Experiments with CST-based Multidocument Summarization

Maria Lucía del Rosario Castro Jorge, Thiago Alexandre Salgueiro Pardo
Núcleo Interinstitucional de Lingüística Computacional (NILC)
Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo
Avenida Trabalhador são-carlense, 400 - Centro
P.O.Box 668. 13560-970, São Carlos/SP, Brazil
{mluciacj,taspardo}@icmc.usp.br

Abstract

Recently, with the huge amount of growing information in the web and the little available time to read and process all this information, automatic summaries have become very important resources. In this work, we evaluate deep content selection methods for multidocument summarization based on the CST model (Cross-document Structure Theory). Our methods consider summarization preferences and focus on the overall main problems of multidocument treatment: redundancy, complementarity, and contradiction among different information sources. We also evaluate the impact of the CST model over superficial summarization systems. Our results show that the use of CST model helps to improve informativeness and quality in automatic summaries.

1 Introduction

In the last years there has been a considerable increase in the amount of online information and consequently the task of processing this information has become more difficult. Just to have an idea, recent studies conducted by IDC showed that 800 exabytes of information were produced in 2009, and it is estimated that in 2012 it will be produced 3 times more. Among all of this information, there is a lot of related content that comes from different sources and that presents similarities and differences. Reading and dealing with this is not straightforward. In this scenario, multidocument summarization has become an important task.

Multidocument summarization consists in producing a unique summary from a set of documents on the same topics (Mani, 2001). A multidocument summary must contain the most relevant information from the documents. For example, we may want to produce a multidocument summary from all the documents telling about the recent world economical crisis or the terrorism in some region. As an example, Figure 1 reproduces a summary from Radev and Mckeown (1998), which contains the main facts from 4 news sources.

> Reuters reported that 18 people were killed in a Jerusalem bombing Sunday. The next day, a bomb in Tel Aviv killed at least 10 people and wounded 30 according to Israel radio. Reuters reported that at least 12 people were killed and 105 wounded. Later the same day, Reuters reported that the radical Muslim group Hamas had claimed responsibility for the act.

Figure 1: Example of multidocument summary (Radev and Mckeown, 1998, p. 478)

Multidocument summarization has to deal not only with the fact of showing relevant information but also with some multidocument phenomena such as redundancy, complementarity, contradiction, information ordering, source identification, temporal resolution, etc. It is also interesting to notice that, instead of only generic summaries (as the one in the example), summaries may be produced considering user preferences. For example, one may prefer summaries including information attributed to particular sources (if one trusts more in some sources) or more context information (considering a reader that has not accompanied some recent important news), among other possibilities.

Proceedings of the 2010 Workshop on Graph-based Methods for Natural Language Processing, ACL 2010, pages 74–82,
Uppsala, Sweden, 16 July 2010.

There are two main approaches for multidocument summarization (Mani and Maybury, 1999): the superficial and the deep approaches. Superficial approach uses little linguistic knowledge to produce summaries. This approach usually has low cost and is more robust, but it produces poor results. On the other hand, deep approaches use more linguistic knowledge to produce summaries. In general terms, in this approach it is commonly used syntactical, semantic and discourse parsers to analyze the original documents. A very common way to analyze documents consists in establishing semantic relations among the documents parts, which helps identifying commonalities and differences in information. Within this context, discourse models as CST (Cross-document Structure Theory) (Radev, 2000) are useful (see, e.g., Afantenos et al., 2004; Afantenos, 2007; Jorge and Pardo, 2009, 2010; Radev and Mckeown, 1998; Radev et al., 2001; Zhang et al., 2002).

It was proposed in Mani and Maybury (1999) a general architecture for multidocument summarization, with analysis, transformation, and synthesis stages. The first stage consists in analyzing and formally representing the content of the original documents. The second stage consists mainly in transforming the represented content into a condensed content that will be included in the final summary. One of the most important tasks in this stage is the content selection process, which consists in selecting the most relevant information. Finally, the third stage expresses the condensed content in natural language, producing the summary.

In this paper, we explore a CST-based summarization method and evaluate the corresponding prototype system for multidocument summarization. Our system, called CSTSumm (CST-based SUMMarizer), produces multidocument summaries from input CST-analyzed news documents. We mainly investigate content selection methods for producing both generic and preference-based summaries. Particularly, we formalize and codify our content selection strategies as operators that perform the previously cited transformation stage. We run our experiments with Brazilian Portuguese news texts (previously analyzed according to CST by human experts) and show that we produce more informative summaries in comparison with some superficial summarizers (Pardo, 2005; Radev et al., 2000). We also use CST to enrich these superficial summarizers, showing that the results also improve. Our general hypothesis for this work is that the deep knowledge provided by CST helps to improve information and quality in summaries.

This work is organized as follows. In Section 2, the main concepts of the CST model are introduced and the works that have already used CST for multidocument summarization are reviewed. In Section 3, we present CSTSumm, while its evaluation is reported in Section 4. Some final remarks are presented in Section 5.

2 Related Work

2.1 Cross-document Structure Theory

Radev (2000) proposed CST model with a set of 24 relations for multidocument treatment in any domain. Table 1 lists these relations.

Table 1: CST original relations

Identity	Judgment
Equivalence	Fulfillment
Translation	Description
Subsumption	Reader profile
Contradiction	Contrast
Historical background	Parallel
Modality	Cross-reference
Attribution	Citation
Summary	Refinement
Follow-up	Agreement
Elaboration	Generalization
Indirect speech	Change of perspective

The established relations may have (or not) directionality, e.g., the equivalence relation (which states that two text segments have similar content) has no directionality while the historical background relation (which states that a segment provides historical information about other) has. Figure 2 shows examples of these two relations among sentences from different sources.

As part of the model, the author proposes a general schema that reveals the possibility of relationship at any level of linguistic analysis. Figure 3 (reproduced from Radev, 2000) illustrates this schema. According to this schema, the documents with CST relations are represented as a graph, whose nodes are text segments (of possibly any level) and the edges are relations. This graph is possibly disconnected, since not all segments present relations with other segments. It is important to say that, in general, only one analysis level is treated. In this work, we only deal with sentences from the input documents, since sentences are well delimited and are standard segments in discourse analysis.

Equivalence relation

Sentence 1: Nine people died, three of them children, and 25 others were wounded last Monday in a blast at a market in Moscow, police said.

Sentence 2: Nine people died, including three children, and 25 others were injured last Monday in an explosion that happened at a market in Moscow, police of Moscow informed.

Historical background relation
(directionality: from Sentence 2 to 1)

Sentence 1: An airplane accident in Bukavu, east of Democratic Republic of Congo, killed 13 people this Thursday in the afternoon.

Sentence 2: Congo has a history of more than 30 airplane tragedies.

Figure 2: Examples of CST relations

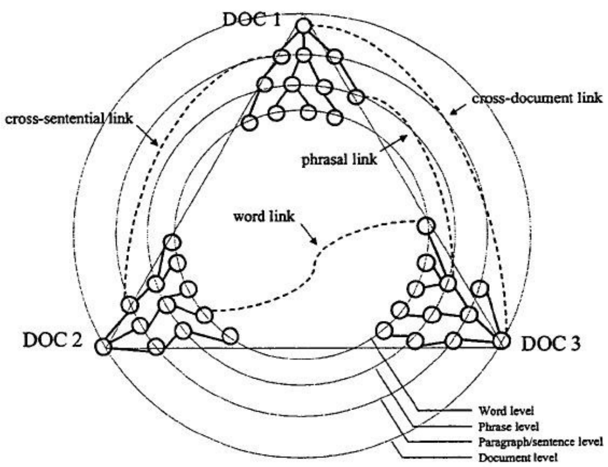

Figure 3: CST general schema (Radev, 2000, p. 78)

2.2 Multidocument Summarization

A few works explored CST for multidocument summarization. A 4-stage multidocument summarization methodology was proposed in Radev (2000). In this methodology, the first stage consists in clustering documents according to their topics. In the second stage, internal analysis (syntactical and semantic, for instance) of the texts may be performed. In the third stage, CST relations are established among texts. Finally, in the fourth stage, information is selected to produce the final summary. For this methodology the author suggests using operators activated by user summarization preferences such as authorship (i.e., reporting the information sources) or contradictory information preference. The author also says that it may be possible to produce generic summaries without considering a particular preference. In this case the criterion used to select information is based on the number of CST relations that a segment has. This criterion is based on the idea that relevant information is more repeated/elaborated and related to other segments across documents. This may be easily verified in practice. In this paper we follow such ideas.

A methodology for enriching multidocument summaries produced by superficial summarizers was proposed by Zhang et al. (2002). The authors incorporated the information given by CST relations to MEAD (Radev et al., 2000) summarization process, showing that giving preference to segments with CST relations produces better summaries. Otterbacher et al. (2002) investigated how CST relations may improve cohesion in summaries, which was tested by ordering sentences in summaries according to CST relations. The idea used behind this ordering is that sentences related by CST relations should appear closer in the final summaries as well as should respect possible temporal constraints indicated by some relations.

Afantenos et al. (2004) proposed another summarization methodology that extracts message templates from the texts (using information extraction tools) and, according to the type of CST relation between two templates, produces a unified message that would represent the summary content. The authors did not fully implement this method.

3 CSTSumm

In this paper we evaluate a CST-based multidocument summarization method by implementing and testing a prototype system, called CSTSumm. It performs content selection operations over a group of texts on the same topic that were previously annotated according to CST. For the moment, we are using manually annotated texts, i.e., the analysis stage of multidocument summarization is only simulated. In the future, texts may be automatically annotated, since a CST parser is under development for Brazilian Portuguese language (Maziero et al., 2010).

Initially, the system receives as input the CST-annotated texts, which are structured as a graph. An initial rank of sentences is then built: the sentences are ordered according to the number of CST relations they present; the more relations a sentence presents, better ranked it will be. Having the initial rank, content selection is performed. In this work, following the idea of Jorge and Pardo (2010), we represent and codify each content selection strategy as an operator. A content selection operator tells how to rearrange the sentences in the rank in order to produce summaries that better satisfy the corresponding user preferences. For instance, if a user requires more context information in the summary, the corresponding operator is activated. Such operator will (i) select in the rank all the sentences that present historical background and elaboration CST relations with better ranked sentences and (ii) improve their position in the rank by putting them immediately after the better ranked sentences with which they are related. This final action would give to these "contextual" sentences more preference for being in the summary, since they are better positioned in the refined rank. Figure 4 shows an example of a hypothetical CST graph (derived from a group of texts), the corresponding initial rank (with relations preserved for clarification) and the transformation that the context operator would do for producing the new/refined rank. It is possible to see that sentence 1, that presents historical information about the sentence 4, gets a better position in the rank (immediately after sentence 4), receiving some privilege to be in the summary.

Besides the context operator, we also have other 3 operators: the contradiction operator (which looks for the contradiction CST relation in order to include in the summary every contradiction in the texts), the authorship operator (which looks for the citation and attribution CST relations in order to include in the summary possible sources that provided the available information), and the evolving events operator (which looks for historical background and follow-up CST relations in order to present the development of the events during a time period).

Independently from the user preference, an extra operator is always applied: the redundancy operator. It removes from the rank all sentences whose information is already expressed in other better ranked sentences. Redundancy is represented by the identity, equivalence, and subsumption CST relations.

After the content selection process, in the last stage – the synthesis stage – the system selects as many sentences from the rank as allowed by the specified compression rate. The compression rate (provided by the user) informs the size of the summary. For instance, a 70% rate indicates that the summary must have at most 30% of the number of words in a text. In this work, given the multidocument nature of the task, we compute the compression rate over the size of the longest text in the group.

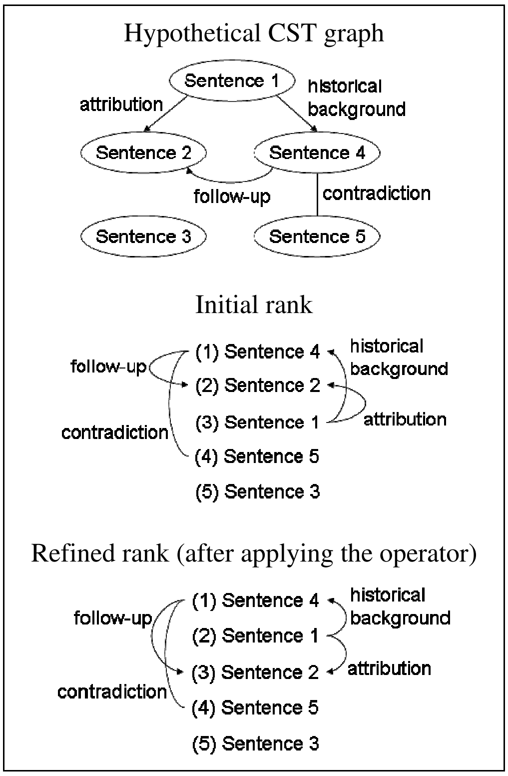

Figure 4: Example of context operator application

Synthesis stage also orders the selected sentences according to a simple criterion that only considers the position of the sentences in the original documents: first sentences appear first in the summary. If two sentences have the same position but in different documents, then the sentences are ordered according to the document number. Finally, we apply a sentence fusion system (Seno and Nunes, 2009) to some selected sentences. This is done when sentences with overlap CST relation among them are selected to the summary. The overlap relation indicates that the sentences have similar content, but also that both present unique content. In this case, it is desired that the sentences become only one with the union of their contents. The fusion system that we use does that. Figure 5 illustrates the fusion process, with the original sentences and a resulting fusion.

Figure 6 shows the general architecture of CSTSumm, which summarizes the whole process described before. Each operator is codified in

XML, where it is specified which relations should be looked in the rank in order to have the correspondent sentences better ranked. It is important to notice that, excepting the redundancy operator, our system was designed to allow the application of only one content selection operator at a time. If more than one operator is applied, the application of the following operator may probably rewrite the modifications in the rank that the previous operator has done. For instance, the application of the contradiction operator after the context operator might include sentences with contradiction above sentences with context information in the rank, altering therefore the rank produced by the context operator. One simple alternative to this design choice is to ask the user to rank his preferences and, then, to apply the corresponding operators in the opposite order, so that the rank produced by the most important preference will not be further altered. Other alternative is to produce more complex operators that combine preferences (and the corresponding CST relations), but some preference on the relations should still be specified.

Sentence 1: According to a spokesman from United Nations, the plane was trying to land at the airport in Bukavu in the middle of a storm.

Sentence 2: Everyone died when the plane, hampered by bad weather, failed to reach the runway and crashed in a forest 15 kilometers from the airport in Bukavu.

Fusion: According to a spokesman for the United Nations, everyone died when a plane that was trying to land at Bukavu airport, hampered by bad weather, failed to reach the runway and crashed in a forest 15 kilometers from the airport.

Figure 5: Example of sentence fusion

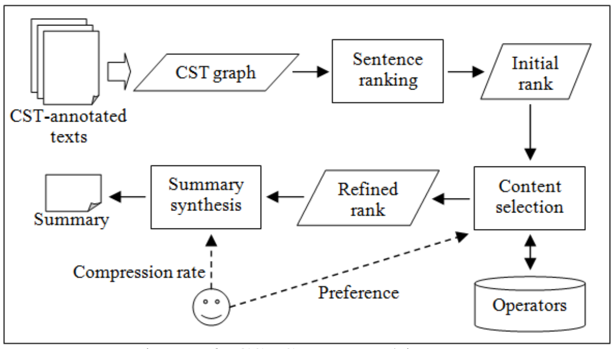

Figure 6: CSTSumm architecture

In Figure 7 we show the algorithm for the application of operators during content selection process. It is important to notice that the selected operator looks for its relations in all pairs of sentences in the rank. Once it finds the relations, it rearranges the rank appropriately, by putting the related sentence more above in the rank.

<u>procedure for application of content selection operators</u>
input data: initial rank, user summarization preference, operators
output data: refined rank
apply the redundancy operator
select one operator according to the user summarization preference
for i=sentence at the first position in the rank to the last but one sentence
    for j=sentence at position i+1 in the rank to the last sentence
        if the operator relations happen among sentences i and j, rearrange the rank appropriately

Figure 7: Algorithm for application of content selection operators

As an illustration of the results of our system, Figure 8 shows an automatic summary produced from a group of 3 texts with the application of the context operator (after redundancy operator was applied) and a 70% compression rate. The summary was translated from Portuguese, the language with which the summarizer was tested.

The Brazilian volleyball team has won on Friday the seventh consecutive victory in the World League, defeating Finland by 3 sets to 0 - partials of 25/17, 25/22 and 25/21 - in a match in the Tampere city, Finland. The first set remained balanced until the middle, when André Heller went to serve. In the last part, Finland again paired the game with Brazil, but after a sequence of Brazilians points Finland failed to respond and lost by 25 to 21. The Brazilian team has won five times the World League in 1993, 2001, 2003, 2004 and 2005.

Figure 8: Example of multidocument summary with context information

4 Evaluation

Our main research question in this work was how helpful CST would be for producing better summaries. CSTSumm enables us to assess the summaries and content selection strategies, but a comparison of these summaries with summaries produced by superficial methods is still necessary. In fact, we not only proceeded to such comparison, but also improved the superficial methods with CST knowledge.

As superficial summarizers, we selected MEAD (Radev et al., 2000) and GistSumm (Pardo et al., 2003; Pardo, 2005) summarizers. MEAD works as follows. Initially, MEAD builds an initial rank of sentences according to a score based on three parameters: position of the sentence in the text, lexical distance of the sentence to the centroid of the text, and the size of the sentence. These three elements are linearly combined for producing the score. GistSumm, on the other side, is very simple: the system juxtaposes all the source texts and gives a score to each sentence according to the presence of frequent words (following the approach of Luhn, 1958) or by using TF-ISF (Term Frequency – Inverse Sentence Frequency, as proposed in Larroca et al., 2000). Following the work of Zhang et al. (2002), we decided to use CST to rearrange (and supposedly improve) the sentence ranks produced by MEAD and GistSumm. We simply add to each sentence score the number of CST relations that the sentence presents:

new sentence score = old sentence score + number of CST relations

The number of sentences is retrieved from the CST graph. This way, the sentence positions in the rank are changed.

For our experiments, we used the CSTNews corpus (Aleixo and Pardo, 2008), which is a corpus of news texts written in Brazilian Portuguese. The corpus contains 50 clusters of texts. Each group has from 2 to 4 texts on the same topic annotated according to CST by human experts, as well as a manual generic summary with 70% compression rate (in relation to the longest text). The annotation process was carried out by 4 humans, with satisfactory agreement, which demonstrated that the annotation task was well defined and performed. More details about the corpus and its annotation process are presented by Maziero et al. (2010).

For each cluster of CSTNews corpus, it was produced a set of automatic summaries corresponding to each method that was explored in this work. To evaluate the informativity and quality of the summaries, we used two types of evaluation: automatic evaluation and human evaluation. For the automatic evaluation we used ROUGE (Lin, 2004) informativity measure, which compares automatic summaries with human summaries in terms of the n-grams that they have in common, resulting in precision, recall and f-measure numbers between 0 (the worst) and 1 (the best), which indicate how much information the summary presents. Precision indicates the amount of relevant information that the automatic summary contains; recall indicates how much information from the human summary is reproduced in the automatic summary; f-measure is a unique performance measure that combines precision and recall. Although it looks simple, ROUGE author has showed that it performs as well as humans in differentiating summary informativeness, which caused the measure to be widely used in the area. In particular, for this work, we considered only unigram comparison, since the author of the measure demonstrated that unigrams are enough for differentiating summary quality. For computing ROUGE, we compared each automatic summary with the corresponding human summary in the corpus.

We computed ROUGE for every summary we produced through several strategies: using only the initial rank, only the redundancy operator, and the remaining preference operators (applied after the redundancy operator). Is is important to notice that it is only fair to use ROUGE to evaluate the summaries produced by the initial rank and by the redundancy operator, since the human summary (to which ROUGE compares the automatic summaries) are generic, produced with no preference in mind. We only computed ROUGE for the preference-biased summaries in order to have a measure of how informative they are. Ideally, these preference-biased summaries should not only mirror the user preference, but also contain the main information from the source texts.

On the other hand, we used human evaluation to measure the quality of the summaries in terms of coherence, cohesion and redundancy, factors that ROUGE is not sensitive enough to capture. By coherence, we mean the characteristic of a text having a meaning and being understandable. By cohesion, we mean the superficial makers of coherence, i.e., the sequence of text elements that connect the ideas in the text, as punctuation, discourse markers, anaphors, etc.

For each one of the above evaluation factors, a human evaluator was asked to assign one of five values: very bad (score 0), bad (score 1), regular (score 2), good (score 3), and excellent (score 4). We also asked humans to evaluate informativity in the preference-biased summaries produced by our system, which is a more fair evaluation than the automatic one described above. The user should score each summary (using the same values above) according to how much he was satisfied with the actual content of the summary in face of the preference made. The user had access to the source texts for performing the evaluation.

Table 2 shows the ROUGE scores for the summaries produced by the initial rank, by the application of the operators, by the superficial summarizers, and by the CST-enriched superficial summarizers. It is important to say that these results are the average results obtained for the automatic summaries generated for all the clusters in the CSTNews corpus.

Table 2: ROUGE results

Content selection method	Precision	Recall	F-measure
Initial rank	0.5564	0.5303	0.5356
Redundancy treatment (only)	0.5761	0.5065	0.5297
Context information	0.5196	0.4938	0.4994
Authorship information	0.5563	0.5224	0.5310
Contradiction information	0.5503	0.5379	0.5355
Evolving events information	0.5159	0.5222	0.5140
MEAD without CST	0.5242	0.4602	0.4869
MEAD with CST	0.5599	0.4988	0.5230
GistSumm without CST	0.3599	0.6643	0.4599
GistSumm with CST	0.4945	0.5089	0.4994

As expected, it may be observed that the best results were achieved by the initial rank (since it produces generic summaries, as happens to the human summaries to which they are compared), which does not consider any summarization preference at all. It is also possible to see that: (a) the superficial summarizers are outperformed by the CST-based methods and (b) CST-enriched superficial summarizers produced better results than the superficial summarizers.

Results for human evaluation are shown in Table 3. These results show the average value for each factor evaluated for a sample group of 48 texts randomly selected from the corpus. We also associated to each value the closest concept in our evaluation. We could not perform the evaluation for the whole corpus due to the high cost and time-demanding nature of the human evaluation. Six humans carried out this evaluation. Each human evaluated eight summaries, and each summary was evaluated by three humans.

Table 3: Results for human evaluation

Content selection method	Coherence	Cohesion	Redundancy	Informativity
Initial rank	3.6 Excellent	3.2 Good	1.8 Regular	3.6 Excellent
Context	2.1 Regular	2.7 Good	3.6 Excellent	2.2 Regular
Authorship	3.3 Good	2.4 Regular	2.8 Good	3 Good
Contradiction	2.4 Regular	2.7 Good	2.5 Regular	3.7 Excellent
Evolving events	2.1 Regular	2.5 Regular	2.6 Good	3.2 Good

It may be observed that informativity factor results are quite satisfactory, since more than 50% of the judges considered that the performance was excellent. For coherence, cohesion and redundancy factors, results were not excellent in all the cases, but they were not bad either. We consider that one of the things that could have had an influence in this case is the performance of the fusion system, since it may generate sentences with some problems of coherence and cohesion. There are also other things that may influence these results, such as the method for ordering sentences that we used in this work. This method does not follow any deep criteria to order sentences and may also lead to coherence and cohesion problems.

These results show that CSTSumm is capable of producing summaries with good informativity and quality. In fact, the results validate our hypothesis that deep knowledge may improve the results, since it deals better with the multidocument phenomena, as the presence of redundant, complementary and contradictory information.

5 Final Remarks

Although we consider that very good results were achieved, there is still room for improvements. Future works include the investigation of better sentence ordering methods, as well as more investigation on how to jointly apply more than one content selection operator.

For the moment, CSTSumm assumes that the texts to be summarized must be already annotated with CST. In the future, as soon as an automatic CST parser is available for Portuguese, it should provide the suitable input to the summarizer.

Finally, it is interesting to notice that, although we have tested our methods with Brazilian Portuguese texts, they are robust and generic enough to be applied to any other language, since both our methods and CST model are language independent.

Acknowledgments

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

References

Afantenos, S.D.; Doura, I.; Kapellou, E.; Karkaletsis, V. 2004. Exploiting Cross-Document Relations for Multi-document Evolving Summarization. In the *Proceedings of SETN*, pp. 410-419.

Afantenos, S.D. 2007. Reflections on the Task of Content Determination in the Context of Multi-Document Summarization of Evolving Events. In *Recent Advances on Natural Language Processing*. Borovets, Bulgaria.

Aleixo, P. and Pardo, T.A.S. 2008. *CSTNews: Um Córpus de Textos Journalísticos Anotados segundo a Teoria Discursiva CST (Cross-Document Structure Theory)*. Série de Relatórios Técnicos do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo no. 326. São Carlos, Brazil .

Jorge, M.L.C and Pardo, T.A.S. 2009. Content Selection Operators for Multidocument Summarization based on Cross-document Structure Theory. In the *Proceedings of the 7th Brazilian Symposium in Information and Human Language Technology*. São Carlos, Brazil.

Jorge, M.L.C. and Pardo, T.A.S. 2010. Formalizing CST-based Content Selection Operations. In the *Proceedings of the 9th International Conference on Computational Processing of Portuguese Language* (Lecture Notes in Artificial Intelligence 6001), pp. 25-29. Porto Alegre, Brazil.

Larocca Neto, J.; Santos, A.D.; Kaestner, A.A.; Freitas, A.A. 2000. Generating Text Summaries through the Relative Importance of Topics. In M.C. Monard and J.S. Sichman (eds.), *Lecture Notes in Artificial Intelligence*, N. 1952, pp. 300-309. Springer, Verlag.

Lin, C-Y. 2004. ROUGE: a Package for Automatic Evaluation of Summaries. In the *Proceedings of the Workshop on Text Summarization Branches Out*. Barcelona, Spain.

Luhn, H.P. 1958. The automatic creation of literature abstracts. *IBM Journal of Research and Development*, Vol. 2, pp. 159-165. Barcelona, Spain.

Mani, I. and Maybury, M. T. 1999. *Advances in automatic text summarization*. MIT Press, Cambridge, MA.

Mani, I. 2001. *Automatic Summarization*. John Benjamins Publishing Co. Amsterdam.

Maziero, E.G.; Jorge, M.L.C.; Pardo, T.A.S. 2010. Identifying Multidocument Relations. In the *Proceedings of the 7th International Workshop on Natural Language Processing and Cognitive Science*. June 8-12, Funchal/Madeira, Portugal.

Otterbacher, J.C.; Radev, D.R.; Luo, A. 2002. Revisions that improve cohesion in multi-document summaries: a preliminary study. In the *Proceedings of the Workshop on Automatic Summarization*, pp 27-36. Philadelphia.

Pardo, T.A.S.; Rino, L.H.M.; Nunes, M.G.V. 2003. GistSumm: A Summarization Tool Based on a New Extractive Method. In N.J. Mamede, J. Baptista, I. Trancoso, M.G.V. Nunes (eds.), *6th Workshop on Computational Processing of the Portuguese Language - Written and Spoken* (Lecture Notes in Artificial Intelligence 2721), pp. 210-218. Faro, Portugal.

Pardo, T.A.S. 2005. *GistSumm - GIST SUMMarizer: Extensões e Novas Funcionalidades*. Série de Relatórios do NILC. NILC-TR-05-05. São Carlos, Brazil.

Radev, D. and McKeown, K. 1998. Generating natural language summaries from multiple on-line sources. *Computational Linguistics*, Vol. 24, N. 3, pp. 469-500.

Radev, D.R. 2000. 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*. Hong Kong.

Radev, D.R.; Jing, H.; Budzikowska, M. 2000. Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation and user studies. In the *Proceedings of the ANLP/NAACL Workshop*, pp. 21-29.

Radev, D.R.; Blair-Goldensohn, S.; Zhang, Z. 2001. Experiments in single and multi-document summarization using MEAD. In the *Proceedings of the 1st Document Understanding Conference*. New Orleans, LA.

Seno, E.R.M. and Nunes, M.G.V. 2009. Reconhecimento de Informações Comuns para a Fusão de Sentenças Comparáveis do Português. *Linguamática*, Vol. 1, pp. 71-87.

Zhang, Z.; Goldenshon, S.B.; Radev, D.R. 2002. Towards CST-Enhanced Sumarization. In the *Proceedings of the 18th National Conference on Artificial Intelligence*. Edmonton.