

Automatic Summarization Using Terminological and Semantic Resources

Jorge Vivaldi¹, Iria da Cunha^{1,2}, Juan-Manuel Torres-Moreno^{2,3}, Patricia Velázquez-Morales⁴

⁽¹⁾ Institut Universitari de Lingüística Aplicada (UPF)

Roc Boronat 138, 08018, Barcelona (Spain)

⁽²⁾ Laboratoire Informatique d'Avignon (UAPV)

339, chemin des Meinajariés, BP 1228, 84911, Avignon (France)

⁽³⁾ École Polytechnique de Montréal

CP 6079 Succ. Centre-ville, Montréal (Qubec) Canada

⁽⁴⁾ VM Labs, 116, Av de Tarascon, 8400, Avignon (France)

{jorge.vivaldi,iria.dacunha}@upf.edu, juan-manuel.torres@univ-avignon.fr, patricia_velazquez@yahoo.com

Abstract

In this paper we present a new algorithm for automatic summarization of specialized texts combining terminological and semantic resources: a term extractor and an ontology. The term extractor provides the list of the terms that are present in the text together their corresponding termhood. The ontology is used to calculate the semantic similarity among the terms found in the main body and those present in the document title. The phrases with the highest score are chosen to take part of the final summary. We evaluate the algorithm with ROUGE, comparing the resulting summaries with the summaries of other summarizers. The sentence selection algorithm was also tested as part of a standalone summarizer. It obtains good results, but there is a space for improvement.

1. Introduction

Nowadays automatic summarization is a very prominent research topic. At the beginning, in the 60s, the research on this field was centred basically on general discourse, although there were some exceptions: the first experiments with technical texts of Luhn (1959) and the summarizer of chemical texts of Pollock and Zamora (1975). In the 90s, several researchers started to work on specialized discourse (as for example Paice (1990), Riloff (1993), Lehman (1995), McKeown and Radev (1995), Abracos and Lopes (1997) and Saggion and Lapalme (2000) among others). But as a rule, they used the same strategies that those employed for general discourse: textual or discourse structure, cue phrases, sentence position, name entities, statistical techniques, machine learning, etc. Afantenos et al. (2005) point out that, specifically, automatic summarization of medical texts has turned to an important research subject, because professionals of the medical domain need to process a great quantity of documents. There is some related relevant work in this field, as for example Damianos et al. (2002), Johnson et al. (2002), Gaizauskas et al. (2001), Lenci et al. (2002) and Kan et al. (2001). But again the techniques that they use are not specific to the medical domain.

In da Cunha (2008), a summarization model of Spanish medical articles is presented. This model takes into account various specialized information of the medical domain: relevant lexical units of this domain and genre, textual information and discourse structure. It obtains good results but the complete implementation was not possible because at present there are not available discourse parsers for Spanish¹.

¹There is a current project in the Laboratoire Informatique d'Avignon (LIA) to develop this parser for the Spanish language, see da Cunha and Torres-Moreno (2010).

A summarization system that takes specific resources into consideration is showed in Reeve and Han (2007). This system uses lexical chains, that is, sequences of words with lexical-semantic relations (basically identity and synonymy) among them. Lexical chains were used before by Barzilay and Elhadad (1997) and Silber and McCoy (2000). The novelty of the approach of Reeve and Han (2007) is that concepts are chained according to their UMLS semantic type. UMLS (Unified Medical Language System) is an thesaurus of the biomedical domain that provides a knowledge representation of this domain. This representation classifies the concepts by semantic types and relations (hierarchical and no hierarchical) among these types². In Reeve and Han (2007), to identify the concepts, the documents are processed using the UMLS Metamap Transfer tool that identify nominal phrases which are automatically mapped into UMLS concepts and semantic types.

Inspired by the work of Luhn (1959) (terms that appear in the title of a scientific text are relevant marks that indicate its main topic), and by the works of Barzilay and Elhadad (1997), Silber and McCoy (2000) and Reeve and Han (2007), we have designed a new summarization strategy for specialized texts. Our hypothesis is that the terms that are semantically related to the terms included in the title of a specialized text are particularly relevant. Therefore, an automatic summarization algorithm that selects the text sentences that include not only the existent terms in the title, but also the sentences including terms that are semantically related with them, should get positive results.

To our knowledge, summarization systems based on the combination of terminological and semantic resources don't exist, so our work contributes to developing a new line of research in the field of automatic summarization. In order to support this idea, on the one hand, we have

²<http://www.nlm.nih.gov/>

designed a new summarization algorithm based on this principle and on the other hand, we have used this relatedness/terminological information as an additional metric of an existent summarizer. Then, we have applied both procedures to a medical corpus and we have evaluated the results using ROUGE Lin (2004).

A previous experiment combining the results of several heterogeneous summarizer in a single result was presented in (da Cunha et al., 2009). In such case, the resulting system combines linguistics and statistical techniques to create a hybrid system that outclasses any single system.

In Section 2. we describe the methodology of our task while in Section 3. we present some external resources required to implement our methodology. In Section 4. we describe the summarization algorithm while in Section 5. we present the corresponding experiments and evaluation results. Finally in Section 6. we issue some conclusions and we describe the future tasks.

2. Methodology

As a first step we have compiled a corpus that includes 50 Spanish medical papers chosen from the Spanish journal *Medicina Clínica*. We consider that this number of texts is enough for a preliminary test. In the future and according to the obtained results, we plan to do more extensive experiments. Then, we have designed the summarization algorithm and we have selected the tools to implement it. Finally, results have been evaluated using ROUGE.

Also, we perform some experimentation partially integrating the proposed methodology in CORTEX (Torres-Moreno et al., 2001), a standalone summarizer that integrates several metrics in the sentence selection algorithm. Again, results has been evaluated using ROUGE.

We have worked with medical papers because this kind of texts is published in journals including their respective abstracts written by their authors, and we have employed them for the evaluation. This fact allows us to compare such authors' abstracts with the summaries produced by our algorithm in order to carry out the final evaluation with the automatic system ROUGE. Human evaluation would be fine, but it is difficult to find people able evaluate summaries, mainly because evaluators should be specialists in the domain (they have to understand the text). This fact makes human evaluation difficult and expensive.

The notion of term that we have adopted in this work is based on the "Communicative Theory of Terminology" (Cabr e, 1999): a term is a lexical unit (single/multiple word) that designates a concept in a given specialized domain. This means that terms have a designative function and, obviously not all the words in a language have such function. To detect which are the words that in a given text have this designative function is the main task of term detector/extractors (see Cabr e et al. (2001) for a review of this subject).

It is also important to note that a term may be ambiguous. This means that sometimes the same lexical unit may designate two or more concepts in the same domain (as "adenoid" or "ganglion" in Medicine) or a different domain (for example, "mouse" in Zoology, Mathematics and in Computer Science). Moreover, a concept may be designated by two or more terms (like, "fever" and "pyrexia" in Medicine or "storage battery" and "accumulator" in Electricity).

3. External Tools and Resources

As mentioned above a number of external tools has been used to implement the SUMMTERM algorithm: YATE, EuroWordNet and CORTEX. In this section we briefly present such tools.

YATE

YATE (see Vivaldi (2001) for details) is a term extraction (TE) tool whose main characteristics are: a) it uses a combination of several term extraction techniques and b) it makes intensive use of semantic knowledge (by means of EWN, see below). Taking into consideration that all the term extraction techniques used were heterogeneous (because they are based on different properties of the term candidates, see below) it applies a combination technique to issue its results. Some of the aspects regarding this technique have been described in Vivaldi and Rodr guez (2001a) and Vivaldi and Rodr guez (2001b) where different ways of combination are presented. This tool was designed to obtain all the terms (from the following set of syntactically filtered term candidates -TC-: <noun>, <noun-adjective> and <noun-preposition-noun>) found in Spanish specialised texts within the medical domain. As mentioned above, YATE is an hybrid TE system that combines the results obtained by a set of TC analysers, described briefly as follows:

- Domain coefficient: it uses the EWN ontology to sort the TC .
- Context: it evaluates each candidate using other candidates present in its sentence context.
- Classic forms: it tries to decompose the lexical units in their formants, taking into account the form characteristics of many terms in the domain.
- Collocational method: it evaluates multiword candidates according their mutual information.
- Other internal/external TC analyzers.

The results obtained by this set of heterogeneous methods are combined using two different strategies: voting and boosting. The voting strategy simply consists in counting the results provided by each TE, using some heuristics and assigning the predominant result as the winner. Boosting is a well known member of the class of "ensemble" classifiers. These classifiers are based on the performance of a set of simple classifiers (weak learners) whose results are combined. In this class of algorithms, the basic learning step is iterated, changing at each iteration the weight of the weak learners in order to optimise an objective function. In our experiments we have used Adaboost, the most common

implementation of Boosting, with a set of instances of our TEs as weak learners.

As mentioned above, YATE has been developed specifically to use EWN as a source of domain information although each domain requires some tuning. It has been developed basically for Medicine but it has been also successfully adapted to other domains like Genomics and Economics, see (Joan et al., 2008).

It produces a list of "term candidates" sorted according their "termhood"³. The user should select, according his/her specific needs, the portion of the list that he/she accepts as valid terms in the domain, while the rest should be revised or rejected.

EuroWordNet (EWN)

EWN⁴ is a multilingual extension of WordNet, a lexical-semantic ontology developed at Princeton University. The basic semantic unit is the synset (synonymy set), grouping together several words that can be considered synonyms in some contexts. Synsets are linked by means of semantic labels (hyperonym, hyponym, meronym, etc.). Due to polysemy, lexical entries can be attached to several synsets. Coverage for medical domain in EWN is enough to develop an adequate Spanish term extractor and to use the relations among terms for our summarization algorithm. The EWN version that we use includes approximately 6,000 synsets (corresponding to 10,000 variants). This is not too bad for Spanish. Nevertheless, other specialized databases, as SNOMED-CT⁵, include 100 times more.

CORTEX

CORTEX⁶ is a single-document summarization system. The main idea is to represent the text in an appropriate way and then to apply numerical processing. CORTEX uses an optimal decision algorithm, combining several metrics, to extract the most representative sentences. The metrics are calculated by using several statistical and numerical algorithms, based on the Vector Space Model. In order to reduce the complexity, a classic pre-processing (Manning and Schütze, 1999) is performed to the document: splitting sentences, stop-word filtering, lemmatisation and/or stemming. A frequency matrix $\gamma_{[p \times N]}$ is constructed in the following way: every element γ_i^μ of this matrix represents the

number of occurrences of the word i in the sentence μ .

$$\gamma = \begin{bmatrix} \gamma_1^1 & \dots & \gamma_i^1 & \dots & \gamma_N^1 \\ \gamma_1^2 & \dots & \gamma_i^2 & \dots & \gamma_N^2 \\ \vdots & \dots & \vdots & \ddots & \vdots \\ \gamma_1^\mu & \dots & \gamma_i^\mu & \dots & \gamma_N^\mu \\ \vdots & \dots & \vdots & \ddots & \vdots \\ \gamma_1^p & \dots & \gamma_i^p & \dots & \gamma_N^p \end{bmatrix}, \quad \gamma_i^\mu \in \{0, 1, 2, \dots\} \quad (1)$$

The matrix ξ is defined as:

$$\xi_i^\mu = \begin{cases} 1 & \text{if } \gamma_i^\mu \neq 0 \\ 0 & \text{elsewhere} \end{cases} \quad (2)$$

Sentences are indexed by μ varying from 1 to p . Every column represents a type word. Words are indexed by a value i varying from 1 to N . CORTEX can use up to $\Gamma = 11$ metrics (Torres-Moreno et al., 2002b; Torres-Moreno et al., 2009) to evaluate the sentence's relevance. For example:

- The angle between the title (or the main topic) and the sentences.
- Hamming Metrics. These two other metrics use the Hamming matrix H , a square matrix $N_L \times N_L$, in which every value $H[i, j]$ represents the number of sentences in which exactly one of the terms i or j is present.
 - The sum of Hamming weights of words per segment \times the number of different words in a sentence.
 - The sum of Hamming weights of the words \times word frequencies.
- The entropy of sentences.
- The interaction (shallow words) among sentences.
- Etc.

The system scores each sentence using a decision algorithm. It computes a decision over the set of normalized metrics. Therefore two averages are calculated, a positive $\lambda_s > 0.5$ and a negative $\lambda_s < 0.5$ tendency (the case $\lambda_s = 0.5$ is ignored). The algorithm that allows to compute the vote of each metric is the following one:

$$\sum_s \alpha = \sum_{v=1}^{\Gamma} (|\lambda_s^v| - 0.5); \quad |\lambda_s^v| > 0.5 \quad (3)$$

$$\sum_s \beta = \sum_{v=1}^{\Gamma} (0.5 - |\lambda_s^v|); \quad |\lambda_s^v| < 0.5$$

Γ represents the number of metrics and v is the index of the metrics. The score of each sentence s is then computed in the following way:

$$\text{if } \left(\sum_s \alpha > \sum_s \beta \right) \quad (4)$$

$$\text{then Score}_{\text{cortex}}(s) = 0.5 + \sum_s \alpha / \Gamma \quad : \text{ retains } s$$

$$\text{else Score}_{\text{cortex}}(s) = 0.5 - \sum_s \beta / \Gamma \quad : \text{ not retains } s$$

³Termhood has been defined in Kageura and Umino (1996) as "the degree that a linguistic unit is related to domain-specific concepts".

⁴<http://www.illc.uva.nl/EuroWordNet>

⁵SNOMED-CT (Systematized Nomenclature of MEDicine, Clinical Terms) is a structured collection of medical terms that has a wide coverage of the clinical domain. The terms are organized in a number of hierarchies where nodes are linked using both vertical and horizontal specialized relations. Since 2002, a Spanish version of this resource is published regularly. (<http://www.ihtsdo.org/>)

⁶CORTEX es otro Resumidor de TEXTos.

4. Algorithm Design

The general idea of this summarization algorithm is to obtain a relevance score for each sentence taking into account both the "termhood" of the terms found in such sentence and the similarity among such terms and those terms present in the title of the document.

Figure 1 shows the general architecture of SUMMTERM, the proposed system. Firstly, a classic linguistic processing is performed over the text (common to most NLP applications: sentence segmentation, tokenization and pos-tagging). Then, the resulting file is processed by the term extractor YATE (Vivaldi and Rodríguez, 2001b; Vivaldi and Rodríguez, 2001a) in order to obtain a sorted list of all the term candidates that are present in the text. Finally the summary is produced, taking into account the term candidates, their similarity/termhood values and the input text. Some configuration information is necessary to define some useful parameters: thresholds, number of sentences in the final abstract, etc.

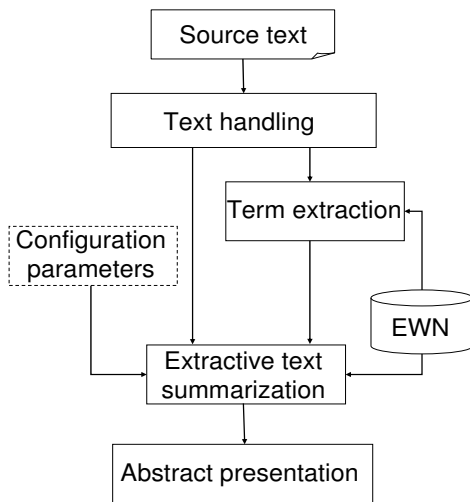


Figure 1: SUMMTERM architecture.

As mentioned above, the term extractor YATE is at first used to find all the terms in a given medical paper. Each term will have an associated termhood. Internally, the extractor keeps track of the terms present in the title from those present in other sections of the document. Then, for each term in the main body a module of YATE measures the semantic distance among such term and all the terms found in the title. As a result, the term receives a score that is calculated using (5):

$$Score(t) = T(t)P + MaxSim(t : t_t)(1 - P) \quad (5)$$

where t is a term unit and $T(t)$ its YATE termhood, t_t is any term belonging to the document title and P is a weighting coefficient (0.5 by default). The idea is that the score associated with each term will take into account its

termhood and its relatedness to the title of the document. The weighting coefficient allows us to give more or less relevance to each factor. The default value assigns equal relevance to both the termhood and the relatedness.

To calculate the similarity among the terms found in the title and those present in the main body of the document, we use information obtained from the hyperonymical paths for each synset (sy) in EuroWordNet. For such purposes we use formula (6):

$$Sim(sy_1, sy_2) = \frac{2 \times \#Common\ Nodes(sy_1, sy_2)}{Depth(sy_1) + Depth(sy_2)} \quad (6)$$

This similarity measure is based on the same measure used in TRUCKS (Maynard, 1999). Its calculation is simple, since it is done by edge counting. It takes into account two basic ideas: a) the shorter the distance among two nodes is, the higher their similarity is and b) the higher the number of common nodes is (therefore lower in the hierarchy), the higher their similarity is.

In practice, the similarity among two medical terms like "vas" and "gland" is calculated as Figure 2 shows.

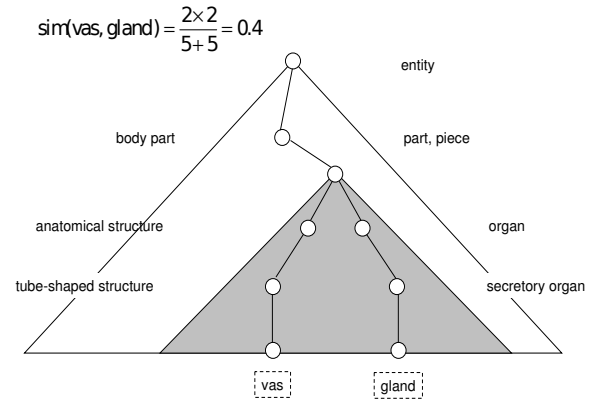


Figure 2: Example of similarity calculation.

In the case of complex terms, all the components (nouns and adjectives) are analyzed, but only the component that offers maximum similarity score is chosen. In the case of adjectives, only relational adjectives are used and the relatedness is calculated using the corresponding noun (*bronchial* → *bronchus*). To obtain the final score FS_i of each sentence i ($i = 1, \dots, k$; k =number of sentences), we add the score of all the terms that the sentence includes, using formula (7):

$$FS(s) = \sum_{t \in L_{CAT}(s)} Score(t) \quad (7)$$

where s is a sentence of the main body, $L_{CAT}(S)$ is the list of terms present in s and t is the term in s .

For example, let's imagine a medical paper with the following title: "Inappropriate visits to the emergency

services of a general hospital”. One of the sentences of that text is: “This is a descriptive study of a random sample of 84329 patients seen in 1999”. This sentence does not include any term from the title, but it contains the term “patient”, that is related semantically to “general hospital” (a term in the title) in EWN. Specifically, the similarity between these two terms is 0.15. Besides, YATE assigns a termhood of 1 to the term “patient”. In this case, considering $P = 0.5$, the score of this sentence would be $FS(s) = 0.15 \times 0.5 + 1 \times 0.5 = 0.575$. This similarity makes sense because patients are hospital users and therefore a relation between both passages exists.

To get the final summary we obtain the score of each sentence of the text and we rank them according to their score. After selecting a number of sentences to be included in the summary, we choose the sentences having the highest score and we put them back to their original order.

Combining CORTEX and SUMMTERM systems

The combination of the CORTEX and the SUMMTERM systems was intuitive and straightforward. In fact, in the Decision Algorithm (see equation 5), new metrics can be plugged-in without modification of strategy. In particular, the SummTerm output’s $FS(s)$ (see equation 7) will be considered as a new metric of Cortex system. For this purpose, $FS(s)$ was normalized to the range of values $[0,1]$. Then, for each sentence s , $\Gamma = 11 + 1 = 12$ metrics were used as inputs to the Decision Algorithm, as showed in figure 3.

In this way, SUMMTERM’s information (the termhood and similarity of terms in direct relationship with the domain) will impact on the Decision Algorithm’s results.

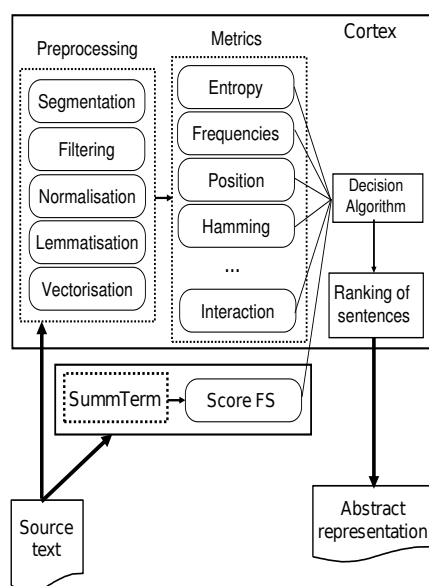


Figure 3: CORTEX + SUMMTERM architecture.

5. Experiments and Results

In order to evaluate SUMMTERM, we applied it over a corpus including 50 medical texts, obtaining the corresponding summaries. Each summary includes 11 sentences. To set the length of the summaries, we compute the average number of sentences in each section present in the author’s summaries (*Introduction, Patients and methods, Results and Discussion* sections). We decided to include in our summaries one additional sentence per section. This decision was made because we have noticed that usually authors give the same information divided into two sentences in the main body but joined in a single sentence in the abstract. In short, it was an empirical decision in order to not lose information.

As we have mentioned before, to carry out the evaluation of SUMMTERM we use the automatic system ROUGE. Nowadays, this is the most used system to evaluate summaries and it is used in the DUC/TAC competitions. It offers a set of statistics that compare peer summaries (the candidate summaries) with models summaries (summaries performed by humans). It counts cooccurrences of ngrams in peer and models to derive a score. Various statistics exist depending on the used n-grams (for example, ROUGE-2 uses bi-grams and ROUGE-SU4 uses skip bi-grams) and on the type of text processing applied to the input texts (e.g., lemmatization, stopwords removal).

To carry out the evaluation with ROUGE (specifically ROUGE-2 and ROUGE-SU4) we employed summaries (with 11 sentences) produced by other summarization systems: CORTEX (Torres-Moreno et al., 2002a), ENERTEX (Fernández et al., 2007; Fernández et al., 2008), Open Text Summarizer (OTS)⁷, Pertinence⁸, Swesum⁹ and Microsoft Word. Also we created 2 baselines in order to include them in the performance comparison. Baseline 1 (Random) contains 11 randomly selected sentences. Baseline 2 (LeadBase) contains the first 11 sentences of the text. Finally, we compared the summaries of all these systems with the abstracts of the authors of the medical articles.

As we can observe in Table 1 and in Figure 4, SUMMTERM is the fourth best system (0.33307 with ROUGE-2 and 0.38450 with ROUGE-SU4). In standalone mode, the best system is CORTEX, followed by ENERTEX, OTS and SUMMTERM. The scores of these four systems are quite similar, while the scores obtained by the remaining systems are much lower.

Remember that CORTEX is a standalone summarizer that combines several metrics. As an additional experiment, we use the SUMMTERM’s sentences scoring as an additional metric of CORTEX summarizer. Then, the final score for a document is calculated by applying the decision algorithm of CORTEX. The resulting system CORTEX+SUMMTERM

⁷<http://libots.sourceforge.net/>

⁸<http://www.pertinence.net/index.html>

⁹<http://swesum.nada.kth.se/index-eng.html>

is then the best one, in terms of ROUGE-SU4 score (it obtains 0.39872 with ROUGE-2 and 0.42351 with ROUGE-SU4).

System	ROUGE-2	ROUGE-SU4
SUMMTERM	0.33307	0.38450
CORTEX	0.39943	0.42342
CORTEX+SUMMTERM	0.39872	0.42351
ENERTEX	0.37752	0.41605
OTS	0.36667	0.40106
Swesum	0.28391	0.33332
Pertinence	0.29756	0.33629
Microsoft Word	0.26208	0.31163
Random	0.26514	0.30356
LeadBase	0.25633	0.29554

Table 1: Numerical results of ROUGE-2 and ROUGE-SU4 evaluation.

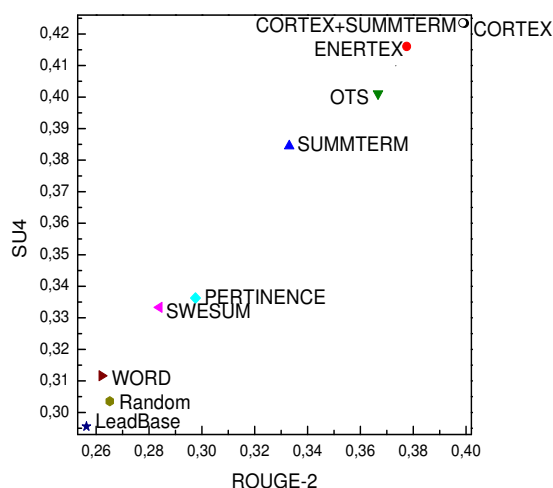


Figure 4: Illustration of ROUGE-2 and ROUGE-SU4 evaluation.

We observe that, in general, SUMMTERM summaries include information of each of the four sections of the medical articles, in the same way authors do. This fact indicates that our algorithm maintain the coherence of the original text, including relevant contents of the whole text, not only of some specific section.

A detailed analysis of the results shows that YATE is too conservative; that is, it provides as terms only those strings that have a high confidence. This is particularly true for multiword terms. Perhaps, given the high degree of specialization of the documents to be summarized, YATE should relax its considerations about what is a good term candidate. New experiments will be conducted in this sense.

Also, the term candidates list shows that terms in this type of texts are usually complex sequences, with a noun with several adjectives appearing more frequently than as usual.

It also happens that one or more components are not included in EWN or not recognized as derived from Latin forms due to their unforeseen complexity. Also the average text length is 1500 words. This length is good enough for making a summary but it is too short to allow YATE to benefit from statistical methods. Improvements in EWN coverage and in the mixing of Latin forms with regular words should also help.

6. Conclusions

In this work we have developed SUMMTERM, an innovative summarization system that combines terminological and semantic resources. We have designed the algorithm then we have evaluated it. It obtains quite good results compared to other systems as an isolated system as well as used in cooperation with an already existent summarizer. But the perception is that its performance could be better: it is necessary to continue working on this research line in order to improve the results of the algorithm.

It should also be noted that our experiment constitutes a preliminary test in order to assess the use of semantic and terminological information for automatic summarization. SUMMTERM has been developed with little effort; therefore we consider the results are fully satisfactory and there is a high improvement margin in the system, both as individual system and as a combination with other systems. As an individual system, such improvement could be done using a specialized ontology (like, for example, SNOMED-CT) in order to measure the similarity among term candidates and the improvement of the semantic similarity measure. The former will allow to reach a better domain organization and the latter to take profit not only of vertical relations (“is_a”) but also of horizontal relations like “causative_agent”, “finding_site” and “due_to”, among others. As part of a standalone summarizer like CORTEX, the above mentioned improvements together with a deeper integration may contribute to improve the final results.

Also YATE, the term extractor itself, should improve and adapt its performance according to the automatic summarization requirements. We plan to do more experiments, changing for example the weighting factor, giving more importance to the termhood or to the semantic similarity. As well, we plan to evaluate summaries of other sizes, as for example summaries of 100 words (in the same way that TAC/DUC competitions do).

Moreover, we plan to apply the algorithm to other domains (assuming EWN includes terms of such domain). In this work, we apply it to the medical domain because the term extractor we use to obtain the terms of the titles and the texts only works for some domains (Medicine, Genomics, Economics).

Finally, we would like to try to improve the resulting extracts of SUMMTERM using a module of sentence compression, since da Cunha and Molina (In press) have been proved that this technique could benefit summarization systems producing better summaries.

7. Acknowledgements

This work has been financed partially by a postdoctoral grant of the Spanish Ministry of Science and Innovation (National Program for Mobility of Research Human Resources; National Plan of Scientific Research, Development and Innovation 2008-2011) given to Iria da Cunha.

We would like to thank as well the funding of the Université d'Avignon et des Pays de Vaucluse (France) given to Jorge Vivaldi to carry out two research stages in the team TALNE (Traitement Automatique de la Lange Naturelle Écrite) of the Laboratoire Informatique d'Avignon (LIA).

8. References

- J. Abracos and G. Lopes. 1997. Statistical methods for retrieving most significant paragraphs in newspaper articles. In *Proceedings of the ACL/EACL'97 Workshop on Intelligent Scalable Text Summarization*, pages 51–57, Madrid, Spain.
- S. Afantenos, V. Karkaletsis, and P. Stamatopoulos. 2005. Summarization of medical documents: A survey. *Artificial Intelligence in Medicine*, 33(2):157–177.
- R. Barzilay and M. Elhadad. 1997. Using lexical chains for text summarization. In *Proceedings of the ACL Workshop on Intelligent Scalable Text Summarization*, pages 10–17, Madrid, Spain.
- M. T. Cabré, R. Estopà, and J. Vivaldi. 2001. Automatic Term Detection: A Review Of Current Systems. In D. Bourigault, C. Jacquemin, and M.-C. L'Homme, editors, *Recent Advances in Computational Terminology*, Amsterdam. John Benjamins Publishing Company.
- M. T. Cabré. 1999. La terminología. Representación y comunicación. In *Recent Advances in Computational Terminology*, Barcelona. IULA-UPF.
- I. da Cunha and A. Molina. In press. Optimización de resumen automático mediante comprensión de frases. In *Asociación Española de Lingüística Aplicada*.
- I. da Cunha and J. M. Torres-Moreno. 2010. Automatic discourse segmentation: Review and perspectives. In *Proceedings of the International Workshop on African Human Languages Technologies*, Djibouti (Africa).
- I. da Cunha, J. M. Torres-Moreno, P. Velázquez, and J. Vivaldi. 2009. Un algoritmo lingüístico-estadístico para resumen automático de textos especializados. *Linguística*, 2:67–79.
- I. da Cunha. 2008. *Hacia un modelo lingüístico de resumen automático de artículos médicos en español*. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona.
- L. Damianos, S. Wohlever, G. Wilson, F. Reeder, T. McEntee, R. Kozierok, and L. Hirschman. 2002. Real users, real data, real problems: The MiTAP system for monitoring bio events. In *Conference on Unified Science & Technology for Reducing Biological Threats & Counter-ing Terrorism*, pages 167–177. México.
- S. Fernández, E. SanJuan, and J. M. Torres-Moreno. 2007. Textual Energy of Associative Memories: Performant Applications of Enertex Algorithm in Text Summarization and Topic Segmentation. In *MICAL*, pages 861–871.
- S. Fernández, E. SanJuan, and J. M. Torres-Moreno. 2008. Enertex : un systme basé sur l'énergie textuelle. In *Proceedings of Traitement Automatique des Langues Naturelles*, pages 99–108, Avignon.
- R. Gaizauskas, P. Herring, M. Oakes, M. Beaulieu, P. Willett, H. Fowkes, and A. Jonsson. 2001. Intelligent access to text: Integrating information extraction technology into text browsers. *Human Language Technology Conference*, pages 189–193.
- A. Joan, J. Vivaldi, and M. Lorente. 2008. Turning a term extractor into a new domain: first experiences. In *Proceedings of the 3rd international conference on language resources and evaluation (LREC08)*, Marrakesh, Marocco.
- D. B. Johnson, Q. Zou, J. D. Dionisio, V. Liu, and W. W. Chu. 2002. Modeling medical content for automated summarization. *Annals of the New York Academy of Sciences*, 980:247–258.
- K. Kageura and B. Umino. 1996. Methods of automatic term recognition: A review. *Terminology*, 3(2):259–289.
- M. Y. Kan, K. R. McKeown, and J. L. Klavans. 2001. Domain-specific informative and indicative summarization for information retrieval. In *Workshop on text summarization (DUC 2001)*, pages 19–26. New Orleans.
- A. Lehman. 1995. *Le résumé des textes techniques et scientifiques, aspects linguistiques et computationnels*. Ph.D. thesis, Université Nancy 2, France.
- A. Lenci, R. Bartolini, N. Calzolari, A. Agua, S. Busemann, E. Cartier, K. Chevreau, and J. Coch. 2002. Multilingual summarization by integrating linguistic resources in the mlis-musi project. In *Proceedings of the 3rd international conference on language resources and evaluation (LREC02)*, pages 1464–1471, Las Palmas, Spain.
- C. Lin. 2004. Rouge: A Package for Automatic Evaluation of Summaries. In *Workshop on Text Summarization Branches Out (WAS 2004)*, pages 25–26. Barcelona.
- H. P. Luhn. 1959. The Automatic Creation of Literature Abstracts. *IBM Journal of Research and Development*, 2:159–165.
- C. D. Manning and H. Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA.
- D. Maynard. 1999. *Term recognition using combined knowledge sources*. Ph.D. thesis, Manchester Metropolitan University, Manchester.
- K. McKeown and D. Radev. 1995. Generating summaries of multiple news articles. In *18th Annual Int. Conference on Research and Development in Information Retrieval (SIGIR-95)*, pages 74–82. Seattle.
- C. D. Paice. 1990. Constructing literature abstracts by computer: Techniques and prospects. *Information Processing and Management*, 26:171–186.
- J. Pollock and A. Zamora. 1975. Automatic abstracting research at the chemical abstracts service. *Journal of Chemical Information and Computer Sciences*, 15:226–232.
- L. H. Reeve and H. Han. 2007. The Use of Domain-Specific Concepts in Biomedical Text Summarization.

- Information Processing and Management*, 43(6):1765–1776.
- E. Riloff. 1993. A Corpus-Based Approach to Domain-Specific Text Summarisation: A Proposal. In B. Endres-Niggemeyer, J. Hobbs, and K. Sparck-Jones, editors, *Workshop on Summarising Text for Intelligent Communication - Dagstuhl Seminar Report (9350)*. Dagstuhl.
- H. Saggion and G. Lapalme. 2000. Concept identification and presentation in the context of technical text summarization. In *Proceedings of the ANLP/NAACL Workshop on Automatic Summarization*, Seattle.
- H. G. Silber and K. F. McCoy. 2000. Efficient Text Summarization Using Lexical Chains. In *Proceedings of the ACM Conference on Intelligent User Interfaces*, pages 252–255, New York.
- J.M. Torres-Moreno, P. Velázquez-Morales, and J. Meunier. 2001. *CORTEX, un algorithme pour la condensation automatique de textes*. In *ARCo*, volume 2, page 365.
- J. M. Torres-Moreno, P. Velázquez-Morales, and J. G. Meunier. 2002a. Condensés de textes par des méthodes numériques. In *Journées internationales d'Analyse statistique des Données Textuelles*, pages 723–734. St. Malo.
- J.M. Torres-Moreno, P. Velázquez-Morales, and J.G. Meunier. 2002b. *Condensés de textes par des méthodes numériques*. *JADT*, 2:723–734.
- J.M. Torres-Moreno, P.-L. St-Onge, M. Gagnon, M. El-Bèze, and P. Bellot. 2009. Automatic Summarization System coupled with a Question-Answering System (QAAS). *CoRR*, abs/0905.2990.
- J. Vivaldi and H. Rodríguez. 2001a. Improving term extraction by combining different techniques. *Terminology*, 7(1):31–47.
- J. Vivaldi and H. Rodríguez. 2001b. Improving term extraction by system combination using boosting. In *Lecture Notes in Computer Science*, volume 2167, pages 515–526.
- J. Vivaldi. 2001. *Extracción de candidatos a término mediante combinación de estrategias heterogéneas*. Ph.D. thesis, Universitat Politècnica de Catalunya, Barcelona.