Automatic Alignment of News Texts and their Multi-document Summaries: Comparison among Methods

Verônica Agostini, Roque Enrique López Condori, Thiago Alexandre Salgueiro Pardo

Interinstitutional Center for Computational Linguistics (NILC)
Institute of Mathematical and Computer Sciences, University of São Paulo
{agostini, rlopez, taspardo}@icmc.usp.br

Abstract. Aligning texts and their multi-document summaries is the task of determining the correspondences among textual segments in the texts and in their corresponding summaries. The study of alignments allows a better understanding of the multi-document summarization process, which may subsidize new summarization models for producing more informative summaries. In this paper, we investigate some approaches for text-summary sentence alignment, including superficial, deep and hybrid approaches. Our results show that superficial approaches may obtain very good results.

Keywords: Sentence alignment, multi-document summarization

1 Introduction

Multi-document summarization aims at producing a summary*, i.e.,* a condensed document, from multiple documents on the same topic. This is very useful to present a reasonable amount of information to a reader, given the large amount of information available on the web, and to organize what is being conveyed.

There are several methods for multi-document summarization (see, *e.g.*, [19], [24]) with most of them focusing on producing extractive summaries (by copying and pasting segments from the source texts, without rewriting them). In general, although useful results are already available, the quality of extracts are still far away from those produced by humans (which usually produce abstracts, by presenting different linguistic material in relation to the source texts) and several problems remain, as dangling anaphors, presence of redundant information, inadequate temporal ordering of events, and inappropriate treatment of contradictory information, among others.

In this scenario, the alignment of source texts/documents and their multi-document summaries is an important task. In this case, to align is to find correspondences among textual segments with different granularity levels, which means that the alignment may occur between single words, n-grams, sentences, paragraphs and even entire documents. In summarization, this task may provide information about human summarization methods, which may help understanding the nature of the phenomenon and improving the automatic process by subsidizing the creation of new summarization rules and models.

In this paper, we report our investigation of some text-summary sentence alignment methods, evaluating both superficial and deep methods. Superficial methods are generally easier to implement and to use, while methods that use more linguistically motivated assumptions are usually more difficult to produce, especially because it is necessary to obtain language-dependent resources. In our work, we used a discourse theory (Cross-document Structure Theory (CST) [26]) in order to linguistically enhance the process. Furthermore, we combine superficial and deep information in a hybrid approach using machine learning. We run our experiments on news texts written in Brazilian Portuguese, which encompass general (day by day) language and allow us to check the robustness of the methods. Our results show that superficial approaches are enough to obtain very good results.

To the best of our knowledge, this is the first attempt of aligning texts and summaries for Brazilian Portuguese. In general, our contribution remains on developing and evaluating superficial, deep and hybrid methods.

The rest of this paper is organized as follows: in Section 2, we present the alignment task and the related work; in Section 3, we present a *corpus* that was manually annotated, used here as a gold standard for the evaluation of the alignment methods; in Section 4, we present our methods to obtain the alignments; in Section 5, we show the obtained results; and, in Section 6, we make some final remarks.

2 Basic Concepts and Related Work

The alignment originated in the machine translation area (see, *e.g.*, [11], [27]), in which the alignment occurs among textual segments (words, phrases or sentences, usually) of a document and its translated version. They are useful for both producing bilingual dictionaries and allowing the development of the current state of the art in statistical machine translation. The alignments may be 1-1, when 1 segment in a text is aligned to 1 segment in the other; 1-N (including 1-2, 1-3, etc.), when 1 segment in a text is aligned to more than 1 segment in the other; and N-N, when more than 1 segment in a text is aligned to more than 1 segment in the other. Furthermore, the alignments may be 1-0, when some information is new in a text and has no correspondence in the other. A sentence alignment example is shown in Figure 1, when two sentences, in English and Portuguese, are aligned.

Source sentence	Target sentence
I would like to humbly thank you all from the bottom of my heart.	*Eu gostaria de agradecer humildemente a todos do fundo do meu coração.*

Fig. 1. Alignment example in translation

Similarly to translation, alignment in summarization consists in finding the correspondence of segments in the source text(s) and in the summary. Figure 2 shows an example, when sentences of two texts are aligned to 1 sentence in a multi-document summary (in a 2-1 alignment). As one may see, sometimes two aligned segments (sentences, in this case) have many words in common, making it easier to find the alignment.

Sentences from the source texts	Sentence in the summary
A tocha passará por vinte países, mas o Brasil não estará no percurso olímpico. (The torch will pass through twenty countries, but Brazil will not be on the Olympic journey.)	*O Brasil não fará parte do trajeto de 20 países do revezamento da tocha.* (Brazil is not part of the path of 20 countries of the torch relay.)
O Brasil não faz parte do trajeto da tocha olímpica. (Brazil is not part of the path of the Olympic torch.)	

Fig. 2. Alignment example in summarization – a simple case

In other cases, it is a harder task, as exemplified in Figure 3 (in a 1-1 alignment), since it is necessary to have world knowledge that "storing supplies" is a way of "preparing for a hurricane".

Sentence from the source text	Sentence in the summary
Na Jamaica, muitos ***estocaram alimentos, água, lanternas e velas****.* (In Jamaica, many people **stored food, water, flashlights and candles**.)	*Vários moradores e turistas nas regiões, inclusive brasileiros, foram retirados dos locais, enquanto outros estão* ***se preparando para a passagem do furacão****.* (Many residents and tourists in the regions, including Brazilians, were evacuated from the places, while others **are preparing themselves for the hurricane**.)

Fig. 3. Alignment example in summarization – a difficult case

There are some automatic methods proposed in the literature to find the alignments. For example, [2] used the Term-length Term-frequency (TLTF) algorithm to align sentences based on the words that these segments have in common. [21] developed an algorithm, to create an extract from a document, in which the idea is to iteratively exclude sentences from the text until the resultant extract is the most similar to the abstract. The idea of their works is to obtain extracts, which would automatically encompass the alignment information (since whole sentences are simply taken from the texts). Another example in summarization is the work of [17], in which the authors proposed a method that uses Hidden Markov Model (HMM) to perform the alignments. The HMM models heuristics based on cut-and-paste operations performed by the human summarizers, which were found by the authors. [9] and [10] use a HMM with the Expectation Maximization algorithm. The HMM is constructed using a generative history, which models how a summary is produced from a text. Other authors perform the alignments among similar texts. Their works may be replicated to the summarization area because documents and their summaries may be considered similar texts. [13] and [14] used machine learning techniques, using features as word co-occurrences, noun phrase matching, WordNet synonyms, common semantic classes for verbs, and shared proper nouns, among others. In [3], similar paragraphs are grouped and a cosine similarity measure is used to align their sentences. In multi-document summarization, we may highlight the work of [15], in which the authors used the dependency tree path using a similarity measure to perform the alignments between a single document and its summary as well as among a set of documents and their summary.

Basically, the authors used news texts in their experiments. Some perform the alignments among sentences or clauses ([2, 3, 15, 21]), others use n-grams ([17]), phrases and words ([9, 10]) or even paragraphs ([3, 13, 14]). The best results are synthesized in Table 1.

Many authors produce manual aligned corpora to be the basis for comparisons with their automatic alignment methods. In order to evaluate our methods, which will be presented in Section 4, we also created a manual aligned *corpus*, which is composed of news texts, as we describe in what follows.

Table 1. Synthesis of the main results in the literature

Work	Granularity	Precision	Recall	F-measure
[21]	clause	74.27%	80.29%	76.47%
[21]	sentence	77.45%	80.06%	78.15%
[17]	n-gram	81.50%	78.50%	79.10%
[13,14]	paragraph	49.30%	52.90%	51.00%
[3]	clause, paragraph	76.90%	55.80%	-
[9,10]	word, phrase	52.20%	71.20%	60.60%
[15]	sentence (single document)	-	-	97.70%
[15]	sentence (multi-document)	-	-	80.80%

3 The *corpus*

CSTNews [5] is a *corpus* consisting of 50 clusters of news texts written in Brazilian Portuguese. Each cluster contains 2 to 3 news texts on the same topic, their multi-document summaries (automatically and manually created), and many other annotations (for example, CST [26] and RST (Rhetorical Structure Theory) [20]). The clusters have, on average, 42 sentences (10 to 89 sentences), and the multi-document summaries have, on average, 7 sentences (3-14). The news texts were collected from online news agencies. We used the news texts and the human multi-document summaries (abstracts) for performing the manual alignment.

The manual alignment was conducted by 2 computational linguists, after a training phase, in daily sections of 1 to 2 hours. The task resulted in some alignment rules, which are presented in detail in [1]. The rules are important to guarantee that the alignment annotation will be consistent enough to be followed. Another alignment example may be seen in Figure 4, where two sentences are aligned because they convey the same topic, which is the traffic jam, and the summary sentence provides specific information about it, *i.e.*, the time of the traffic jams and their extension.

Overall, after the *corpus* annotation, most of the alignments showed to be of type 1-2, *i.e.*, 1 sentence in the summary aligned to 2 sentences in the texts, which was expected. This was followed by the 1-1 alignment type. The distribution of alignment types may be seen in Table 2.

Sentence in the summary
A Companhia de Engenharia de Tráfego (CET) anunciou que o índice de congestionamento era de 54 quilômetros às 8h, 113 km às 9h e 110 km meia hora depois, valores bem acima das médias para os horários, que eram de 36, 82 e 76 quilômetros respectivamente, mas não havia registro de acidentes graves, apesar de haver feridos. (The Traffic Engineering Company (CET) announced that the extent of traffic jam was 54 km at 8am, 113 km at 9am and 110 km at half an hour later, values much above the averages for those times, which were 36, 82 and 76 km respectively, although there were no serious accidents, but with some injured people.)
Sentence in the source text
Com o asfalto molhado, o trânsito ficou mais lento e o congestionamento ficou o dobro da média. (With the wet pavement, the traffic got slow and the traffic jam the double of the average.)

Fig. 4. Alignment example based on rules 1 and 6

Table 2. Alignment types in the *corpus*

Alignment types												
1-0	**1-1**	**1-2**	**1-3**	**1-4**	**1-5**	**1-6**	**1-7**	**1-8**	**1-9**	**1-10**	**1-11**	**1-12**
2	71	91	72	33	37	13	6	6	1	1	2	1

We obtained some extreme cases, like 1-0 types (due to summary sentences that were not annotated because these were information inferred by the human summarizers and were not in the source texts), and 1-12 types, when one sentence in the summary synthesizes 12 other sentences in the texts.

Besides that, we computed the agreement among the annotators, using the kappa agreement measure [6, 7], considering 5 clusters of the *corpus*. We obtained 0.831 of agreement, which ranges from 0 (no agreement) to 1 (total agreement). This value is very good, and reflects the reliability of the annotation. This value also indicates that, although the alignment process is subjective, the annotators know it very well.

Furthermore, we also labeled all the detected alignments, in a task that we called typification of the alignments (as shown in details in [4]). Basically, we annotated the transformations that were performed to summarize the documents. There are two major groups of types: form and content. Two aligned sentences must have 1 form type, which may be: (i) identical, when the two sentences are the same; (ii) partial, when they have many words in common, but are note identical; or (iii) different, when they have few words in common. The alignment pair may have more than one content type, which may be: (i) specification, when the summary sentence contains some more specific information related to the document sentence; (ii) generalization, when the summary sentence contains some information that is a generalization related to the document sentence; (iii) contradiction, when the sentences present some information that is contradictory; (iv) inference, when the summary sentence expresses information that was inferred from the document sentence; (v) neutral, when there is some information that does not result from a transformation; and (vi) other, when the annotators do not agree with the previous alignment. This disagreement may happen because the alignment is a subjective task. Besides that, we annotated information that is related to the onomastics (toponomastics, related to place names, and anthroponomastics, related to person names). In general, kappa measures were 0.717, 0.318 and 0.452 for form, content and joint form and content annotations, respectively, which are good considering the difficulty of the task.

One example of alignment typification may be seen in Figure 5, which shows the labels: (i) partial, because the sentences have some words in common; (ii) neutral, because some information in the pair does not have a transformation; (iii) generalization, because the names of the states were generalized to "many states" in the summary sentence; and (iv) toponomastics, because there are names of places in the document sentence.

Sentence in the summary	Types of the alignment	Sentence in the source text
Mais de 300 policiais federais de vários estados participaram das buscas e prisões durante a operação. (More than 300 federal police officers from many states took part in the searches and arrests during the operation.)	(i) **Partial** form overlap (ii) **Neutral** (iii) **Generalization** of content (iv) **Toponomastics**	*A PF divulgou que mais de 300 policiais federais do Amazonas, Distrito Federal, Mato Grosso, Acre e Rondônia fazem parte das investigações da "Operação Dominó".* (The FP reported that more than 300 federal police officers from Amazonas, Distrito Federal, Mato Grosso, Acre and Rondônia took part in the "Operação Dominó".)

Fig. 5. Typification example

In the CSTNews, the most frequent labels were partial and neutral for form and content labels, respectively, as Table 3 shows. Onomastics was a rare phenomenon.

Table 3. Occurrence of alignment types in the corpus

Category	Type	Number of occurrences	Percentage (%)
Form	Partial	871	86.06
	Different	82	8.10
	Identical	59	5.83
Content	Neutral	955	94.36
	Generalization	80	7.90
	Specification	47	4.64
	Contradiction	37	3.65
	Inference	29	2.86
	Other	6	0.59
Onomastics	Anthroponomastics	23	2.27
	Toponomastics	4	0.39

4 Methods

We followed 3 approaches to perform the automatic alignment. The first approach encompasses 3 methods that use superficial information about the texts, like the sentence position and the number of words in the sentences; the second approach is more linguistically motivated and uses a discourse theory; and the third one is a hybrid approach that combines features from the two approaches before in a machine learning solution. Each approach is described in what follows.

4.1 The Superficial Approach

The first approach consists in three superficial methods that may be used together or separately. The methods are based on: (i) word overlap, which measures the amount of common words in two sentences; (ii) relative distance (or relative position), which indicates the distance between sentences in the summary and in the text; and (iii) relative size, which measures the difference, in characters, for two sentences. Some examples may be seen in Figures 6 and 7.

Sentence 1	Sentence 2
O agressor morreu, mas ainda não foi confirmado se ele foi baleado pela polícia ou se cometeu suicídio. (18 words) (The attacker died, but it is not yet confirmed whether he was shot by the police or committed suicide.)	*Ainda não se sabe se ele cometeu suicídio ou foi morto por policiais.* (13 words) (It is still unknown whether he committed suicide or was killed by police officers.)

Fig. 6. Example for word overlap

Sentence 1	Sentence 2
Na sexta-feira, choveu 12 centímetros em algumas regiões, e há previsão de mais tempestades hoje. (98 characters) (On Friday, it rained 12 cm in some areas, and there are more storms forecast today.)	*Na sexta-feira, choveu muito acima do esperado e há previsão de mais tempestades hoje.* (86 characters) (On Friday, it rained much more than expected and there are more storms forecast today.)

Fig. 7. Example for relative size

In this work, the value of word overlap ranges from 0 to 1 and is computed with the following formula (applied for the example in Figure 6):

$$\frac{words\ in\ common\ in\ the\ two\ sentences * 2}{number\ of\ words\ in\ the\ 1st\ sentence + number\ of\ words\ in\ the\ 2nd\ sentence} = \frac{9*2}{18+13} = 0.58 \quad (1)$$

Indeed, it is easy to think that two sentences that convey the same topic will have words in common. For some cases, this method works really well.

For the relative size measure, the calculations are made as follows for the sentences in Figure 7. For this measure, 0 indicates that two sentences are equal in character size. This measure is inspired by the machine translation area, in which, if two sentences have approximate sizes, it is reasonable to assume that they convey similar meaning.

$$\frac{difference\ in\ number\ of\ characters\ in\ the\ sentences}{number\ of\ characters\ in\ the\ longest\ sentence} = \frac{98-86}{98} = 0.12 \quad (2)$$

The third measure, relative position, only uses the information about the positions of the two sentences (in the summary and the text), and the size of the texts. For example, when considering a summary sentence in position "2" in a hypothetical summary with 3 sentences and a sentence in position "3" in a hypothetical text with 7 sentences, the calculations are made as follows. First, it is necessary to find a value range, to know how much a text is longer than the summary:

$$value\ range = \frac{number\ of\ sentences\ in\ the\ text}{number\ of\ sentences\ in\ the\ summary} = \frac{7}{3} = 2.33 = 2 \tag{3}$$

Then, using the value range and the sentence positions (lower position = 2 and higher position = 3), the final calculations are made as follows.

$$\frac{lower\ position - \lceil \frac{higher\ position}{value\ range} \rceil}{value\ range\ number - 1} = \frac{2 - \lceil \frac{3}{2} \rceil}{3-1} = \frac{2-2}{2} = 0 \tag{4}$$

The value range number is the number of ranges in a text. In this case, it was 3. For this measure, 0 indicates that the two sentences are in equivalent positions in their corresponding texts. This measure comes from the assumption that, when one or more texts are summarized, the order of the information in the texts will be respected.

4.2 The deep approach

This approach uses Cross-document Structure Theory (CST) [26] to guide the alignments. CST indicates the discourse relations among passages of different texts. We assume that two sentences have an alignment if they show at least one CST relation. In this work, the CST relations among sentences in the texts and the summaries are recovered by the CSTParser [22], which has an overall performance of 68.57% for news texts. For certain cases, it is quite easy to verify that this assumption is true, as may be seen in Figure 8.

Sentence in the summary	CST relation	Sentence in the source text
A outra brasileira, Joana Costa, ficou na quinta posição, com 4m20, mostrando que o nervosismo pode atrapalhar as competições em casa. (The other Brazilian, Joana Costa, was in the fifth position, with 4m20, showing that nervousness can derail competitions at home.)	Overlap	*Já a outra brasileira que participou da prova, Joana Costa, não subiu ao pódio, uma vez que não alcançou a marca da cubana.* (The other Brazilian who attended the trial, Joana Costa, didn't step to the podium, since she didn't reach the Cuban mark.)

Fig. 8. Example with CST relation

In this example, two sentences have the relation "Overlap" between them, which means that the two sentences have some information in common and, at the same time, the two sentences have some information that is unique to each one. The alignment may denote exactly this kind of information, so it is reasonable to think that the existence of CST relations may indicate alignments.

It is important to notice that this approach may possibly incorporate errors from the parser if it fails to recognize some relations.

4.3 Hybrid Approach: the Use of Machine Learning

This third approach combines the previous ones in a machine learning solution to perform the alignments. The features are related to the superficial methods and the CST-based method for a pair of sentences (one from the texts and one from the summary), namely: word overlap, relative size, relative position, number of CST relations and type of CST relations[1]; the class values were "yes", in the case of an alignment, or "no", if an alignment does not occur. We use 4 different learning techniques in WEKA environment [12], which are J48 [25], OneR [16], SVM [8] and Naïve Bayes [18]. Our database is composed by 15689 examples related to all possible alignments in the manually annotated *corpus*. 93.55% of them (14678) are examples of pairs with "no" class, and the other 6.44% (1011) are examples of pairs with "yes" class. The database is very unbalanced, but, even with this real life scenery, the results obtained were good, as we show in the next section.

5 Results

The main results obtained by the methods for detecting the alignments are synthesized in Table 4.

Table 4. Main results

Method	Precision (%)	Recall (%)	F-measure (%)
Word overlap	71.8	61.3	66.2
Relative Position	12.7	68.0	21.4
Relative Size	10.1	63.3	17.5
CST method	55.0	67.8	60.7
Jing and McKeown (1999)	35.6	80.5	49.4
Machine learning (J48)	78.7	50.7	61.7
Machine learning (OneR)	86.2	47.6	61.3

We show average precision, recall and f-measure computed over the results for the clusters, showing only the results for the alignment cases. We do not show the values for the cases for which there were no alignments because all the methods performed very well in excluding "invalid" sentence pairs (*i.e.*, pairs that should not be aligned), achieving results over 90%. It is also important to say that, for the machine learning cases (the two last lines in the table), 10-fold cross-validation was performed. Therefore, it is necessary some reservation before directly comparing the results of machine learning with the other results.

---

[1] The CST relation types are referred to the types in a typology created in [23]. They may be redundancy, complement, contradiction, source/authorship or style.

Overall, one may see that our best method was the superficial method word overlap, which reached an F-measure of 66.2%. The CST method was also good, achieving an F-measure of 60.7%. Its results were lower than the ones for word overlap and machine learning methods probably because of errors from the CSTParser.

Regarding machine learning, although we have tested some other techniques too, we only show the ones that produced the best results, the J48 and OneR techniques. In particular, OneR used the word overlap feature for composing its classification rule, showing once more the discriminative power of such information.

We also made an experiment manually balancing our database (by oversampling the minority class), and the results were very good: J48 and OneR achieved F-measures of 97.2% and 86.4%, respectively. However, we do not appreciate such solution because it introduces a large bias (since the minority class is replicated several times) and it does not correspond to the actual data that we find in the real world.

We also analyzed which features were the most important ones, using information gain for feature selection. This method ranked the features in the following order: word overlap, type of CST relations, number of CST relations, relative position and relative size, with word overlap being the best. However, running the machine learning techniques with the best features did not improve the results. It is also interesting to notice that the relative size and the relative position obtained bad results as methods, but they improved the machine learning results when used as features.

Furthermore, we reproduced the method proposed by Jing and McKeown [17], which is a very popular method. One may see that, although it achieved a good recall, its precision was low, and, overall, it was outperformed by the other methods.

Finally, Table 5 shows the percentage of alignment types that were correctly identified by our best overall method, the word overlap method.

Table 5. Accuracy over alignment types

Category	Type	Number of occurrences	Number of detected alignments	Percentage of detected alignments (%)
Form	Partial	871	530	60.85
	Different	82	1	1.22
	Identical	59	58	98.31
Content	Neutral	955	586	61.36
	Generalization	47	13	23.75
	Specification	80	19	27.66
	Contradiction	37	12	32.43
	Inference	29	11	37.93
	Other	6	0	0.00
Onomastics	Anthroponomastics	23	4	17.39
	Toponomastics	4	3	75.00

As expected, identical alignments are simple to identify. Partial and neutral alignments (which are the most common alignment types) are also correctly detected with good accuracy. “Generalization” and “specification” showed to be challenges for alignment detection, as well as the “different” cases.

6 Final Remarks

For future works, we highlight the possibility of performing the alignments among more refined textual segments, like n-grams or words. Alignments among n-grams or words may reveal more information about how humans summarize texts, but they are more likely to produce worse results. We also envision the possibility of incorporating semantic features for aiding in the alignment detection.

Acknowledgments

The authors are grateful to FAPESP and CAPES for supporting this work.

References

1. Agostini, V., Camargo, R. T., Di Felippo, A. & Pardo, T. A. S.: *Manual Alignment of News Texts and their Multi-document Human Summaries*. In S.M. Aluísio and S.E.O. Tagnin (eds.), New Language Technologies and Linguistic Research: A Two-Way Road, pp. 148-170. Cambridge Scholars Publishing (2014)
2. Banko, M., Mittal, V., Kantrowitz, M., & Goldstein, J.: Generating Extraction-Based Summaries from Hand-Written Summaries by Aligning Text Spans. In the *Proceedings of the 4th Conference of the Pacific Association for Computational Linguistics*, 5p (1999)
3. Barzilay, R., & Elhadad, N.: Sentence Alignment for Monolingual Comparable Corpora. In the *Proceedings of the Empirical Methods for Natural Language*, pp. 25-32 (2003)
4. Camargo, R. T., Agostini, V., Di Felippo, A., & Pardo, T. A. S.: Manual Typification of Source Texts and Multi-document Summaries Alignments. *Procedia – Social and Behavioral Sciences*, Vol. 95, pp. 498-506 (2013)
5. Cardoso, P. C. F., Maziero, E. G., Castro Jorge, M. L. C., Seno, E. M. R., Di Felippo, A., Rino, L. H. M., Nunes, M. G. V., & Pardo, T. A. S.: CSTNews - A Discourse-Annotated Corpus for Single and Multi-Document Summarization of News Texts in Brazilian Portuguese. In the *Proceedings of the 3rd RST Brazilian Meeting*, pp. 88-105. October 26, Cuiabá/MT, Brazil (2011)
6. Carletta, J.: Assessing Agreement on Classification Tasks: The Kappa Statistic. Computational Linguistics, v. 22, n. 2, pp. 249-254 (1996)
7. Cohen, J.: A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, v. 20, n. 1, pp. 37-46 (1960).
8. Cortes, C.; Vapnik, V. Support-vector networks. Machine learning, 20(3), pp. 273-297 (1995)
9. Daumé III, H., & Marcu, D.: A Phrase-Based HMM Approach to Document/Abstract Alignment. In the *Empirical Methods in Natural Language Processing (EMNLP)*, 8p (2004)

10. Daumé III, H., & Marcu, D.: Induction of Word and Phrase Alignments for Automatic Document Summarization. *Computational Linguistics*, v. 31, n. 4, pp. 505-530 (2005)
11. Gale, W. A., & Church, K. W.: A program for aligning sentences in bilingual corpora. *Computational Linguistics*, v. 19, n. 1, pp. 75-102 (1993)
12. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H.: *The WEKA Data Mining Software: An Update; SIGKDD Explorations*, Volume 11, Issue 1 (2009)
13. Hatzivassiloglou, V., Klavans, J. L., & Eskin, E.: Detecting Text Similarity over Short Passages: Exploring Linguistic Feature Combinations via Machine Learning**.** In the *Proceedings of the Empirical Methods for Natural Language Processing*, pp. 203-212 (1999)
14. Hatzivassiloglou, V., Klavans, J. L., Holcombe, M. L., Barzilay, R., Kan, M., & McKeown, K. R.: SIMFINDER: A Flexible Clustering Tool for Summarization**.** In the *Proceedings of the NAACL Workshop for Summarization*, pp. 41-49 (2001)
15. Hirao, T., Suzuki, J., Isozaki, H., & Maeda, E.: Dependency-based Sentence Alignment for Multiple Document Summarization. In the *COLING '04 Proceedings of the 20th international conference on Computational Linguistics*, pp. 446-452 (2004)
16. Holte, R. C.: Very simple classification rules perform well on most commonly used datasets. *Machine learning*, 11(1), pp. 63-90 (1993)
17. Jing, H., & McKeown, K.: The Decomposition of Human-Written Summary Sentences. In the Proceedings of the 22nd Annual International ACMSIGIR Conference on Research and Development in Information Retrieval, pp. 129-136 (1999)
18. John, G. H.; Langley, P. Estimating continuous distributions in Bayesian classifiers. In the *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pp. 338-345 (1995)
19. Mani, I.: *Automatic Summarization. Natural Language Processing Vol. 3.* Amsterdam/Philadelphia: John Benjamins Publishing Company. 285p (2001)
20. Mann, W. C., & Thompson, S. A.: *Rhetorical structure theory: A theory of text organization*. Tech. rep. ISI/RS-87-190, University of Southern California, 83p (1987)
21. Marcu, D.: The automatic construction of large-scale corpora for summarization research. In the *Proceedings of the 22nd Conference on Research and Development in Information Retrieval*, pp. 137-144 (1999)
22. Maziero, E. G., & Pardo, T. A. S.: Multi-Document Discourse Parsing Using Traditional and Hierarchical Machine Learning. In the *Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology*, pp. 1-10.October 24-26, Cuiabá/MT, Brazil (2011)
23. Maziero, E. G., Castro Jorge, M. L. C., & Pardo, T. A. S.: Identifying Multidocument Relations. In the *Proceedings of the 7th International Workshop on Natural Language Processing and Cognitive Science - NLPCS*, pp. 60-69. June 8-12, Funchal/Madeira, Portugal (2010)
24. Nenkova, A.; McKeown, K. Automatic summarization. *Foundations and Trends in Information Retrieval 5* (2-3), pp. 103-233 (2011)
25. Quinlan, J. R.: *C4.5: programs for machine learning (Vol. 1)*. Morgan kaufmann (1993)
26. Radev, D.R.: A common theory of information fusion from multiple text sources, step one: Cross-document structure. In the *Proceedings of the 1st ACL SIGDIAL Workshop on Discourse and Dialogue*, pp. 74-83 (2000)
27. Yamada, K., & Knight, K.: A syntax-based statistical translation model. In the *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 523-530.Toulouse, France, July (2001)