IJCLA VOL. 2, NO. 1–2, JAN-DEC 2011, PP. 225–237
RECEIVED 31/10/10 ACCEPTED 26/11/10 FINAL 11/02/11

Translating the Penn Treebank with an Interactive-Predictive MT System

MARTHA ALICIA ROCHA[1] AND JOAN ANDREU SÁNCHEZ[2]

[1] *Instituto Tecnológico de León, México*
[2] *Universidad Politécnica de Valencia, Spain*

ABSTRACT

The Penn Treebank corpus is a commonly used corpus in the Computational Linguistics community. This corpus is manually annotated with lexical and syntactical information. It has been extensively used for Language Modeling, Probabilistic Parsing, PoS Tagging, etc. In recent years, with the increasing use of Syntactic Machine Translation approaches, the Penn Treebank corpus has also been used for extracting monolingual linguistic information for further use in these Machine Translation systems. Therefore, the availability of this corpus adequately translated to other languages can be considered an challenging problem. The correct translation of the Penn Treebank corpus by using Machine Translation techniques and then amending the errors in a post-editing phase can require a large human effort. Since there is not parallel text for this dataset, the translation of this corpus can be considered as a translation problem in the absence of in-domain training data. Adaptation techniques have been previously considered in order to tackle this problem. In this work, we explore the translation of this corpus by using Interactive-Predictive Machine Translation techniques, that has proved to be very efficient in reducing the human effort that is needed to obtain the correct translation.

1 INTRODUCTION

The Penn Treebank corpus [1] is a very used corpus in many applications of Computational Linguistic. This corpus consists of approximately fifty thousand sentences that were manually annotated with lexical information and syntactical information. This corpus has been used for many task related to Language Modeling [2–4], Probabilistic Parsing [5–8], PoS Tagging [9, 10], etc. One important advantage of this corpus is that its annotated information allows machine learning techniques to obtain very accurate and profitable linguistic information. Examples of these machine learning techniques are Maximum Entropy [10, 5], Probabilistic Estimation [4, 11] and Grammar Learning [2, 8]. Since this corpus is manually annotated and reviewed by human experts, it allows a reliable comparison with the gold annotation.

In the last years, syntax has become important for Machine Translation (MT) [12–15]. Some of these works are based on the availability of syntactically annotated corpus [13, 14, 16]. Most of these works use the Penn Treebank corpus for learning a syntactical model and then the Syntactic MT system is used in another task different from the Penn Treebank corpus. Therefore, a very interesting step forward in Syntactic MT would be to have the Penn Treebank correctly translated and apply Syntactic MT techniques to this translated corpus. In addition, another very challenging problem would be to use a parallel treebank to study new approaches of Syntactic MT.

The manual translation of the Penn Treebank corpus can be a very tedious and expensive task, and therefore it seems appropriate to carry out this task by using a MT system. The MT system can provide initial translations of the sentences, and a human could then review the full translation. This conventional post-editing review process for obtaining a high quality translation can be also expensive and tedious. In addition, it should be taken into account that there is not parallel text for the Penn Treebank corpus and the translation systems should be trained with out-of-domain data. Therefore, the translation quality of the output of the MT system could be low. Adaptation techniques for translating the Penn Treebank corpus were considered in [17] to alleviate this problem, but the final obtained results showed that a lot of errors were yet present in the final translations, and a lot of effort should be carried out to obtain correct high-quality translated sentences.

Therefore in this work we studied Interactive-Predictive MT (IPMT) techniques [18] for translating the Penn Treebank corpus. In IPMT, the human translator and the MT system work together in order to obtain correct high-quality translations. The human translator provides feedback to the MT translation and the system takes into account this feedback in order to constrain the search space and to avoid further errors. IPMT was proven to be very effective for MT translation in [19] for in-domain tasks. But some tasks require to translate texts for which parallel text is not available. In such situation, the IPMT system starts with out-of-domain training data. However, [19] did not explore IPMT in the translation of an out-of-domain task. In this work we explored IPMT for translating the Penn Treebank corpus in which the initial system is trained with out-of-domain data.

2 INTERACTIVE-PREDICTIVE MACHINE TRANSLATION

Statistical MT has evolved rapidly in the last years, specially after the appearing the seminal papers of [20, 21]. Those MT systems were mainly based on words as basic unit translation. Currently, the state-of-the-art statistical MT systems use phrases as basic translation units [22, 23], although in recent years the syntax-based approach has provided very promising results [15].

In statistical MT, the problem can be stated as:

$$\widehat{\mathbf{y}} = \arg\max_{\mathbf{y}} \Pr(\mathbf{y}|\mathbf{x}) = \arg\max_{\mathbf{y}} \Pr(\mathbf{x}|\mathbf{y}) \Pr(\mathbf{y}), \tag{1}$$

where $\mathbf{x}$ is a given input source sentence, $\Pr(\mathbf{y})$ is the language model probability and $\Pr(\mathbf{x}|\mathbf{y})$ is the translation model probability. The maximization is carried out over all possible output target sentences according to the language model. Expression (1) has been also stated in a different way by considering a *log linear* model [22]. In (1), both statistical models $\Pr(\mathbf{y})$ and $\Pr(\mathbf{x}|\mathbf{y})$ are usually trained on a very large training corpus.

For tasks in which both the training data and the test data are in the same domain, statistical MT systems that are based on the previously mentioned approaches are able to provide good translations results if the two languages share common structure and enough training data is available . The output sentences provided by the systems allow the user to understand the source sentence. However, the output sentence usually contains a lot of errors. If error-free translated sentences are required, then a human translator should review and correct the errors in a post-editing process. Interactive-Predictive MT (IPMT) intends to reduce the human effort that is needed to carry out this correction process.

In IPMT, the system and the user participate in a tight way in order to obtain the correct translation of an input sentence. First, the system provides an initial translation of the input sentence. Then, the user amends the first detected erroneous word. This action implicitly involves to accept as correct a given prefix of the output sentence. This loop continues until the translation that the user has in mind is obtained. Every time the user provides a correction, the system incorporate the correction introduced by the user to the translation system in order to constrain the search space and to avoid further errors [18]. Next time that the system provides a new translation, it takes into account the correct prefix. In this way expression (1) is modified as follows:

$$\widehat{\mathbf{y}}_s = \arg\max_{\mathbf{y}_s} \Pr(\mathbf{y}_s | \mathbf{x}, \mathbf{y}_p), \tag{2}$$

where $\mathbf{y}_p$ and $\mathbf{y}_s$ are, respectively, the prefix and the suffix of the target sentence, and $\mathbf{y}_p$ includes the feedback provided by the user. Note that if $\mathbf{y} = \mathbf{y}_p\mathbf{y}_s$, the expression (1) and expression (2) are similar. This time, the search is carried out over the set of suffixes $\mathbf{y}_s$ that complete $\mathbf{y}_p$. Clearly, the feedback information $\mathbf{y}_p$ provided by the user is the opportunity to get better $\mathbf{y}_s$. The search process in the IPMT approach is carried out over a word graph. This word graph is obtained automatically after the system proposes the first hypothesis.

Figure 1 illustrates an example of how the IPMT system interacts with the human in a editing activity. In each iteration Iter-i the system uses a validated prefix $\mathbf{y}_p$ that it completes with a suffix $\mathbf{y}_s$ to compose the best hypothesis $\widehat{\mathbf{y}}$ of the following iteration. In the following iteration the user validates implicitly a new prefix $\mathbf{y}_p$ by typing an incorrect word $\mathbf{w}$, and again the system suggests a suitable continuation $\mathbf{y}_s$ for the following iteration. This process is repeated until a complete translation of the source sentence is reached. In the final translation, the 3 words typed by the user are underlined. In this example the estimated post-editing effort would be 13/23 (57%), produced by the errors: insert *"de"*, remove *"juntará el tablero"*, insert *"se unirá a la junta"*, remove *"29"*, and insert *"el 29 de"*. The corresponding interactive estimate is 3/23 (13%). This results in an estimated user effort reduction of 77%.

In the example that we have showed above the user corrects an error every time by typing a new correct word. If the new composed prefix that includes the typed word is not in the word graph, then the most probable path is computed by using an error-correcting algorithm. In [19], other ways of amending errors were studied. They proposed to use Mouse Actions (MA) as an additional feedback in order to obtain the correct translation. These MA could be of two ways: non-explicit positioning MA and interaction-explicit MA. In *non-explicit positioning MA*, when the

Source ($\mathbf{x}$) :		Pierre Vinken , 61 years old , will join the board as a nonexecutive director Nov. 29 .
Target ($\mathbf{y}$):		Pierre Vinken , de 61 años de edad , se unirá a la junta como director no ejecutivo el 29 de noviembre .
Iter-0	$\mathbf{y}_p$	()
	$\mathbf{y}_s$	Pierre Vinken , 61 años de edad , juntará el tablero como director no ejecutivo noviembre 29 .
Iter-1	$\widehat{\mathbf{y}}$	Pierre Vinken , 61 años de edad , juntará el tablero como director no ejecutivo noviembre 29 .
	$\mathbf{y}_p$	Pierre Vinken ,
	$\mathbf{w}$	<u>de</u>
	$\mathbf{y}_s$	61 años de edad , juntará el tablero como director no ejecutivo noviembre 29 .
Iter-2	$\widehat{\mathbf{y}}$	Pierre Vinken , de 61 años de edad , juntará el tablero como director no ejecutivo noviembre 29 .
	$\mathbf{y}_p$	Pierre Vinken , de 61 años de edad ,
	$\mathbf{w}$	<u>se</u>
	$\mathbf{y}_s$	unirá a la junta como director no ejecutivo noviembre 29 .
Iter-3	$\widehat{\mathbf{y}}$	Pierre Vinken , de 61 años de edad , se unirá a la junta como director no ejecutivo noviembre 29 .
	$\mathbf{y}_p$	Pierre Vinken , de 61 años de edad , se unirá a la junta como director no ejecutivo
	$\mathbf{w}$	<u>el</u>
	$\mathbf{y}_s$	29 de noviembre .
Final	$\widehat{\mathbf{y}}$	Pierre Vinken , de 61 años de edad , se unirá a la junta como director no ejecutivo el 29 de noviembre .
	$\mathbf{w}$	#
	$\mathbf{y}_p \equiv \mathbf{y}$	Pierre Vinken , de 61 años de edad , se unirá a la junta como director no ejecutivo el 29 de noviembre .

Fig. 1. Simulated example of IPMT interaction to translate the sentence of the Penn Treebank *"Pierre Vinken , 61 years old , will join the board as a nonexecutive director Nov. 29 ."*.

user positions the cursor in the place where he wants to type a new word, he is implicitly indicating that the word after the cursor is incorrect. Then, the search engine can start to look for a new suffix in which the first word is different from the current word after the cursor. In *interaction-explicit MA*, the user asks for a new suffix each time he presses the mouse. Each new suffix has to start with a word different from initial words that have appeared in previous rejected suffixes. This is formalized as follows:

$$\widehat{\mathbf{y}}_s = \underset{\mathbf{y}_s : \mathbf{y}_{s_1} \neq \mathbf{y}_{s_1}^{(i)} \forall i \in \{1..n\}}{\arg\max} \Pr(\mathbf{y}_s | \mathbf{x}, \mathbf{y}_p, \mathbf{y}_s^{(1)}, \mathbf{y}_s^{(2)}, \ldots, \mathbf{y}_s^{(n)}), \quad (3)$$

where $\mathbf{y}_{s_1}^{(i)}$ is the first word of the *i*-th suffix discarded and $\mathbf{y}_s^{(1)}$, $\mathbf{y}_s^{(2)}$, …, $\mathbf{y}_s^{(n)}$ are the n suffixes discarded. Note that in some sense the *non-explicit positioning MA* approach is included in the *interaction-explicit MA* approach.

In the following section we present experiments in which we use an IPMT framework for the translation of the Penn Treebank corpus.

3 Experiments

In this section we describe the experiments that we carried out to translate the Penn Treebank using an IPMT system with the interaction-explicit MA approach previously described.

3.1 *Datasets*

The Penn Treebank corpus is an annotated corpus that has approximately $49,000$ sentences. From this corpus, we used $1,141$ sentences from section 23 for the experiments, since this section is usually used for testing. These sentences were manually translated to Spanish by human experts without the help of any MT system: 500 of them were obtained from [17], and the other 641 were obtained from [24][3]. The source sentences of both datasets did not overlap each other. The 500 translated sentences from [17] were reviewed by two native Spanish speakers. The 641 sentences from [24] were translated by other human experts different from [17]. We called the $1,141$ sentences dataset Small Parallel Penn Treebank set (SPPT), we called SPPT-R the dataset from [17], and we called SPPT-M

[3] This translated corpus is available at `http://www.dsic.upv.es/~jandreu/SPTB.tgz`

the dataset from [24]. Note that the SPPT dataset was the union of the SPPT-R and SPPT-M datasets. We distinguished between SPPT-M and SPPT-R because they were translated in different places and in different contexts, and we wanted to check if there were notably differences in the translation experiments due to these differences. Table 1 presents the main characteristics of these datasets.

Table 1. Characteristics of the SPPT-R, SPPT-M, and SPPT datasets.

	SPPT-R		SPPT-M		SPPT	
Total of	**English**	**Spanish**	**English**	**Spanish**	**English**	**Spanish**
Sentences	500	500	641	641	1,141	1,141
Running Words	12,172	13,175	15,133	16,847	27,305	30,022
Vocabulary	2,613	2,946	3,613	4,120	4,918	5,862

Since there is no in-domain parallel data to train a system for translating the Penn Treebank, we had to use an out-of-domain parallel text. We used the second version of the *Europarl* bilingual corpus [25]. This corpus was used for training a phrase-based translation model. For these experiments we used only the sentences that had less than or equal to 40 words. The main characteristics of this training set can be seen in Table 2. All the experiments were carried out with uncapitalized text and appropriately tokenized.

Table 2. Characteristics of *Europarl* corpus.

Total of	**English**	**Spanish**
Sentences	730,740	730,740
Running Words	15,242,854	15,702,800
Vocabulary	64,076	102,821

3.2 *Assessment Metrics*

For assessment, we used some of the metrics defined and used in [18, 19]. The quality of the interactive translation is given by the Word Stroke

Ratio (WSR) defined as the number word-strokes that a user would need to perform in order to obtain the reference translation divided by the total number of words in the reference sentence. Note that a word-stroke is considered as a single action. We expect that the number of word strokes that are necessary to obtain the correct translation decreases in the IPMT system as the prediction changes their predictions and provides more accurate suffixes. Note that the same metric can be used for a non IPMT, and in such case it is similar to the usual Word Error Rate (WER) metric. For this reason we used this metric also for the post-editing experiments.

Another metric that we used was the Word Click Ratio (WCR) defined as the number of mouse actions per word that the user had to perform before accepting a new prediction with respect to using exclusively the keyboard in IPMT system.

The also measured the Effort Reduced (ER) as the relative difference between two evaluations in WSR.

3.3 *Results*

For the IPMT experiments, first of all, an English-Spanish phrase-based MT system was built. This was carried out by means of the public software THOT[4], GIZA++[5], and SRILM[6]. THOT and GIZA++ were used for training the translation model ($\Pr(\mathbf{x}|\mathbf{y})$ in expression (1)), and SRILM was used for obtaining a 5-gram language model ($\Pr(\mathbf{y})$ in expression (1)). A multi-stack decoder [26] was used to generated the word graphs and the hypotheses. Note that no process was carried out in order to adjust the parameter of the log-lineal model, because in such case we should leave some data for development and the dataset was not very large.

Experiments were carried out with the three datasets previously described, that is, the SPPT, SPPT-R, and SPPT-M datasets. In all the experiments we computed the WSR, WCR and ER metrics. Table 3 shows the obtained results. We compared two scenarios: first scenario corresponds to a classical post-editing scenario (column Post-edit in Table 3). Second scenario corresponds to an IPMT scenario in which the user carries out just a single MA (column IPMT in Table 3). Column ER shows an estimation of the percentage of ER achieved by using the IPMT scenario with respect to post-editing scenario. It is important to note that the

[4] `http://sourceforge.net/projects/thot/`
[5] `http://www-i6.informatik.rwth-aachen.de/Colleagues/och/software/GIZA++.html`
[6] `http://www-speech.sri.com/projects/srilm/`

percentage of ER increased for the three datasets. Note also that the results in the post-editing scenario reflect some lexical differences between SPPT-R and SPPT-M. Values in column Post-edit reflects that the initial hypothesis provided by the system had a lot of error, but it must be taken into account that the models were trained on out-of-domain data. *Europarl* is a dataset related to speech transcription of the European parliament. However, Section 23 of Penn Treebank contains a lot of stock market jargon and proper nouns, ant therefore a lot of out-of-vocabulary words.

Table 3. Obtained results (in percentage) with the SPPT, SPPT-R, and SPPT-M datasets using conventional post-editing against IPMT with a single MA.

	Post-edit	**IPMT**	
Dataset	**WER**	**WSR**	**ER**
SPPT-R	70.3	61.2	13.1
SPPT-M	74.3	65.5	11.8
SPPT	72.5	63.6	12.3

It is important to remark that in these experiments we obtained slightly better results than those reported in [19] for an analogous experiment. Thus, Figure 4 in [19], reported ER reduction with just one click of about 10% when translating from Spanish to English but starting from 60%, 7% from German to English starting from 70% of WER, and 10% when translating from French to English starting from 60% of WER. No experiment was reported in [19] from English to Spanish.

Figure 2 shows the WSR, WCR and ER for only the SPPT dataset as a function of the maximal number of MA allowed per incorrect word by the user before writing the correct word. We omitted the other datasets because the results were very similar. Point 0 in the WSR plot corresponds approximately to the conventional post-editing scenario, and coincides with row SPPT in column Post-edit of Table 3; point 1 coincides with IPMT column of the same row.

The difference between point 0 and point 1 in ER plot corresponds to the percentage 12.3 in Table 3. Note that the WSR decreased notably with just four MA. This reduction is along the line of the results reported in [19], or slightly better. WCR plot shows that for the number of average MA per word kept almost constant as the number of maximum MA

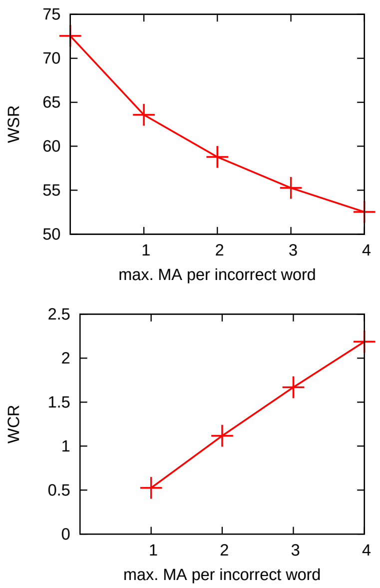

Fig. 2. WSR (top left), ER (top right) and WCR (bottom left) as a function of the maximal clicks of MA allowed by the user before write a new word.

allowed increased. Thus, for just 1 MA the quotient between WCR and max. MA was 0.52 (0.52/1.0), and for 4 MA this value was 0.54 (2.18/4).

A good trade-off is obtained when the maximum number of clicks is around 2 clicks, because a significant amount of effort is saved with a low number of extra clicks per word. These results showed that an adequate number of clicks to improve efficiency and reduce post-editing effort properly is a maximum of 2 or even 3 MA.

4 Conclusions

In this work we studied the use of IPMT for translating the Penn Treebank corpus. We followed the IPMT approach explored in [19] in which MA were considered as an input modality. We explored those ideas for a task in which there is not in-domain training data, and therefore out-of-domain training data had to be used. We proved that this input modality was informative enough in order to the obtain large human effort reduction even in this context. As a final comment, we can conclude that the

IPMT framework is a realistic approach to obtain high-quality translations in absence of in-domain training data.

Note that IPMT is stated for an scenario in which an expert translator collaborates on-line with a MT system in order to provide high quality translations. The user translates a sentence each time. Therefore, for future work we intend to explore learning techniques that make easier the use of the IPMT framework in order to deal with the translation of the Penn Treebank corpus. Thus, some techniques that we consider for future include Active Learning techniques [27, 28] that would allow us to select the sentences to be translated each time, and Online Learning techniques [11, 29] that would allow the models to be learned each time a sentence was translated.

ACKNOWLEDGEMENTS This work has been partially supported by the EC (FEDER/FSE) and the Spanish MEC/MICINN under the MIPRCV program "Consolider Ingenio 2010" (CSD2007-00018), the MITTRAL project (TIN2009-14633-C03-01), and the Generalitat Valenciana under grant Prometeo/2009/014. The first author is supported by "División de Estudios de Posgrado e Investigación" and by "Metrología y Sistemas Inteligentes" research group of Instituto Tecnológico de León.

REFERENCES

1. Marcus, M., Santorini, B., Marcinkiewicz, M.: Building a large annotated corpus of english: the penn treebank. Computational Linguistics **19**(2) (1993) 313–330
2. Chelba, C., Jelinek, F.: Structured language modeling. Computer Speech & Language **14**(4) (2000) 283–332
3. Charniak, E., Knight, K., Yamada, K.: Syntax-based language models for statistical machine translation. In: Proc. of MT Summit IX, New Orleans, USA (September 2003)
4. Benedí, J., Sánchez, J.: Estimation of stochastic context-free grammars and their use as language models. Computer Speech & Language **19**(3) (2005) 249–274
5. Charniak, E.: A maximum-entropy-inspired parser. In: Proc. of NAACL-2000. (2000) 132–139
6. Roark, B.: Probabilistic top-down parsing and language modeling. Computational Linguistics **27**(2) (2001) 249–276
7. Collins, M.: Head-driven statistical models for natural language parsing. Computational Linguistics **29**(4) (2003) 589–637

8. Klein, D., Manning, C.D.: A* parsing: Fast exact viterbi parse selection. In: In Proceedings of HLT-NAACL 03. (2003) 119–126
9. Brill, E.: Transformation-based error-driven learning and natural language processing: a case study in part-of-speech tagging. Comput. Linguist. **21**(4) (1995) 543–565
10. Ratnaparkhi, A.: A maximum entropy model for part-of-speech tagging. In: Proc. Empiriral Methods in Natural Language Processing, University of Pennsylvania (May 1996) 133–142
11. Liang, P., Klein, D.: Online em for unsupervised models. In: Proceedings of HLT-NAACL 09, Stroudsburg, PA, USA (2009) 611–619
12. Wu, D.: Stochastic inversion transduction grammars and bilingual parsing of parallel corpora. Comput. Linguist. **23** (September 1997) 377–403
13. Yamada, K., Knight, K.: A syntax-based statistical translation model. In: In: Meeting on Association for Computational Linguistics, Stroudsburg, PA, USA (2001) 523–530
14. Yamada, K., Knight, K.: A decoder for syntax-based statistical mt. In: In: Meeting on Association for Computational Linguistics, Stroudsburg, PA, USA (2002) 303–310
15. Chiang, D.: Hierarchical phrase-based translation. Computational Linguistics **33**(2) (2007) 201–228
16. Petrov, S., Haghighi, A., Klein, D.: Coarse-to-fine syntactic machine translation using language projections. In: Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, Honolulu, Hawaii, Association for Computational Linguistics (October 2008) 108–116
17. Rocha, M.A., Sánchez, J.A.: Machine translation of the penn treebank to spanish. In: I Joint SIG-IL/Microsoft Workshop on Speech and Language Technologies for Iberian Languages, Porto Salvo, Portugal (3-4 September 2009)
18. Barrachina, S., Bender, O., Casacuberta, F., Civera, J., Cubel, E., Khadivi, S., Ney, A.L.H., Tomás, J., Vidal, E.: Statistical approaches to computer-assisted translation. Computational Linguistics **35**(1) (2009) 2–28
19. Sanchis-Trilles, G., Ortiz-Martínez, D., Civera, J., Casacuberta, F., Vidal, E., Hoang, H.: Improving interactive machine translation via mouse actions. In: EMNLP 2008: conference on Empirical Methods in Natural Language Processing, Waikiki, Honolulu, Hawaii (October 25 – 27 2008)
20. Brown, P.F., Cocke, J., Della Pietra, S., Della Pietra, V.J., Jelinek, F., Lafferty, J.D., Mercer, R.L., Roossin, P.S.: A statistical approach to machine translation. Computational Linguistics **16**(2) (1990) 79–85
21. Brown, P., Pietra, S.D., Pietra, V.D., Mercer, R.: The mathematics of statistical machine translation: parameter estimation. Computational Linguistics **19**(2) (1993) 263–311
22. Och, F.J., Ney, H.: The alignment template approach to statistical machine translation. Comput. Linguist. **30** (December 2004) 417–449

23. Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., Herbst, E.: Moses: Open source toolkit for statistical machine translation. In: Proc. of the ACL Companion Volume Proceedings of the Demo and Poster Sessions. (June 2007) 177–180
24. Mellebeeck, B.: TransBooster: Black Box Optimisation of Machine Translation Systems. Phd thesis, School of Computing, Dublin City University (2007)
25. Koehn, P., Monz, C., eds.: Proceedings on the Workshop on Statistical Machine Translation. Association for Computational Linguistics, New York City (June 2006)
26. Ortiz, D., García-Varea, I., Casacuberta, F.: Generalized stack decoding algorithms for statistical machine translation. In: Proceedings on the Workshop on Statistical Machine Translation, HLT-NAACL 2006, New York City, USA (June 4–9 2006) 64–71
27. Burr, S., Mark, C.: An analysis of active learning strategies for sequence labeling tasks. In: Proc. of the Conference on Empirical Methods in Natural Language Processing EMNLP, ACL press (2008) 1069–1078
28. Haffari, G., Roy, M., Sarkar, A.: Active learning for statistical phrase-based machine translation. In: Human Language Technologies: The 2009 Annual Conference of the North America Chapter of the ACL, ACL press (June 2009) 415–423
29. Ortiz-Martinez, D., Garcia-Varea, I., Casacuberta, F.: Online learning for interactive statistical machine translation. In: Proceedings of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT). (Jun 2010) 173–177

MARTHA ALICIA ROCHA
DEPARTAMENTO DE SISTEMAS Y COMPUTACIÓN,
INSTITUTO TECNOLÓGICO DE LEÓN,
MÉXICO
E-MAIL: <MROCHA@DSIC.UPV.ES>

JOAN ANDREU SÁNCHEZ
INSTITUTO TECNOLÓGICO DE INFORMÁTICA,
UNIVERSIDAD POLITÉCNICA DE VALENCIA,
SPAIN
E-MAIL: <JANDREU@DSIC.UPV.ES>