International Journal of Modern Physics C
Vol. 19, No. 4 (2008) 583–598

World Scientific
www.worldscientific.com

COMPLEX NETWORKS ANALYSIS OF MANUAL AND MACHINE TRANSLATIONS

DIEGO R. AMANCIO*, LUCAS ANTIQUEIRA†, THIAGO A. S. PARDO*, LUCIANO da F. COSTA†, OSVALDO N. OLIVEIRA, Jr.† and MARIA G. V. NUNES*

**Instituto de Ciências Matemáticas e de Computação*
Universidade de São Paulo, P. O. Box 668, 13560-970
São Carlos, São Paulo, Brazil

†Instituto de Física de São Carlos
Universidade de São Paulo, P. O. Box 369, 13560-970
São Carlos, São Paulo, Brazil

Received 19 October 2007
Accepted 8 November 2007

Complex networks have been increasingly used in text analysis, including in connection with natural language processing tools, as important text features appear to be captured by the topology and dynamics of the networks. Following previous works that apply complex networks concepts to text quality measurement, summary evaluation, and author characterization, we now focus on machine translation (MT). In this paper we assess the possible representation of texts as complex networks to evaluate cross-linguistic issues inherent in manual and machine translation. We show that different quality translations generated by MT tools can be distinguished from their manual counterparts by means of metrics such as in- (ID) and out-degrees (OD), clustering coefficient (CC), and shortest paths (SP). For instance, we demonstrate that the average OD in networks of automatic translations consistently exceeds the values obtained for manual ones, and that the CC values of source texts are not preserved for manual translations, but are for good automatic translations. This probably reflects the text rearrangements humans perform during manual translation. We envisage that such findings could lead to better MT tools and automatic evaluation metrics.

Keywords: Complex networks; machine translation; network measurements; translation quality.

PACS Nos.: 89.75.Hc, 89.20.Ff, 89.75.Da.

1. Introduction

Complex systems represented by graphs are now referred to as complex networks,[1–4] which have become prominent since the discovery that many real-world networks cannot be modeled by random networks.[5,6] Watts and Strogatz, for instance, showed that the electric power grid in the USA and the network of actors in Hollywood movies are small-world networks.[7] Several networks such as the World Wide Web are scale-free,[8] thus possessing highly connected hubs that have been proven essential for the dynamics of phenomena modeled by the networks. There is growing evidence that the structure, function and evolution of these complex networks are governed by fundamental principles. For example, Costa *et al.*[9] have shown that Zipf's law can be explained by correlating structure and dynamics in a fully correlated network and with the properties of scale-free networks.

In view of the above, it is not surprising that complex networks have been used to analyze phenomena from various fields, stemming from social sciences to biology (an overview of the many applications of complex networks may be found in Refs. 1–4). Such applications include linguistics and computational linguistics, the fields to which our research may be inserted. Sigman and Cecchi[10] used complex networks to analyze lexical resources and Costa[11] studied human-induced words association, while Dorogovtsev and Mendes[12] modeled language evolution. Antiqueira *et al.*[13] evaluated network metrics to identify text authorship, Antiqueira *et al.*[14] and Pardo *et al.*[15] performed text and summary quality analyses, establishing correlations between network measures and text properties. Building upon these previous works, in this paper we use complex networks concepts to address some Machine Translation (MT) issues.

Machine translation, which basically converts a text produced in a source language into another text in a target language,[16] was probably the first non-numeric application of computers. The MT research area has experienced several translation methods and paradigms:[17] direct and indirect methods, with linguistic transference or not, following superficial or deep paradigms, using translation rules, stored translation templates/examples and/or statistics. The area started with handcrafted translation rules for each language pair intended to be automatically translated. This is a very expensive and time consuming task, which is still carried out today and produces very limited results. State-of-the-art results are obtained with relatively simple statistical models proposed by the IBM research group[18,19] and their several extensions (e.g., Refs. 20–23), for which large translation corpora are required to train the models. These methods result in statistical translation tables that are automatically and quickly learned. New language pair translators are easily produced from new bilingual corpora. Very good translations have been achieved even for very distant languages.

Despite its relatively long history, evaluation of MT has been controversial. Manual evaluation of MT results is obviously expensive, time consuming, and prone to human errors and inconsistencies. Automatic metrics have been proposed to avoid such problems and to allow for objectivity, reproducibility and fair comparison among different MT systems. The most representative automatic measure is BLEU (Bilingual Evaluation Understudy).[24,25] Basically, it counts the number of common n-grams present in both automatic translation and one or more reference (manual) translations. The higher the BLEU value (it falls between 0 and 1) the better the translation. The authors of the metric showed that it is as good as a human in ranking translations by their qualities.

NIST (National Institute of Standards and Technology) conducts annual contests on machine translation in order to measure and advance the state-of-the-art in the area. The contest adopts BLEU as the main evaluation metric and considers hard language pairs in its evaluation tasks, for example, Chinese-English and Arabic-English language pairs. Researchers in the area assume that a translation method that works for very different languages will also work for more similar languages. Particularly, Chinese and Arabic languages are challenging by their rich morphology and long-distance word reordering needed in translation. In recent years, the contest results (which are publicly available at NIST webpage[26]) show that statistical machine translation methods are the best ones and they significantly overcome traditional rule-based methods. It is interesting to see the results for 2005 contest, where statistical and rule-based systems competed: Google statistical translator[27] was the best one (and still is nowadays) with a BLEU score of 0.5131, while Systran,[28,29] the most famous rule-based translator, achieved a much worse BLEU score of 0.1079 for Arabic-English language pairs. Google translator improved Systran results by approximately 376%. For Chinese-English language pairs, the difference was lower: 0.3531 for Google against 0.1471 for Systran (Google was 140% better than Systran). Such numbers give an idea of the state-of-the-art results and the improvements achieved over the years with new and more advanced translation methods and more robust evaluation criteria.

Here we employ complex networks for representing source texts and their translations to assess cross-linguistic issues: the translation nature and quality for two language pairs. By nature, we mean whether the translation is produced by a human (manual) or by a MT system (automatic). In relation to quality, we aim at distinguishing good and bad automatic translations from different systems. We made comparisons between manually and automatically generated translations for Portuguese-English and Portuguese-Spanish pairs of texts (ranging from 20 to 50) collected from Revista Pesquisa FAPESP, a scientific magazine. Our approach is based on the use of network measurements that help characterize a piece of text. Word adjacency networks were created for each text, and the changes in network structure owing to translations of different nature and quality were evaluated. This is a language-independent approach that allows a translation to be analyzed by only observing the connectivity patterns between nodes. These patterns were captured by measurements usually employed in characterizing complex networks,[4] and were compared mainly using the Pearson correlation coefficient,[30] which gives the degree of similarity between the measurements of two networks. We show that structural differences between low and high-quality translations can be identified, mainly in Portuguese-Spanish pairs, in spite of the use of a linguistically superficial approach. Before discussing the results obtained in the experiments (Sec. 3), we describe in the next section the methodology employed. Sec. 4 concludes the paper with final remarks.

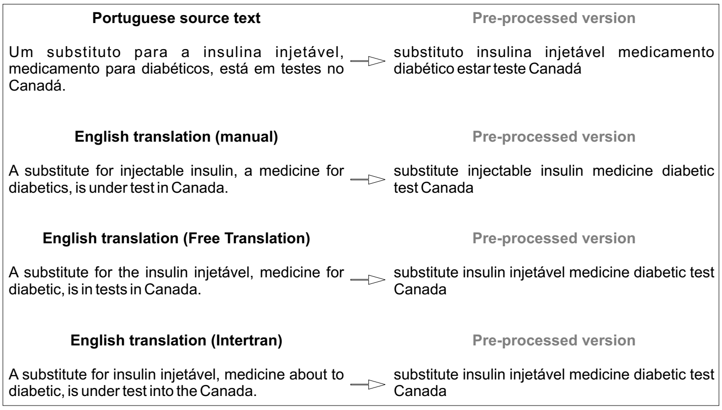


Fig. 1. Manual and machine translations of a Portuguese text into English (left column). The source sentences and their manual translation were taken from the magazine Pesquisa FAPESP.[31] The automatic translations were performed by Free Translation[32] and Intertran.[33] The automatically pre-processed versions of each excerpt (stopwords removed and remaining words lemmatized) are shown in the right column.

2. Methodology

In order to carry out the experiments, we must first represent text as graphs. We model both source and target texts as graphs according to the procedures established by Antiqueira *et al.*[14] Briefly, the texts were initially pre-processed to remove auxiliary verbs and functional words (stopwords). Then, the remaining words were lemmatized, i.e., they were reduced to their base form. An example of the pre-processing step is shown in Fig. 1. After this step, the set of unique words corresponds to the set of network nodes (vertices). Moreover, there is an edge between two nodes if the corresponding words are adjacent in the pre-processed text, even if they are in subsequent sentences. The edges weights represent the number of times those word pairs co-occur. The graph obtained was represented by a weighted matrix W of dimension $N \times N$, where $W(j,i)$ refers to the weight of the edge $i \rightarrow j$, i and j are adjacent words and N is the total number of distinct words in the text after the pre-processing step. Only the immediate adjacency was considered, but the results would probably not be strongly affected if further neighboring words were considered.[15]

The experiments were aimed at correlating graph metrics with manual and automatic translations in order to distinguish them. For some experiments, we had to map the nodes of the source network onto the target network, which required the two pieces of text to be word aligned (see Ref. 34 for an overview on text alignment). The alignment was performed using the system LIHLA,[35] which does a pairwise mapping between words in the source (S) and target (T) texts. When a group of n_s words in S is translated into a group of n_t words in T, LIHLA tries to create an alignment between the words n_s and n_t. When a single alignment pair considers more than one word in S or T, these words are grouped into one single node in the respective network. Some nodes may not belong to any alignment, and this occurs when the mapping is not possible, i.e., when a subset of words of S is not directly translated into words of T (or vice-versa). In additional experiments, we compared the results from the LIHLA automatic alignment process with a manual alignment, and the differences were not significant in terms of the metrics of the networks.

The first two graph measurements are based on the concept of degree. Note that these measurements are obtained separately for the source text and for the corresponding translation (the alignment is not considered yet). The weights of the links that come into a node i are used to compute the in-degree of i:

$$\mathrm{ID}(i) = \sum_{j=1}^{N} W(i,j)\,. \tag{1}$$

Similarly, the out-degree of a node i considers the links that emerge from i:

$$\mathrm{OD}(i) = \sum_{j=1}^{N} W(j,i)\,. \tag{2}$$

These measurements are also computed for the entire graph. In this case, the mean $\overline{\mathrm{ID}(i)}$ is the in-degree ID of the whole network, and the mean $\overline{\mathrm{OD}(i)}$ is the global out-degree OD. For the graphs considered, the average ID and OD are always the same.

Another frequently used network measurement is the clustering coefficient $\mathrm{CC}(i)$, which quantifies the level of connectivity between the neighbors of a node i. Let N_c be the number of nodes that receive a connection from i, then the total number of possible connections between these nodes is $N_c(N_c - 1)$. The actual number of directed connections between the N_c neighbors is denoted by B. The clustering coefficient of node i is the number of connections between the N_c nodes divided by the maximum possible number of links between them:

$$\mathrm{CC}(i) = \frac{B}{N_c(N_c - 1)}\,, \tag{3}$$

where $0 \le \mathrm{CC}(i) \le 1$. If $N_c = 0$ or $N_c = 1$ then $\mathrm{CC}(i) = 0$. The global clustering coefficient CC is taken as the mean $\overline{\mathrm{CC}(i)}$.

Shortest paths were also obtained for the networks representing the texts. A shortest path between two nodes i and j is defined as the minimum cost $d(i,j)$ necessary to reach node j starting from node i. The cost is defined as the sum of the weights of the edges included in the path. The quantity $d(i,j)$ is then used to define the mean minimum distance between a node i and the remaining nodes of the network:

$$\mathrm{SP}(i) = \frac{\sum_{j \neq i} d(i,j)}{N - 1}\,. \tag{4}$$

Moreover, the shortest paths $d(i,j)$ may be used to obtain a single measure for an entire network. In this case, any pair of distinct nodes is considered, as the equation below shows:

$$\mathrm{SP} = \frac{\sum_{i=1}^{N}\sum_{j\neq i} d(i,j)}{N(N-1)}\,. \tag{5}$$

We used three variations of Eqs. (4) and (5), regarding the computation of the cost $d(i,j)$:

(i) SP1(i) and SP1, which take a unity value for all edge weights,
(ii) SP2(i) and SP2, which consider the complement of the original weights $W_{\max} - W(i,j) + 1$ (where $W_{\max}$ is the maximum value of matrix W) and
(iii) SP3(i) and SP3, which take the inverted weights $1/W(i,j)$.

The first case removes the weights of the edges, and thus can be useful to assess the influence of the absence of weights on shortest paths. The other two cases are used to give priority to the stronger and more important edges (those with higher weights), since in the usual way these edges would be avoided in the shortest paths. Finally, when there is no path between nodes i and j, then $d(i,j) = \infty$. Since the measures of Eqs. (4) and (5) are used to quantify the overall connectivity of a network, we had to define useful values of the proximity $d(i,j)$ for this particular case, instead of infinity. Thus, when no path exists between i and j, we consider that:

(i) $d(i,j) = N$, for SP1(i) and SP1, and
(ii) $d(i,j) = N\bar{W}$, for SP2(i), SP2, SP3(i) and SP3 ($\bar{W}$ is the mean weight of all network edges).

The corpus used in the experiments comprised texts published in Revista Pesquisa FAPESP,[36] a magazine published in Portuguese, English and Spanish. In the experiments that employ text alignment and individual measurements for each node of a graph (Secs. 3.2 and 3.3), 20 Portuguese texts and the corresponding manual translations into English and Spanish were used, resulting in 40 aligned text pairs. In the experiment that uses only global (average) graph measurements without text alignment (Sec. 3.1), 50 texts in Portuguese were employed for English translations and 20 for Spanish translations. Each text had approximately 10 sentences. Automatic translations were obtained from the following systems: Free Translation,[32] Google Translator[27] and Intertran[33] for Portuguese-English translations; Intertran and Apertium[37] for Portuguese-Spanish translations. An example of the output generated by Free Translation and Intertran is shown in Fig. 1.

3. Results and Discussion

3.1. *Comparing average metrics for translated texts*

Figure 2 shows a histogram for the number of texts whose average out-degree (OD) for the MT text exceeds the value for the manually translated text. The original texts were in English, and the translated versions into Portuguese were obtained

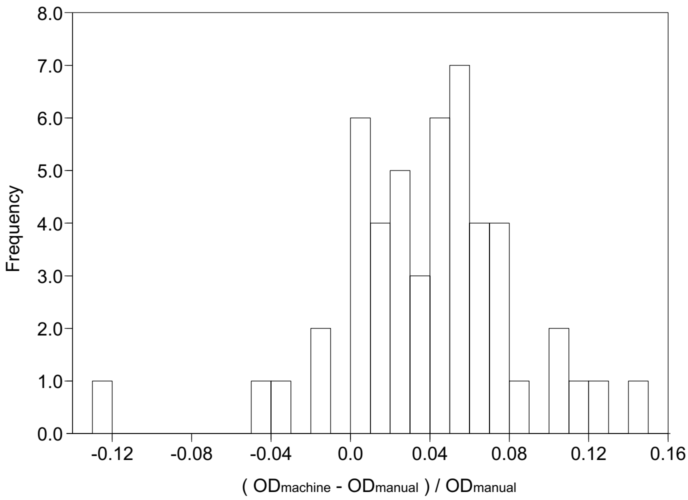


Fig. 2. Histogram of the relative change in the average OD between a MT text and a manually translated text ($(\mathrm{OD}_{\mathrm{machine}} - \mathrm{OD}_{\mathrm{manual}})/\mathrm{OD}_{\mathrm{manual}}$) from English into Portuguese.

either manually (Human) or with the Google translator, which is essentially a version of Systran.[28,29] The average OD was therefore calculated for the networks built from the translated texts only, without using text alignment. It is readily seen that the average OD is consistently higher for the MT text, even though the difference is not large. Approximately 90% of all texts fell into this pattern. Considering OD, it was also possible to distinguish the two types of translations, since 70% of the Portuguese texts have higher degree than the corresponding texts in Spanish. Furthermore, in subsidiary experiments we noted that other metrics could also be used to distinguish between manual and machine translations, as it is in the case of the shortest paths SP (results not shown). In fact, a combined use of these metrics along with text alignment would allow distinction between the two types of translation with high accuracy (see next section).

3.2. *Comparing mapped networks of aligned texts*

One important issue in evaluating the characteristics of translated texts is to compare whether the corresponding networks for the source and target texts can be mapped onto each other. This requires the analysis of individual node measurements and aligned texts. The networks were compared by plotting the measurements extracted from the source and target networks, from which the Pearson correlation coefficient[30] and the slope (angular coefficient) were obtained. Our hypothesis is that if the measurements of the target network are similar to the ones from the source network, their scatter-plot should be approximately linear with slope close to 1, i.e., with angle close to 45 degrees (see Fig. 3 for a scatter-plot using OD for a Portuguese-Spanish pair). It is possible to measure the strength of a linear

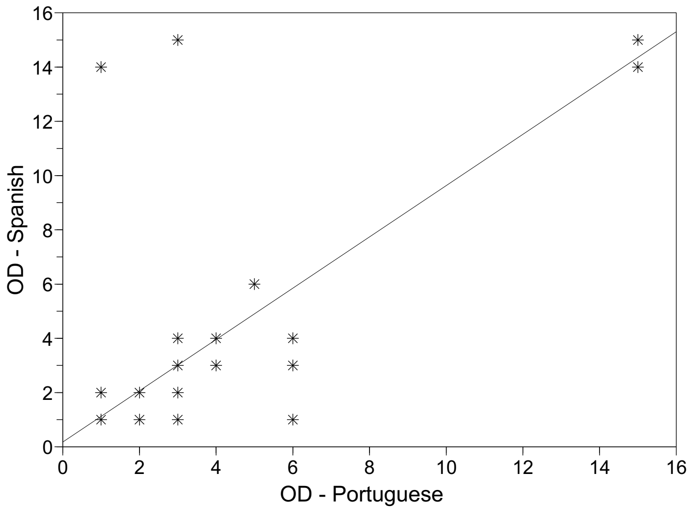
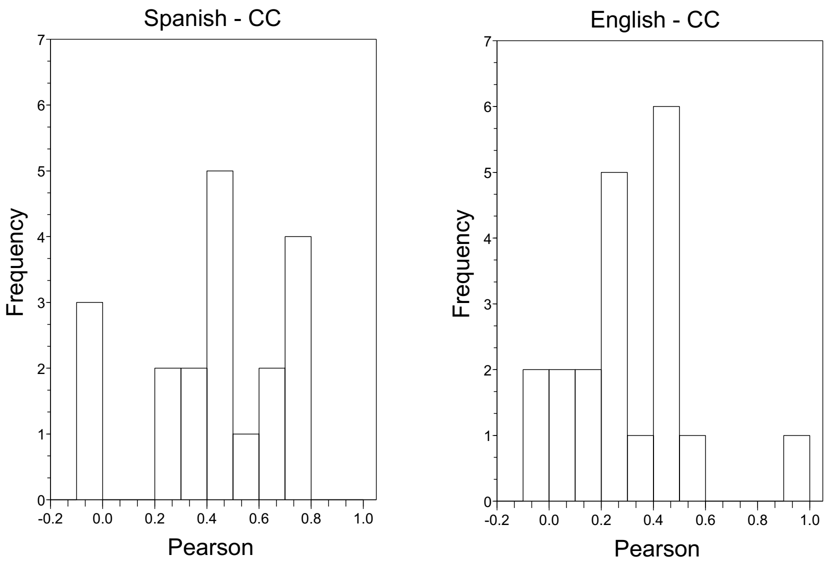


Fig. 3. Scatter-plot obtained from the alignment between a source text in Portuguese and its translation into Spanish generated by Apertium. The horizontal axis represents the measurement OD obtained for each node of the network created from the translated text (in Portuguese). The vertical axis represents the same measurement obtained for the source text (in Spanish). The Pearson correlation coefficient is high (0.958), with angular coefficient of 0.954.

relationship with the Pearson coefficient, whose absolute value is close to 0 when the correlation is far from linear or when there is no correlation at all, and it is close to 1 when the correlation is close to linear. When the Pearson coefficient is close to 1, we take the slope of the straight line that fits the data to check whether the measurements are preserved from one network onto another. If preservation occurs, the slope is close to 1.

With this methodology, we found that the clustering coefficient (CC) and shortest paths (SP) are not preserved in manually translated texts for neither Portuguese-Spanish nor Portuguese-English translations (Fig. 4; the results for SP were not included), probably because these parameters are affected by text rearrangement at the supra-sentential level, which is common in human translation. In contrast, CC values are preserved in good-quality MT but not in low-quality MT (Fig. 5). Figures 4(a) and 4(b) show small values for the Pearson coefficient for CC for manually translated texts from Portuguese into Spanish and Portuguese into English, respectively. Figures 5(a) and 5(b), on the other hand, show that high Pearson coefficients are obtained for the good-quality MT system for Spanish (Apertium), but not for a low-quality system (Intertran). In addition, similar results were obtained for texts translated automatically from Portuguese into English, and the behavior observed for CC also applied to SP values (results not shown). It will be discussed later on that one important feature of the good-quality MT systems is the preservation of the network topology (captured using network measurements) of the source text in the translated text.

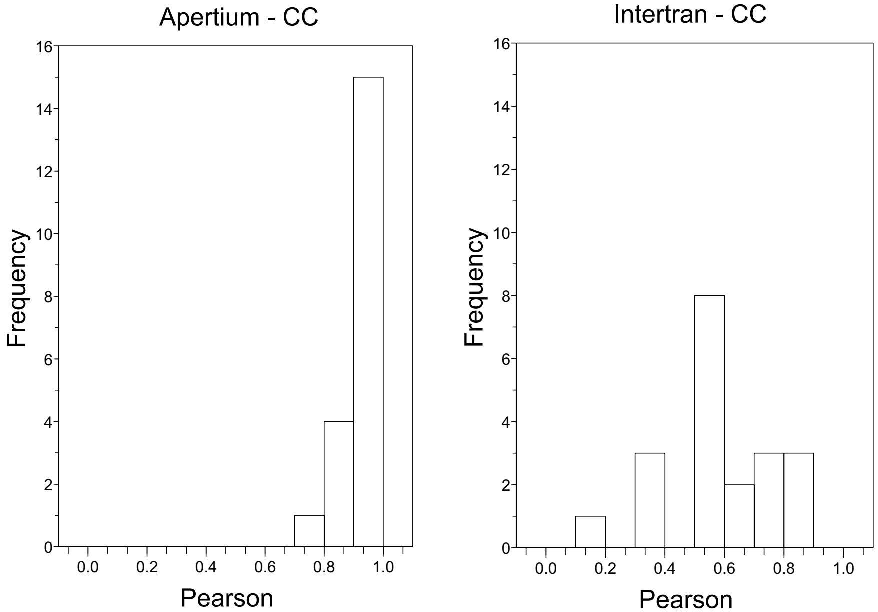


(a) (b)

Fig. 4. Pearson coefficient for CC when mapped from a source text in Portuguese and the target text in (a) Spanish and (b) English, both obtained by human translation.

Language dependence appeared in the analysis of in- and out-degrees, which are largely preserved in good translations for Portuguese-Spanish texts. Network nodes were essentially mapped in a one-to-one basis from one language to the other, with practically the same values for ID and OD. The scatter plot of Fig. 3 illustrates the good correlation between OD for Portuguese and Spanish (translated) texts. This applied to both human translated texts and texts produced with a good-quality MT tool, as indicated by the high Pearson coefficients in Fig. 6. This is not true for the low-quality MT tool, perhaps because mapping of grammatical structures was inaccurate, which is illustrated in the top panel of Fig. 6. Furthermore, in the bottom panel, one can see that the angular coefficient is also close to 1 for all translations, probably deriving from the very similar sentence structures of the two languages. In Figs. 3 and 6 we show only the results for OD, since essentially the same results were obtained for the ID measurement.

In the English-Portuguese texts, again ID and OD values were preserved for the quality MT, denoted by the high Pearson coefficient in Fig. 7 (results shown for OD only). However, in contrast to the Spanish-Portuguese pair, the Pearson coefficient for the human translation was lower, closer to a lower-quality MT (Intertran). Furthermore, the correlation between pairs of aligned networks for a given text was not as good as in the case of Portuguese-Spanish texts, therefore reflecting the larger structural differences between English and Portuguese. Interestingly, in human translated texts ID and OD values were consistently lower for English texts, which result from the differences in word usage for the two languages. The latter



(a) (b)

Fig. 5. Pearson coefficient for CC when mapped from a source text in Portuguese and the target text in Spanish, where the latter was obtained with a good-quality MT system (a) and with a low-quality MT system (b).

finding is illustrated in the bottom panel of Fig. 7, where the angular coefficients are lower than 1 for the human translation. For the good-quality MT the angular coefficients are again close to 1, as the automatic system apparently does not impose sentence rearrangements that would affect either ID or OD.

Table 1 summarizes the results for Spanish and English translations regarding topological preservation with automatic alignment. The Pearson coefficient, along with the angular coefficient of the straight line obtained by linear regression (both on average), are shown for the measurements OD, ID and CC. Preservation is higher for the Apertium and human translations (Spanish), when *OD* and ID are considered, since both coefficients are close to 1. For lower quality Spanish translations (Intertran) the same pattern is not verified, so the preservation of the measurements does not occur. For the CC cases, no translation seems to preserve the measurement. When considering English translations, the closer we get to measurement preservation is in the Free Translation case, with Pearson coefficients near 0.8. The human English translations also appear to preserve OD and ID, but with lower coefficients than Spanish translations. In general, Spanish translations of higher quality (human and Apertium) tend to preserve OD and ID, whereas for English translations this tendency is less strong.

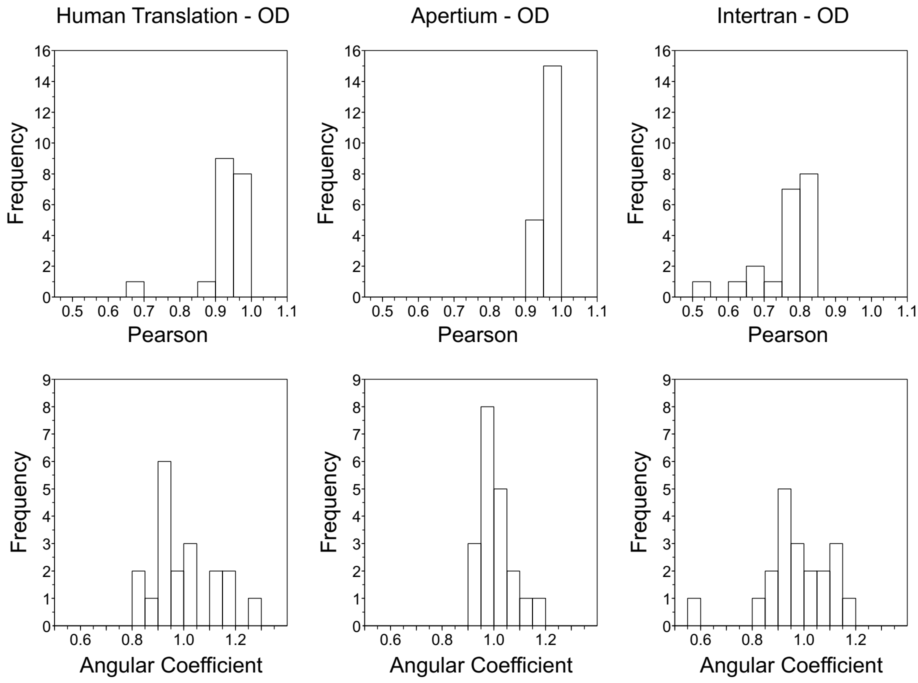


Fig. 6. Pearson coefficient (top panel) and angular coefficient (bottom panel) obtained for the metric out-degree (OD) for texts translated from Portuguese into Spanish. Note the high Pearson coefficients for the human and good-quality MT translations, whereas lower Pearson coefficients were obtained for the low-quality MT system. The angular coefficient was close to 1 in all cases.

Table 1. Pearson and angular coefficients obtained, on average, for the experiments with alignments. For Spanish translations (human and Apertium) the measures OD and ID are likely to be preserved in the translations, since both coefficients (in bold) approach 1. Data on shorted paths (SPs) were omitted because its behavior is similar to that of CC.

Type of Translation	Avg. Pearson Coefficient			Avg. Angular Coefficient		
	OD	ID	CC	OD	ID	CC
Spanish — Human	**0.93**	**0.94**	0.44	**0.99**	**1.00**	0.44
Spanish — Apertium	**0.97**	**0.98**	0.93	**1.01**	**1.03**	0.85
Spanish — Intertran	0.76	0.80	0.58	0.98	1.48	1.01
English — Human	0.79	0.83	0.31	0.81	0.82	0.33
English — Free Translation	0.81	0.89	0.59	0.96	0.99	0.57
English — Intertran	0.88	0.89	0.71	0.82	0.83	0.98

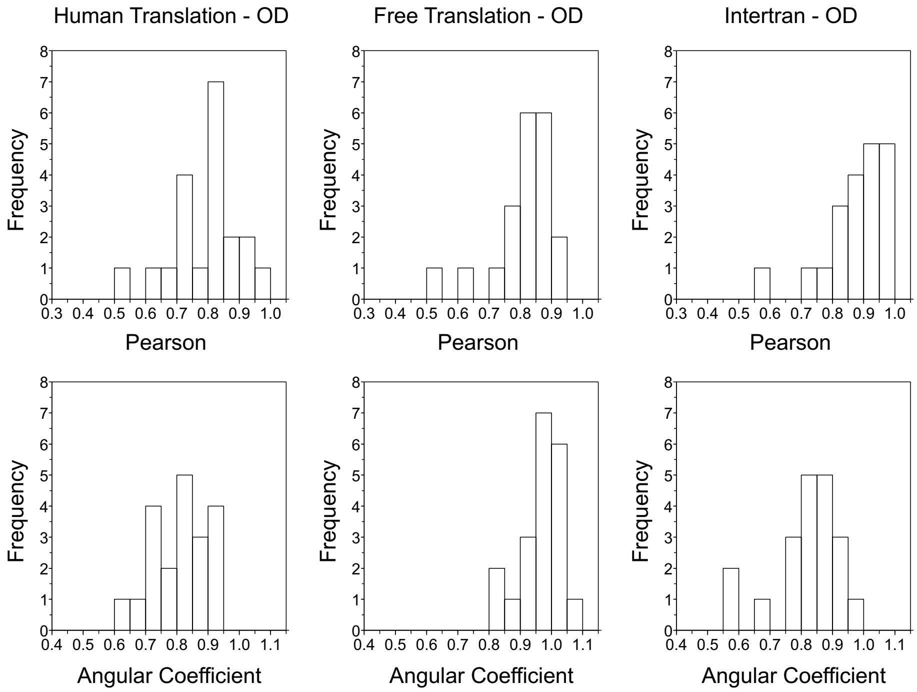


Fig. 7. Pearson coefficient (top panel) and angular coefficient (bottom panel) obtained for the metric out-degree (OD) for texts translated from Portuguese into English. Degrees are only preserved for quality MT.

3.3. *Distinguishing between manual and machine translations*

For the purpose of separating different types of translations, we used the network measurements in a clustering experiment with the visualization software PEx (Projection Explorer),[38,39] which is capable of creating projections into two dimensions for n-dimensional data. We analyzed groups of texts according to its classes (namely human or automatic translations, for both English and Spanish). Starting from a graphical representation created by PEx, it is possible to distinguish regions that belong to specific translation classes, as can be seen in Fig. 8, in which each node in the graph represents one translation pair Portuguese-Spanish, and the x and y axes indicate the positioning of the text pairs obtained from projection of multidimensional values (network measurements). This plot was created by PEx using six measures (Pearson and angular coefficient for ID, OD and CC) for each Portuguese-Spanish translation pair and was identified as small circles. The algorithm used was the Interactive Document Map, along with a City Block distance.[40] In contrast to the Euclidean distance, which is the square root of the squared differences of coordinates, the City Block distance is the sum of the absolute differences of coordinates. The edges are produced with the KNN (k-Nearest Neighbor) algorithm, which puts in the same cluster the k nodes with the shortest City Block distances. In our

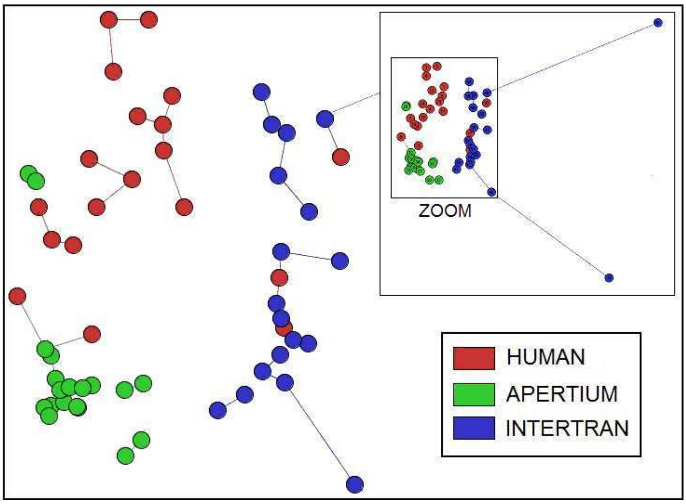


Fig. 8. PEx plot for three types of Portuguese-Spanish translations. The measurements used for each translation pair were the ones obtained by assessing topology preservation with text alignment, i.e., the Pearson and angular coefficients for OD, ID and CC. The projection algorithm of PEx allowed a good separation between the three types of translation.

experiment, $k = 2$, and therefore the edges connect the points with their nearest neighbors in a bidimensional plot.[41]

The three classes considered in the experiment are (i) human translation, (ii) Apertium translation, and (iii) Intertran translation, which can be spatially separated from each other in Fig. 8 with reasonably good accuracy (note that higher distances between translations indicate stronger differences in translation — as far as the six measures adopted are concerned). Thus, this 2D plot may be used to classify the nature and quality of Portuguese-Spanish translations (the clustering for Portuguese-English translations was less clear — results not shown).

4. Conclusions

The metrics obtained from complex networks used to model texts have been correlated to the quality of machine translation systems and employed to distinguish between human and machine translation for the Portuguese-Spanish and Portuguese-English pairs of languages. Among the main results, we observed that the values for the cluster coefficient (CC) and shortest paths (SP) are not preserved in manually translated texts for neither pairs of languages, probably because these parameters are affected by text rearrangement at the supra-sentential level, which is common in human translation. In contrast, CC values were preserved in quality MT but not in poor MT. Language dependence appeared in the analysis of in- and out-degrees, which were largely preserved in good translations. For Portuguese-Spanish texts, network nodes were essentially mapped in a one-to-one basis from one language to the other, with practically the same values for in- (ID) and out-degrees

(OD). This applied to both human translated texts and texts produced with a quality MT tool. It was not the same for the poor quality MT tool, perhaps because mapping of grammatical structures was inaccurate. In the Portuguese-English texts, again ID and OD values were preserved for quality MT, and not for a poor MT system or human translation, thus reflecting some language dependence. Indeed, the correlation between the two corresponding networks for a given text was not as good as in the case of Portuguese-Spanish texts, owing to the larger structural differences between English and Portuguese. In human translated texts ID and OD values were consistently lower for English texts, probably due to the differences in word usage for these languages. Such differences in these parameters were not observed in the texts generated by the quality MT system; apparently, in spite of the good quality in translation, treating the sentence at the sentence level in the MT tool has hindered the generation of more concise texts.

One may also infer that some differences in network topologies are directly related to intrinsic differences in the structures of the languages analyzed, with negligible effects from translation. This is consistent with the finding that network topologies were closer for the texts in Spanish and Portuguese than for the English-Portuguese pair. On the basis of the results presented here, we may now extend the study of modeling texts in a source and in a target language to identify topological differences that are only caused by poor translation, which should be distinguished from those intrinsic to language differences. This would increase the ability of a computational system using complex networks to assess the quality of translations.

As a proof-of-principle, we employed a clustering algorithm within a visualization software, PEx, to distinguish between manual and machine translations, which may also allow distinction between low and high-quality MT systems. Based on these results, one may envisage the development of a robust machine-learning system using complex network metrics to assess the quality of MT systems, probably with regression algorithms that may allow one to assign scores to the quality of the translation.

Acknowledgments

The authors acknowledge the financial support from FAPESP and CNPq, and are grateful to Dr. Rosane Minghim for the use of the PEx suite. Luciano da F. Costa thanks FAPESP (05/00587-5) and CNPq (308231/03-1).

References

1. R. Albert and A. L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
2. M. E. J. Newman, *SIAM Rev.* **45**, 167 (2003).
3. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
4. L. da F. Costa, F. A. Rodrigues, G. Travieso and P. R. Villas Boas, *Adv. Phys.* **56**, 167 (2007).
5. P. Erdös and A. Rényi, *Publ. Math. Debrecen* **6**, 290 (1959).

6. P. J. Flory, *J. Am. Chem. Soc.* **63**, 3083, 3091, 3096 (1941).
7. D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
8. A. L. Barabási and R. Albert, *Science* **286**, 509 (1999).
9. L. da F. Costa, O. Sporns, L. Antiqueira, M. G. V. Nunes and O. N. Oliveira Jr., *Appl. Phys. Lett.* **91**, 054107 (2007).
10. M. Sigman and G. A. Cecchi, *Proc. Natl. Acad. Sci. USA* **99**, 1742 (2002).
11. L. da F. Costa, *Int. J. Mod. Phys. C* **15**, 371 (2004).
12. S. N. Dorogovtsev and J. F. F. Mendes, *Proc. Royal Soc.* **268**, 2603 (2001).
13. L. Antiqueira, T. A. S. Pardo, M. G. V. Nunes, O. N. Oliveira Jr. and L. da F. Costa, in *Proc. Fourth Workshop in Information and Human Language Technology (TIL'06)* (2006).
14. L. Antiqueira, M. G. V. Nunes, O. N. Oliveira Jr. and L. da F. Costa, *Physica A* **373**, 811 (2007).
15. T. A. S. Pardo, L. Antiqueira, M. G. V. Nunes, O. N. Oliveira Jr. and L. da F. Costa, in *Proc. International Conference on Communications, Circuits and Systems* (2006), pp. 2678–2682.
16. D. J. Arnold, L. Balkan, S. Meijer, R. L. Humphreys and L. Sadler, *Machine Translation: An Introductory Guide* (Blackwells-NCC, London, 1993).
17. B. J. Dorr, P. W. Jordan and J. W. Benoit, *Adv. Comput.* **49**, 2 (1999).
18. P. F. Brown, J. Cocke, S. Della Pietra, V. J. Della Pietra, F. Jelinek, J. D. Lafferty, R. L. Mercer and P. S. Roossin, *Comput. Linguist.* **16**, 79 (1990).
19. P. F. Brown, S. Della Pietra, V. J. Della Pietra and R. L. Mercer, *Comput. Linguist.* **19**, 263 (1993).
20. F. J. Och, C. Tillmann and H. Ney, in *Proc. Joint Conference on Empirical Methods in Natural Language Processing and Very Large Corpora* (1999), pp. 20–28.
21. F. Och, D. Gildea, S. Khudanpur, A. Sarkar, K. Yamada, A. Fraser, S. Kumar, L. Shen, D. Smith, K. Eng *et al.*, in *Proceedings of HLT/NAACL* (2004), pp. 161–168.
22. K. Yamada and K. Knight, in *Proc. Meeting of the Association for Computational Linguistics* (2001), pp. 523–530.
23. P. Koehn, F. J. Och and D. Marcu, in *Proc. HLT/NAACL* (2003), pp. 127–133.
24. K. Papineni, S. Roukos, T. Ward and W. J. Zhu, in *Proc. Meeting of the Association for Computational Linguistics* (2002), pp. 311–318.
25. R. Soricut and E. Brill, in *Proc. Meeting of the Association for Computational Linguistics* (2004), p. 613.
26. NIST Open Evaluation of Machine Translation, http://www.nist.gov/speech/tests/summaries/index.htm.
27. Google Translator, http://www.google.com/language_tools?hl=en.
28. P. Senellart and J. Senellart, in *Proc. XML Conference e Exposition* (2005), pp. 1–14.
29. Systran, http://www.systransoft.com.
30. J. Neter, M. H. Kutner, C. J. Nachtsheim and W. Wasserman, *Applied Linear Statistical Models* (Irwin, Chicago, 1996).
31. Revista Pesquisa FAPESP, 86 (2003).
32. Free Translation, http://www.freetranslation.com.
33. Intertran, http://intertran.tranexp.com/Translate/result.shtml.
34. J. Véronis, *Parallel text processing: Alignment and use of translation corpora* (Kluwer Academic Publishers, Dordrecht, 2000), Chap. 1, pp. 1–24.
35. H. M. Caseli, M. G. V. Nunes and M. L. Forcada, *Procesamiento Del Lenguaje Natural* **35**, 237 (2005).
36. Revista Pesquisa FAPESP, http://www.revistapesquisa.fapesp.br.
37. Apertium, http://xixona.dlsi.ua.es/prototype/pt/.

38. G. P. Telles, R. Minghim and F. V. Paulovich, *Comput. Graph.* **31**, 327 (2007).
39. PEx, http://www.lcad.icmc.usp.br/∼paulovic/pex/.
40. E. F. Krause, *Taxicab Geometry* (Dover, New York, 1986).
41. B. V. Dasarathy, *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques* (IEEE Computer Society Press, Los Alamitos, 1990).