

# Monter un système de traduction automatique statistique basé sur les séquences de mots: Le cas de la campagne d'évaluation IWSLT

**Philippe Langlais**  
RALI/DIRO  
Université de Montréal  
felipe@iro.umontreal.ca

En collaboration avec **Michael Carl**, IAI, Saarbrücken (carl@iai-uni-sb.de)  
et **Oliver Streiter**, National University of Kaohsiung, Taiwan  
(ostreiter@nuk.edu.tw)

## Outline

- Overview of the IWSLT04 Evaluation Campaign
- Our participation at IWSLT04
  - Few words on the core engine we considered
  - Our phrase extractors
  - Experiments with phrase-based models (PBMs)
- Overview of the participating systems
- Conclusions

# Part I

## Overview of the IWSLT Evaluation Campaign



## International Workshop on Spoken Language Translation

- Two goals:
  - evaluating the available translation technology
  - methodology for evaluating speech translation technologies
- Two pairs of languages: Chinese/English and Japanese/English
- Three tracks: Small, Additional and Unrestricted
- 14 institutions, 20 CE-MT systems, 8 JE ones

Online proceedings: <http://www.slt.atr.co.jp/IWSLT2004/>

## The different Tracks

resources	small	additional	unrestricted
IWSLT 2004 corpus	✓	✓	✓
LDC resources, tagger, chunker, parser	×	✓	✓
other resources	×	×	✓

## Provided corpora

type	language	<i>sent</i>	avr. length	<i>token</i>	<i>types</i>
train	Chinese	20 000	9.1	182 904	7 643
	English	20 000	9.4	188 935	8 191
dev	Chinese	506	6.9	3 515	870
	English	506	7.5	67 410	2 435
test	Chinese	500	7.6	3 794	893

## Translation Domain

The provided corpora were from the BTEC<sup>1</sup> corpus (<http://cstar.atr.jp/cstar-corpus>):

- it 's just down the hall . i 'll bring you some now . if there is anything else you need , just let me know .
- no worry about that . i 'll take it and you need not wrap it up .
- do you do alterations ?
- the light was red .
- we want to have a table near the window .

The Chinese part was tokenized by the organizers.

---

<sup>1</sup>Basic Travel Expressions Corpus

## Participants

SMT	7	ATR-SMT, IBM, IRST, ISI, ISL-SMT, RWTH, TALP	
EBMT	3	HIT, ICT, UTokyo	
RBMT	1	CLIPS	
Hybrid	4	ATR-HYBRID	(SMT + EBMT)
		IAI	(SMT + TM)
		ISL-EDTRL	(SMT + IF)
		NLPR	(RBMT + TM)

## Automatic Evaluation

- 5 automatic metrics computed:
  - NIST/BLEU, *n*-gram precision
  - mWER, *edit distance to the closest reference*
  - mPER, *position independent mWER*
  - GTM, *unigram-based F-measure*
- ↪ 16 references per Chinese sentence
- translations were converted to lower case, punctuations removed



## Automatic Evaluation: Results

0 = perfect				0 = bad					
mWER		mPER		BLEU		NIST		GTM	
0.455	RWTH	0.390	RWTH	0.454	ATR-S	8.55	RWTH	0.720	RWTH
0.469	ATR-S	0.404	ISL-S	0.414	ISL-S	8.34	ISL-S	0.624	ISL-S
0.471	ISL-S	0.420	ATR-S	0.408	RWTH	7.85	IAI	0.685	IAI
0.488	ISI	0.425	ISI	0.374	ISI	7.74	ISI	0.672	ISI
0.507	IRST	0.430	IRST	0.349	IRST	7.48	ATR-S	0.670	ATR-S
0.532	IAI	0.451	IAI	0.346	IBM	7.12	IBM	0.665	IBM
0.538	IBM	0.452	IBM	0.338	IAI	7.09	IRST	0.647	TALP
0.556	TALP	0.465	TALP	0.278	TALP	6.77	TALP	0.644	IRST
0.616	HIT	0.500	HIT	0.209	HIT	5.95	HIT	0.601	HIT

- IAI was tuned on the NIST score only
- best run submitted with 8.00 NIST score

## Human Evaluation

---

### Fluency

---

- 5 Flawless English
  - 4 Good English
  - 3 Non-Native English
  - 2 Disfluent English
  - 1 Incomprehensible
- 

---

### Adequacy

---

- 5 All Information
  - 4 Most Information
  - 3 Much Information
  - 2 Little Information
  - 1 None
-

## Human Evaluation: Results

Fluency		Adequacy	
3.820	ATR-S	3.338	RWTH
3.356	RWTH	3.088	IRST
3.332	ISL-S	3.084	ISI
3.120	IRST	3.056	HIT
3.074	ISI	3.048	ISL-S
2.948	IBM	3.022	TALP
2.914	IAI	2.950	ATR-S
2.792	TALP	2.938	IAI
2.504	HIT	2.906	IBM

→ You won't miss much if you leave now !

## Human Evaluation: a few Facts

- "This indicates that the quality of two systems whose difference in either fluency or adequacy is less than 0.8 cannot be distinguished.", (Akiba,2004).
- Another way of comparing the systems is also provided in the paper (with more or less the same ranking).
- BLEU correlates with fluency, NIST with adequacy (but both are supposed to correlate well with overall human jugements. . . ).

## Part II

# Our participation at IWSLT 2004



## Motivations

- How far can we go in one month of work, starting from (almost) scratch and relying intensively on available packages ?
- Interested by the perspective taken by the organizers: validation of existing evaluation methodologies. See also the CESTA project (TECHNOLANGUE):

<http://www.technolanguage.net/>

To play safe, we participated in:

- the Chinese-to-English track using only the 20K sentences provided

## The core engine

We used an off-the-shelf decoder: Pharaoh (Koehn,2004), based on:

$$\hat{e} = \operatorname{argmax}_{e,I} \prod_{i=1}^I \phi(c_i|e_i)^{\lambda_\phi} d(a_i - b_{i-1})^{\lambda_d} p_{lm}(e)^{\lambda_{lm}} \omega^{|e| \times \lambda_\omega}$$

- a flat PBM (*e.g* small boats  $\leftrightarrow$  bateau de plaisance 0.82)
- we used SRILM (Stolcke,2002) to produce a 3-gram  
ngram-count -interpolate -kndiscount1 -kndiscount2 -kndiscount3
- a set of parameters (one for the PBM, one for the language model, one for the length penalty and one for the built-in distortion model)

## Our phrase-based extractors

We tried two different methods of extraction:

**WABE:** Word-Alignment Based Extractor

— Relying on viterbi alignments computed from IBM model 3

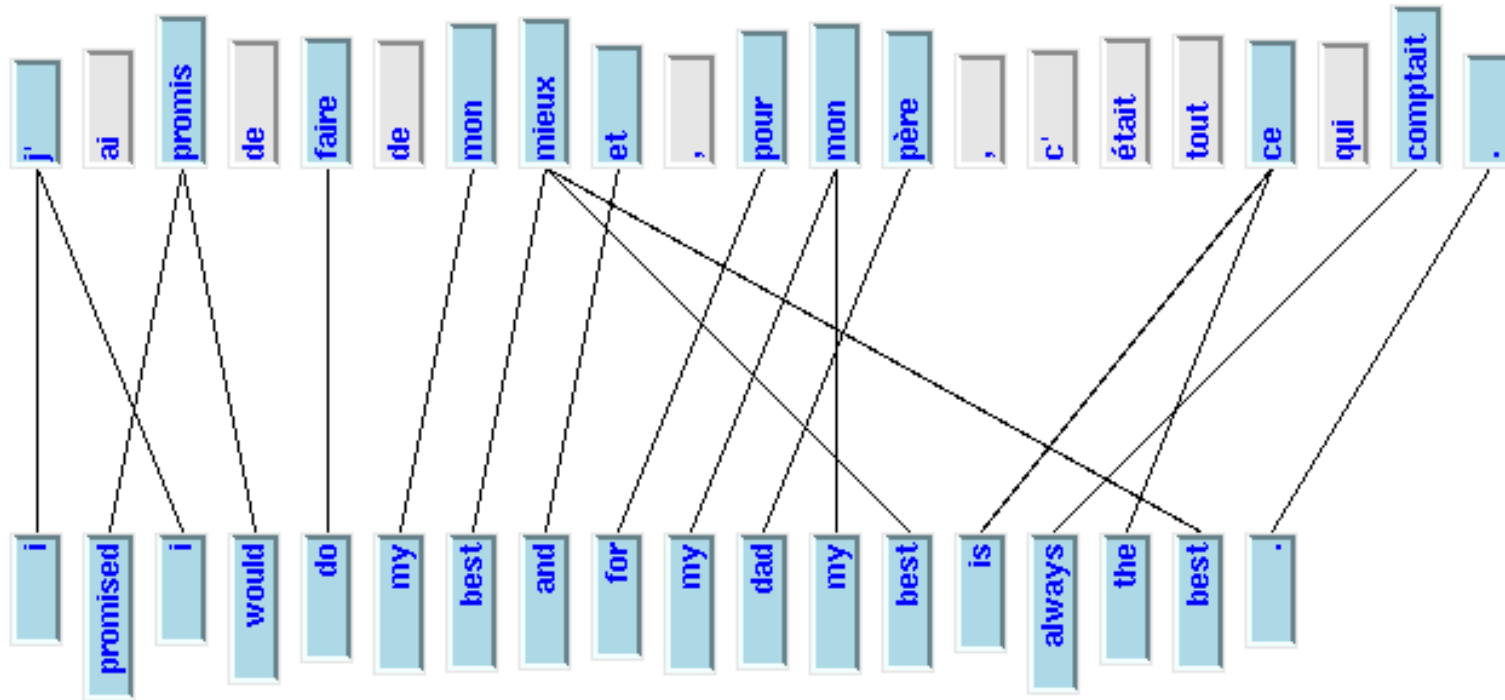
We used Giza++ (Och and Ney, 2000) to obtain them

**SBE:** String-Based Extractor

— Capitalizing on redundancies in the training corpus at the sentence level



## An example of word Alignment



Does not work perfectly (see <http://www.cs.unt.edu/~rada/wpt/>), but nothing to code if you use giza++ !

## WABE: Word-alignment based extractor

Yet another version of (Koehn et al.,2003; Tillmann,2003) and others. Basically:

- Considers the intersection of the word links obtained by viterbi alignment in both directions (C-E, E-C)
- (more or less) carefully extends this set of links with links belonging to the union of both sets (C-E,E-C)

A few meta-parameters control the phrases acquired in this way:

**length-ratio** : `ratio = 2`

**min-max src/tgt length** : `min=1, max=8`

## WABE: An example

.	.	.	.	.	.	.	.	.	.	.	.	.	X
SUNNY	.	.	.	.	.	.	.	.	.	.	.	X	.
MAINLY	.	.	.	.	.	.	.	.	.	X	.	.	.
OTHERWISE	.	.	.	.	.	.	.	.	X	.	.	.	.
PATCHES	.	.	.	.	.	X	↖	.	.	.	.	.	.
FOG	.	.	.	.	X	.	.	.	.	.	.	.	.
MORNING	.	.	.	.	.	.	.	X	.	.	.	.	.
..	.	.	X	.	.	.	.	.	.	.	.	.	.
TODAY	.	X	↖	.	.	.	.	.	.	.	.	.	.
NULL	.	.	.	.	.	↖	.	.	.	.	.	.	.
	N	A	H	.	B	D	B	E	M	P	G	E	.
	U	U	U	.	A	E	R	N	A	U	E	N	
	L	J	I		A		O		T	I	N	S	
	L	O			N		U		I	S	E	O	
		U			C		I		N		R	L	
		R			S		L		E		A	E	
		D					L		E		L	I	
							A				E	L	
							R				M	L	
							D				E	L	
											N	E	
											T		

## WABE: An example

FOG ||| BANCS ||| 1  
 MAINLY SUNNY . ||| GENERALEMENT ENSOLEILLE . ||| 1  
 OTHERWISE MAINLY SUNNY ||| PUIS GENERALEMENT ENSOLEILLE ||| 1  
 OTHERWISE ||| PUIS ||| 1  
 MORNING FOG PATCHES OTHERWISE MAINLY ||| BROUILLARD EN MATINEE PUIS GENERALEMENT ||| 1  
 MORNING FOG PATCHES ||| BROUILLARD EN MATINEE ||| 1  
 TODAY .. MORNING FOG ||| AUJOURD HUI .. BANCS ||| 1  
 . ||| . ||| 1  
 MORNING FOG PATCHES OTHERWISE MAINLY SUNNY . ||| MATINEE PUIS GENERALEMENT ENSOLEILLE . ||| 1  
 MORNING FOG PATCHES OTHERWISE ||| BROUILLARD EN MATINEE PUIS ||| 1  
 MORNING FOG PATCHES OTHERWISE MAINLY SUNNY . ||| BROUILLARD EN MATINEE PUIS GENERALEMENT ENSOLEILLE . ||| 1  
 .. ||| .. ||| 1  
 OTHERWISE MAINLY ||| PUIS GENERALEMENT ||| 1  
 MORNING FOG PATCHES OTHERWISE ||| MATINEE PUIS ||| 1  
 TODAY .. ||| AUJOURD HUI .. ||| 1  
 OTHERWISE MAINLY SUNNY . ||| PUIS GENERALEMENT ENSOLEILLE . ||| 1  
 TODAY ||| AUJOURD HUI ||| 1  
 SUNNY ||| ENSOLEILLE ||| 1  
 PATCHES ||| BROUILLARD EN ||| 1  
 MORNING FOG PATCHES OTHERWISE MAINLY ||| MATINEE PUIS GENERALEMENT ||| 1  
 SUNNY . ||| ENSOLEILLE . ||| 1  
 MORNING FOG PATCHES OTHERWISE MAINLY SUNNY ||| BROUILLARD EN MATINEE PUIS GENERALEMENT ENSOLEILLE ||| 1  
 .. MORNING FOG ||| .. BANCS ||| 1  
 MORNING ||| MATINEE ||| 1  
 MORNING FOG PATCHES OTHERWISE MAINLY SUNNY ||| MATINEE PUIS GENERALEMENT ENSOLEILLE ||| 1  
 MAINLY ||| GENERALEMENT ||| 1



## SBE: String-based extractor

If two strings are in a relation of translation and if part of them also are, then we can induce a specific translation relation between the other parts.

SHOWERS BEGINNING THIS EVENING  
AVERSES DE PLUIE DEBUTANT CE SOIR

SHOWERS BEGINNING THIS EVENING AND ENDING OVERNIGHT .  
AVERSES DE PLUIE DEBUTANT CE SOIR ET CESSANT AU COURS DE LA NUIT .

AND ENDING OVERNIGHT .  
ET CESSANT AU COURS DE LA NUIT .

54 461 parameters out of 20K sentences

## Word-based translation versus PB translation

engine	NIST	BLEU%	MWER	MSER
<i>ibm2+3g</i>	5.0726	26.57	60.56	94.47
Pharaoh	5.5646	26.16	59.70	94.27
wbm by Pharaoh	4.8417	15.54	64.95	97.63

- *ibm2+3g* is an extension of the decoder described by (Niessen et al., 1998)
- Pharaoh was run with its default setting; each parameter of the FPBM was scored by relative frequency

## Tuning the decoder

$\lambda_d$	$\lambda_\phi$	$\lambda_w$	$\lambda_{lm}$	NIST	BLEU%	MWER	MSER
1	1	0	1	5.5646	26.16	59.70	94.27
1	1	-1.5	1	6.3470	25.63	58.93	94.27
.2	.9	-1.5	.8	6.8401	28.44	56.25	94.07

$\lambda_d$ , distortion weight  $([0, 1])$      $\lambda_\phi$ , transfer weight  $([0, 1])$

$\lambda_w$ , word penalty  $([-3, 3])$      $\lambda_{lm}$ , language model weight  $([0, 1])$

We applied a poor man's strategy (sampling uniformly the parameter ranges)

↪ a relative gain over the default configuration (line 1) of 23%

↪ 61% of this gain obtained by tuning only the word penalty parameter

## Merging different FPBMs

config	$ p $	NIST	BLEU%	MWER	MSER
WABE		6.8401	28.44	56.25	94.07
+ WBM		7.0766	31.38	54.88	93.28
+ SBE		7.0926	31.78	54.56	92.69

Merging 2 models was done crudely by:

- copying  $p_i(s|t)$ ,  $\forall s$  whenever  $t$  has not been seen in one model,
- averaging them in case both  $p_1(s|t)$  and  $p_2(s|t)$  exist,
- normalizing

↪ a relative gain of 3.7%



## The weakness of relative frequency

min	max	$ model $	%f1	%f2	%f3+	% $p = 1$
1	8	166 481	90.6	4.9	4.5	74.6
2	8	153 512	92.7	4.3	3.0	78.5
2	4	73 369	87.0	7.1	5.9	68.7

- %f1, %f2 and %f3+ stand for the percentage of parameters (pairs of phrases) seen 1, 2 or at least 3 times in the TRAIN corpus.
- % $p = 1$  stands for the percentage of parameters that have a relative frequency of 1.

## Scoring phrases with IBM model 1

model	NIST	BLEU%	MWER	MSER
relfreq	7.0926	31.78	54.56	92.69
ibm	7.3067	32.98	53.86	92.49
relfreq&ibm	7.3118	34.48	52.73	91.90
relfreq&pn-ibm	7.4219	34.6	53.02	91.70

- baseline model (line 1) = merged FPBM of 306 585 parameters trained by relative frequency.
- rating these parameters by IBM model 1 yields a relative improvement in the NIST score of 3%
- pn-ibm: do not normalize parameters where  $|\{s : p(s|t)\exists\}| = 1$  holds

## Specific models

config	NIST	BLEU%	MWER	MSER
relfreq&ibm	<b>7.3118</b>	<b>34.48</b>	<b>52.73</b>	<b>91.90</b>
A	7.1862	34.21	53.12	91.18
Q	6.4995	34.92	52.12	93.00
specific-lm	<b>7.4702</b>	<b>33.64</b>	<b>53.27</b>	<b>91.90</b>
A	7.3229	33.66	53.08	90.85
Q	6.7010	33.58	53.55	93.50

- around 40% of the training sentences were interrogatives
- ⇒ specific language model combined with the general one (specific tuning over 6 parameters)

(we did not observe improvements by modelling specific FPBMs)

## Translations we submitted before the deadline

*ibm2+3g* word-based translation engine,

*straight* a WABE FPBM

*merge* the best model obtained by merging word and phrase associations

*QA* the one submitted for manual evaluation

*manual* to measure the usefulness of the automatic translations for human post-editing

**Task:** selecting one translation among the generated ones and enhancing its quality through slight modifications

## The *manual* experiment

- 423 (84.6%) were just selections of one of the automatic translations.
- Out of these 423 translations, 85 (20%) were produced by the word-based engine (*ibm2+3g*).

trans1	take a bath for a twin room .
trans2	please take a bath for a double .
trans3	take a bath of double .
trans4	take one twin room with bath .
trans5	have a bath for double .
trans6	have a twin room with bath , please .
trans7	have a double room with bath , please .
manual	<b>please, a twin room with bath .</b>

## Translations we submitted before the deadline

config	BLEU%	NIST	GTM	WER	PER
<i>ibm2+3g</i>	27.27	6.55	62.49	58.12	48.82
<i>straight</i>	30.92	7.52	66.93	56.05	47.90
<i>merge</i>	35.32	8.00	68.60	51.74	43.86
<i>QA</i>	33.89	7.85	68.55	53.24	45.14
<i>manual</i>	36.93	8.13	68.42	49.62	42.53

↪ the ordering of the variants was (almost) consistent with the one observed on the CSTAR corpus

## Part III

# Overview of the participating systems



## Systems' overview: EBM-like approaches

**ATR-HYBRID** 2 EBM-like systems (syst1: edit-distance-based, syst2: grammar-based) + selection of the best output

Sumita, Akiba, Doi, Finch, Imakura, Okuma, Paul, Shimohata, Watanabe

**HIT** extraction of patterns from word-alignment (via bilingual lexicon) + segmentation + statistically flavored selection

Yang, Zhao, Liu, Shi, Jiang

**ICT** EBM-like, many resources used (taggers, training corpora, bilingual word and phrases): closest sentence identification + source segmentation + target material selection + recursively finding non aligned parts

Hou, Deng, Zou, Yu, Liu, Xiong, Liu

**UTokyo** EBM-like with Japanese and English parsers, alignment via bilingual dictionaries

Aramaki, Kurohashi



## Systems' overview: AT-like approaches

**ISI-USC** Och's NIST 2004 system (alignment template model, discriminative training, 12 features)

Ettelaie, Knight, Marcu, Munteanu, Och, Thayer, Tipu

**ITC-IRST** Chinese segmentation + pre/post processing (week days, numbers, etc.) + maxent

Bertoldi, Cattoni, Cettolo, Federico

**RWTH** maxent + simplex (word 3g, word-class 5g, etc.)

Bender, Zens, Matusov, Ney

## Systems' overview: FPBM-like approaches

**ATR-SMT:** HMM phrase-based SMT (PBM analog in spirit to IBM model 4 + log-linear decoder + Simplex)

Sumita, Akiba, Doi, Finch, Imakura, Okuma, Paul, Shimohata, Watanabe

**IAI** PBM + Pharaoh decoder Langlais, Carl , Streiter

**IBM** Extension of Tillmann's engine (R-alignments, reordering of source sentences, unknown Chinese word segmentation, decoder with skip, etc.)

Lee, Roukos

**ISL-SMT** PB-SMT, online phrase acquisition based on a sentence segmentation process based on a variant of IBM-1

Vogel, Hewavitharana, Kolss, Waibel

## Systems' overview: others

**CLIPS** Systran web 5.0 and Systran premium 5.0 !

Blanchon, Boitet, Brunet-Manquat, Tomokiyo, Hamon, Hung, Bey

**NLPR** 3 systems (template based, SMT, and interlingua based) + adhoc selection of their output Zuo,Zhou, Zong

**ISL-EDTRL** statistical transfer rules (unclear how they were trained), simplified English as an Interlingua (ex: please give me → give me . . . please; he had spoken → he spoke) Reichert, Waibel

**TALP** X-grams transducers (X-gram: billette/ticket, clase/second-class ticket) + pre/post-processing (weekdays, cities, dates, etc.)  
De Gispert, B. Marino

## Human Evaluation: The full picture

U-Fluency		U-Adequacy	
3.776	IRST	3.662	ISL-S
3.776	ISL-S	3.526	IRST
3.400	NLPR	3.254	ISL-E
3.036	IBM	3.188	HIT
2.954	ISI	3.082	ICT
2.934	ISL-E	2.996	IBM
2.718	ICT	2.960	CLIPS
2.648	HIT	2.800	NLPR
2.570	CLIPS	2.784	ISI

S-Fluency		S-Adequacy	
3.820	ATR-S	3.338	RWTH
3.356	RWTH	3.088	IRST
3.332	ISL-S	3.084	ISI
3.120	IRST	3.056	HIT
3.074	ISI	3.048	ISL-S
2.948	IBM	3.022	TALP
2.914	IAI	2.950	ATR-S
2.792	TALP	2.938	IAI
2.504	HIT	2.906	IBM

A

3.256	IRST	3.110	IRST
2.846	ISI	2.725	ISI

## Conclusions

- Is phrase-based translation  $\equiv \text{Pharaoh}(\text{Giza++}^{\lambda_g} \times \text{SRILM}^{\lambda_s})$  ?  
↳ at least a decent system can be obtained this way
- Things we tried that did not work better:
  - splitting the training sentences into shorter ones
  - replacing proper names by NAME
- Many factors to be tried:
  - word alignment procedure (Simard and Langlais, 2003)
  - other scoring functions (Zao et al., 2004)
- Unclear why we had a low adequacy score, but a high NIST one.

## References

- [1] Y. Akiba, M. Federico, N. Kando, H. Nakaiwa, M. Paul and J. Tsujii “Overview of the IWSLT04 Evaluation Campaign”, In Proceedings of International Workshop on Spoken Language Translation (IWSLT), Kyoto, Japan, pp. 1–12
- [2] Koehn P., “Pharaoh: a Beam Search Decoder for Phrase-Based SMT”, To appear in Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA), 2004
- [3] Stolcke A., “SRILM - An Extensible Language Modeling Toolkit”, In Proceedings of the International Conference for Speech and Language Processing (ICSLP), Denver, Colorado, September 2002
- [4] Och F.J. and Ney H., “Improved Statistical Alignment Models”, in Proceedings of the Conference of the Association for Computational Linguistics (ACL), Hongkong, China, pp. 440–447, 2000
- [5] Koehn P., Och F.J. and Marcu D., “Statistical Phrase-Based Translation”, In

- Proceedings of the Human Language Technology Conference (HLT), pp. 127–133, 2003
- [6] Tillmann C., “A Projection Extension Algorithm for Statistical Machine Translation”, In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2003
- [7] Niessen S., Vogel S. Ney H. and Tillmann C., “A DP based Search Algorithm for Statistical Machine Translation”, in Proceedings of the International Conference On Computational Linguistics (COLING), pp. 960–966, 1998
- [8] Simard M. and Langlais P., “Statistical Translation alignment with Compositionality Constraints”, HLT-NAACL Workshop: Building and Using Parallel Texts: Data Driven Machine Translation and Beyond, Edmonton, Canada, May 31, pp.19–22, 2003
- [8] Zhao B., Vogel S. and Waibel A., “Phrase Pair Rescoring with Term Weightings for Statistical Machine Translation”, In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Barcelona, Spain, July 2004

## WABE

**Require:**  $\mathcal{P}, \mathcal{R}, minLength, maxLength, ratio$

**Ensure:**  $res$  contains all the pairs of phrases

```
1: Initialization
2:  $res \leftarrow \{\}$ 
3: for all  $x : 1 \rightarrow |S|$  do  $T[x] \leftarrow \{\}$ 
4: for all  $y : 1 \rightarrow |T|$  do  $S[y] \leftarrow \{\}$ 
5:
6: Step1:  $\mathcal{P}$ -projection
7: for all  $(x, y) \in \mathcal{P}$  do  $add(x, y)$ 
8:
9: Step2: Extension
10: for  $p : 1 \rightarrow 2$  do
11:   repeat
12:      $a \leftarrow \{\}$ 
13:     for  $s : 1 \rightarrow |S|$  do
```





```
14:      for all  $t \in T[s]$  do
15:          if  $p = 2$  then
16:              neighbor( $x-1, y-1$ ); neighbor( $x+1, y-1$ );
17:              neighbor( $x-1, y+1$ ); neighbor( $x+1, y+1$ );
18:          else
19:              neighbor( $x-1, y$ ); neighbor( $x+1, y$ );
20:              neighbor( $x, y-1$ ); neighbor( $x, y+1$ );
21:          for all  $(x, y) \in a$  do     $add(x, y)$ 
22:      until  $|a| = 0$ 
23:
24: Step3: Collect independent boxes
25:  $b \leftarrow \{\}$ 
26: for  $x : 1 \rightarrow |S|$  do
27:      $X \leftarrow \{x\}; Y \leftarrow \{\}$ 
28:     repeat
29:          $X_m \leftarrow X; Y_m \leftarrow Y$ 
```



```

30:   for all  $x \in X$  do    $Y \leftarrow Y \cup T[x]$ 
31:   if  $Y \neq Y_m$  then
32:     for all  $y \in Y$  do    $X \leftarrow X \cup S[y]$ 
33:   until  $X = X_m$  and  $Y = Y_m$ 
34:    $b \leftarrow b \cup \left\{ \begin{array}{l} (\min\{x : x \in X\}, \max\{x : x \in X\}), \\ (\min\{y : y \in Y\}, \max\{y : y \in Y\}) \end{array} \right\}$ 
35:    $x \leftarrow \max\{x : x \in X\} + 1$ 
36:
37: Step4: Combine boxes
38: for  $i : 1 \rightarrow |b|$  do
39:   let  $((x_{m_i}, x_{M_i}), (y_{m_i}, y_{M_i})) = b_i$ 
40:    $add(x_{m_i}, x_{M_i}, y_{m_i}, y_{M_i})$ 
41:   for  $j : i + 1 \rightarrow |b|$  do
42:     let  $((x_{m_j}, x_{M_j}), (y_{m_j}, y_{M_j})) = b_j$ 
43:     if  $x_{M_i} + 1 = x_{m_j}$  then
44:        $add(x_{m_i}, x_{M_j}, y_{m_i}, y_{M_j})$ 

```