAIN n'ont qu'une traduction
- Traduire une nouvelle phrase en allant chercher dans la mémoire les phrases les plus proches — similarité via distance d'édition —
 - 57% des différentes phrases de TEST présentes dans TRAIN
 - 75% ont une phrase proche d'au plus une opération d'édition
 - 84% ont une phrase proche d'au plus deux opérations d'édition

- Peupler la mémoire de toutes les phrases sources de TRAIN.
 - ≈ 85% des phrases de DEV sont présentes dans TRAIN
 - ≈ même couverture avec un dixième seulement des phrases source de TRAIN
- Garder les 5 traductions les plus fréquentes
 - ≈ 90% des phrases de TRAIN n'ont qu'une traduction
- Traduire une nouvelle phrase en allant chercher dans la mémoire les phrases les plus proches — similarité via distance d'édition —
 - 57% des différentes phrases de TEST présentes dans TRAIN
 - 75% ont une phrase proche d'au plus une opération d'édition
 - 84% ont une phrase proche d'au plus deux opérations d'édition

Mémoire de Traduction

Résultats

	TEST			
	WER %	SER %	NIST	BLEU %
memo	5.53	23.73	10.9578	87.69
memo \neq	11.20	49.37	10.8610	78.36

	TEST-H			
	WER %	SER %	NIST	BLEU %
memo	21.02	95.50	9.4048	68.37
memo \neq	22.58	96.73	9.2936	66.21

La traduction de référence aurait un score NIST de 12.2798 (TEST)

Traduction statistique & MÉTÉO

Comment cuisiner un système "phrase-based" à la mode (Koehn et al., 2003)

- Entraîner un modèle de langue avec SRILM (Stolcke, 2002)
ppx de 4.3 sur le corpus TEST !
- Obtenir un modèle "phrase-based"
 - Giza-ifier le corpus d'entraînement ($E \rightarrow F$ and $F \rightarrow E$) (Och et Ney, 2000)
 - Koehn-ifier l'alignement de mots produit par Giza
 - Noter chaque paire de segments (paramètre)
dans notre cas : fréquence relative et score IBM-model 2
- Ajuster les coefficients embarqués dans le décodeur (\approx exhaustivement sur DEV)
- Traduire avec PHARAOH (Koehn, 2004)

Traduction statistique

Extrait d'un modèle "phrase-based"

segment cible	segment source	fréq. rel.
COMBINATION OF THE SNOW	EFFET CONJUGE DE LA NEIGE	0.2
SNOW COMBINED WITH THE	EFFET CONJUGE DE LA NEIGE	0.2
SNOW ALONG WITH	EFFET CONJUGE DE LA NEIGE	0.2
COMBINATION OF SNOW	EFFET CONJUGE DE LA NEIGE	0.4
TO STRONG SOUTH	A FORT DU SUD	0.0027
DEVELOP AHEAD OF	SE LEVERONT A L AVANT D	0.25
WILL DEVELOP AHEAD OF	SE LEVERONT A L AVANT D	0.75
ZERO IN THE AFTERNOON .	DE ZERO EN APRES-MIDI .	1

≈ 2 M. de paramètres (max. de 8 mots de long)
47% des paires vues une seule fois

Traduction statistique

Résultats

	TEST			
	WER %	SER %	NIST	BLEU
memo	5.53	23.73	10.9578	87.69
memo \neq	11.20	49.37	10.8610	78.36
smt	5.33	25.27	11.2683	88.52
smt \neq	8.97	44.96	11.6223	83.17

	TEST-H			
	WER %	SER %	NIST	BLEU
memo	21.02	95.50	9.4048	68.37
memo \neq	22.58	96.73	9.2936	66.21
smt	11.24	57.77	11.0598	81.56
smt \neq	12.42	63.28	11.0259	79.67

Traduction statistique

Résultats

	TEST			
	WER %	SER %	NIST	BLEU
memo	5.53	23.73	10.9578	87.69
memo \neq	11.20	49.37	10.8610	78.36
smt	5.33	25.27	11.2683	88.52
smt \neq	8.97	44.96	11.6223	83.17

	TEST-H			
	WER %	SER %	NIST	BLEU
memo	21.02	95.50	9.4048	68.37
memo \neq	22.58	96.73	9.2936	66.21
smt	11.24	57.77	11.0598	81.56
smt \neq	12.42	63.28	11.0259	79.67

Traduction par consensus (Bootstrapping)

- une idée proposée par (Bangalore et al., 2001)
 - alignement multiple de séquences
 - création d'un automate
 - génération de (peut-être nouvelles, peut-être meilleures) traductions
- nous avons essayé cette idée sur les sorties de la mémoire et/ou celles du moteur statistique

Traduction par consensus

Alignement multiple de séquences

HIGH 12 EARLY THIS MORNING / MAXIMUM 12 TOT CE MATIN

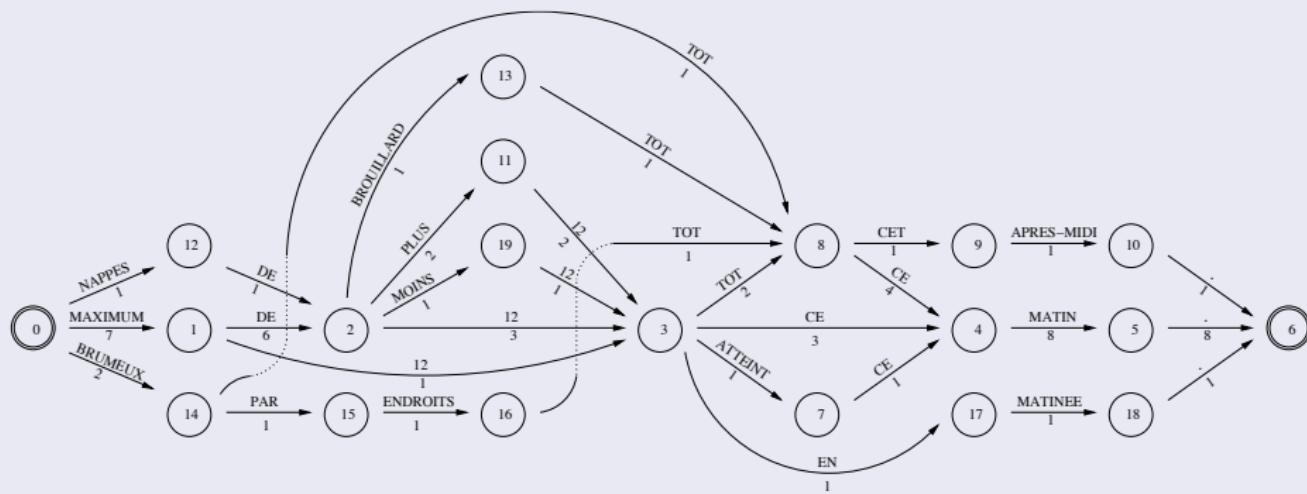
MAXIMUM	DE		12		CE	MATIN
MAXIMUM			12	ATTEINT	CE	MATIN
MAXIMUM	DE		12	TOT	CET	APRES-MIDI
MAXIMUM	DE	PLUS	12	TOT	CE	MATIN
NAPPES	DE	BROUILLARD		TOT	CE	MATIN
BRUMEUX	PAR	ENDROITS		TOT	CE	MATIN
MAXIMUM	DE		12		EN	MATINEE
BRUMEUX				TOT	CE	MATIN
MAXIMUM	DE	PLUS	12		CE	MATIN
MAXIMUM	DE	MOINS	12		CE	MATIN

→ nous avons adapté CLUSTALW (Thompson et al., 1994) à cet effet.

Traduction par consensus

Production d'un automate

HIGH 12 EARLY THIS MORNING / MAXIMUM 12 TOT CE MATIN



Manipulation d'automates avec CARMEL (Knight et Al-Onaizan, 1999)

Traduction par consensus

Génération de nouvelles traductions

HIGH 12 EARLY THIS MORNING / MAXIMUM 12 TOT CE MATIN

MAXIMUM DE PLUS 12 CE MATIN .

MAXIMUM DE 12 CE MATIN .

MAXIMUM DE PLUS 12 TOT CE MATIN .

MAXIMUM DE 12 TOT CE MATIN .

MAXIMUM DE TOT CE MATIN .

MAXIMUM DE ENDROITS TOT CE MATIN .

MAXIMUM DE BROUILLARD TOT CE MATIN .

MAXIMUM DE MOINS 12 CE MATIN .

MAXIMUM DE PLUS 12 EN MATINEE .

MAXIMUM DE PLUS 12 ATTEINT CE MATIN .

Traduction par consensus

Résultats

- Nous avons seulement amélioré les sorties de la mémoire (sur DEV)

	WER	SER	NIST	BLEU
memoire	18.69	94.82	9.7853	66.56
+ consensus	18.97	85.53	9.9314	68.86

- Décevant

Réordonnancement & MÉTÉO

(Rescoring)

- Proposé pour la traduction interactive par ([Grandrabur and Foster, 2003](#))
- Flexible (post-traitement)
- Apprentissage supervisé :
 - Calculer une liste N -best ($N = 1000$) $\Rightarrow a_j, j \leq N$ alternatives
 - Calculer des traits pour chaque alternative $\Rightarrow v_j$
 - Étiqueter chaque alternative comme correcte \oplus ou non \ominus
- Entraîner un réseau de neurones à prédire $p(\oplus|v_j)$
- Ré-ordonner avec ce nouveau score

Entraînements réalisés avec TORCH ([Collobert et al., 2002](#))

Rescoring

Quelques-uns des traits retenus pour entraîner le réseau

- ratio des longueurs source et cible (normalement homogène)
- le meilleur score natif pour une alternative
- quelques statistiques sur la longueur des segments composant l'alternative
- estimée de la probabilité à posteriori

$$c(t_j) = \frac{\sum_i p_j^i}{\sum_j \sum_i p_j^i}$$

- score IBM modèles 1 & 2

Rescoring

Résultats

	TEST			
	WER %	SER %	NIST	BLEU
smt	5.33	25.27	11.2683	88.52
smt+	5.27	25.52	11.3090	88.50
reject	4.77	21.85	11.2471	90.06
oracle	3.29	17.36	11.7187	94.22

les traductions natives et celles ré-ordonnées diffèrent dans 17% des cas

Combinaison d'approches

Un scénario (volontairement) simple

- Laisser faire la mémoire si la phrase à traduire y est présente
- S'en remettre au moteur de traduction (rescoré) sinon

	WER	SER	NIST	BLEU
memo	5.53	23.73	10.9578	87.69
rerank	4.55	21.48	11.2449	90.56
combo	4.40	19.37	11.4133	91.20

Est-ce bon ou pas ?

Nous n'avons trouvé qu'une seule évaluation rigoureuse du système MÉTÉO-2 : ([Macklovitch, 1985](#))

- 1257 phrases produites au cours d'une même journée par *Environnement Canada*.
- E.M. a compté le nombre de fois où la traduction produite par le système différait de la version révisée par un professionnel
- seulement 11% des traductions étaient différentes

Est-ce bon ou pas ?

≈ 11% SER alors que notre meilleur système plafonne à 20% ...

Mais :

- Notre échantillon de test est beaucoup plus grand (36 228 phrases produites pendant un mois)
- Inconsistance dans les traductions de référence (bruit d'environ 7%)
- Une évaluation informelle (à l'interne) a montré que 77% des "mauvaises" traductions produites étaient en fait acceptables (syntaxiquement et sémantiquement).

Exemples d'erreurs commises fréquemment

Et donc facilement réparables

S	TODAY .. PERIODS OF RAIN ENDING NEAR NOON
R	AUJOURD HUI .. PLUIE PASSAGERE CESSANT EN MI-JOURNEE .
T	AUJOURD HUI .. PLUIE PASSAGERE CESSANT VERS MIDI .
A	=====SS=
E	VERS MIDI~~EN MI-JOURNEE
S	40 PERCENT CHANCE OF FLURIES LATE THIS AFTERNOON
R	40 POUR CENT DE PROBABILITE D AVERSES DE PLUIE CET APRES-MIDI
T	POSSIBILITE DE 40 POUR CENT D AVERSES DE PLUIE CET APRES-MIDI
A	II====DD=====
E	POSSIBILITE DE~~φ
	φ~~DE PROBABILITE
S	TEMPERATURE NEAR MINUS 29 .
R	TEMPERATURES PRES DE MOINS 29 .
T	TEMPERATURES DE PRES DE MOINS 29 .
A	=I=====
E	DE~~φ

Exemples d'erreurs commises fréquemment

Et donc facilement réparables

S	TODAY .. PERIODS OF RAIN ENDING NEAR NOON
R	AUJOURD HUI .. PLUIE PASSAGERE CESSANT EN MI-JOURNEE .
T	AUJOURD HUI .. PLUIE PASSAGERE CESSANT VERS MIDI .
A	=====SS=
E	VERS MIDI~~EN MI-JOURNEE
S	40 PERCENT CHANCE OF FLURIES LATE THIS AFTERNOON
R	40 POUR CENT DE PROBABILITE D AVERSES DE PLUIE CET APRES-MIDI
T	POSSIBILITE DE 40 POUR CENT D AVERSES DE PLUIE CET APRES-MIDI
A	II === DD =====
E	POSSIBILITE DE~~ ϕ ϕ~~DE PROBABILITE
S	TEMPERATURE NEAR MINUS 29 .
R	TEMPERATURES PRES DE MOINS 29 .
T	TEMPERATURES DE PRES DE MOINS 29 .
A	=I=====
E	DE~~ ϕ

Exemples d'erreurs commises fréquemment

Et donc facilement réparables

S	TODAY .. PERIODS OF RAIN ENDING NEAR NOON
R	AUJOURD HUI .. PLUIE PASSAGERE CESSANT EN MI-JOURNEE .
T	AUJOURD HUI .. PLUIE PASSAGERE CESSANT VERS MIDI .
A	=====SS=
E	VERS MIDI~~EN MI-JOURNEE
S	40 PERCENT CHANCE OF FLURIES LATE THIS AFTERNOON
R	40 POUR CENT DE PROBABILITE D AVERSES DE PLUIE CET APRES-MIDI
T	POSSIBILITE DE 40 POUR CENT D AVERSES DE PLUIE CET APRES-MIDI
A	II====DD=====
E	POSSIBILITE DE~~φ
	φ~~DE PROBABILITE
S	TEMPERATURE NEAR MINUS 29 .
R	TEMPERATURES PRES DE MOINS 29 .
T	TEMPERATURES DE PRES DE MOINS 29 .
A	=I=====
E	DE~~φ

Conclusions

- Nous avons développé et évalué différents systèmes de traductions sur **la** tâche par excellence en traduction automatique
- Ces approches donnent des performances comparables, même si une comparaison directe est malheureusement impossible
- Des améliorations possibles (et simples)
- Nous étudions une tâche MÉTÉO plus délicate mais néanmoins réelle : la traduction d'alertes météorologiques émises par Environnement Canada (traduites manuellement).

Remerciements

- Rick Jones et Marc Besner d'*Environnement Canada*
- Les auteurs des nombreux *packages* utilisés
(GIZA, SRILM, PHARAOH, CARMEL, CLUSTALW)
- Elliott Macklovitch
- C.B. pour la pertinence de ses commentaires . . .
- Ce travail est financé par le CRSNG (Conseil de Recherches en Sciences Naturelles et Génie) et le FQNRT (Fond de Recherche sur la Nature et les Technologies)

Reference I

-  W.J. Hutchins and H.L. Somers. 1992. *An Introduction to Machine Translation*. Academic Press.
-  A. Grimaila and J. Chandioux, 1992. *Made to measure solutions*, chapter 3, pages 33–45. J. Newton, ed., *Computers in Translation : A Practical Appraisal*, Routledge, London.
-  Wahlster, editor. 2000. *Verbmobil : Foundations of speech-to-speech translations*. Springer Verlag, Berlin, Germany, July.
-  Y. Akiba, M. Federico, N. Kando, H. Nakaiwa, M. Paul, and J. Tsujii. 2004. Overview of the IWSLT04 evaluation campaign. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 1–12, Kyoto, Japan.
-  S. Bangalore, G. Bordel, and G. Riccardi. 2001. Computing consensus translation from multiple machine translation systems. In *Proceedings of ASRU*, pages 351–354, Trento, Italy, Dec.
-  R. Collobert, S. Bengio, and J. Mariéthoz. 2002. Torch : a modular machine learning software library. Technical Report IDIAP-RR 02-46, IDIAP.
-  S. Gandrabur and G. Foster. 2003. Confidence estimation for text prediction. In *Proceedings of CoNLL*, Edmonton, may.
-  K. Knight and Y. Al-Onaizan, 1999. *A Primer on Finite-State Software for Natural Language Processing*, August. <http://www.isi.edu/licensed-sw/carmel/carmel-tutorial2.pdf>.
-  P. Koehn, F.J. Och, and D. Marcu. 2003. Statistical phrase-based translation. In *Proceedings of HLT*, pages 127–133.
-  P. Koehn. 2004. Pharaoh : a beam search decoder for phrase-based smt. In *Proceedings of AMTA*, pages 115–124.
-  T. Leplus, P. Langlais, and G. Lapalme. 2004. Weather report translation using a translation memory. In *Proceedings of the 6th AMTA*, pages 154–163, Washington DC, USA, September.

Reference II

-  E. Macklovitch. 1985. A linguistic performance evaluation of meteo 2. Technical report, Canadian Translation Bureau, Aug.
-  F.J. Och and H. Ney. 2000. Improved statistical alignment models. In *Proceedings of ACL*, pages 440–447, Hongkong, China.
-  A. Stolcke. 2002. SRILM - an extensible language modeling toolkit. In *Proceedings of ICSLP*, Denver, Colorado, Sept.
-  J.D. Thompson, D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W : Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22) :4673–4680.