A summary planner based on a three-level discourse model

Thiago Alexandre Salgueiro Pardo
Lucia Helena Machado Rino
{taspardo, lucia}@dc.ufscar.br
Universidade Federal de São Carlos – Departamento de Computação
Rodovia Washington Luís, Km 235. São Carlos – Brazil

Abstract

In this paper we present a computational view of an automatic summary generator based on a discourse model that combines semantic, rhetorical and intentional knowledge in order to produce coherent summary structures. We particularly address the summary planning process.

1. Introduction

This paper presents an implementation proposal of an automatic summarization (AS) model (Rino, 1996a; Rino and Scott, 1996), primarily focusing on text, or summary, planning. This model takes into account three diverse representational levels aiming at discourse production, namely: the intentional, the rhetorical, and the semantic ones. Unlike many other works based on empirical methods (e.g., Larocca Neto et al., 2000; Black and Johnson, 1988) or fundamental, based on some sort of rhetorical structuring (e.g., Marcu, 1998a, 1998b; O'Donnell, 1997; Hovy, 1988; McKeown, 1985) or even on very limited communicative settings (e.g., Cawsey, 1993; Maybury, 1992; Moore and Paris, 1993), this work is based upon complementary knowledge sources aiming at intertwining discourse demands at both the informative and intentional settings.

In this article, we present the summary planning module of our *Discourse Modeling Summarizer* (Figure 1), or simply, *DMSumm*. We thoroughly show how discourse structuring is carried out on the basis of the 3-level discourse model (DM) and assess some of the corresponding automatically built summary plans (*SummPlans*), showing also their hypothetical linguistic realizations. The results are assessed in the light of two basic constraints: gist preservation, defined in a specific way in Rino's AS context, and communicative goal satisfaction, which is a general constraint in the discourse production scenario.

According to Figure 1, the DMSumm architecture follows the classical pipelined three-step text generation (TG) scenario. Interpretation is not a process under focus here. Instead, our DMSumm input is considered to be a message resulting from interpreting a given source text. Such a message is a composition of three information units: the central proposition (CP), the communicative goal (CG) and the knowledge base (KB), this referring back to the source text content itself. In this way, any DMSumm input message conveys the main components for discourse production: a) the CP brings about the primary discourse motivation; b) the CG builds up on the CP, by choosing discourse segments that contribute to the discourse thread; c) the KB conveys information, i.e., all the knowledge available. In DMSumm, both CP and CG are punctual information; the KB is a hierarchical semantically structured component that also includes the CP.

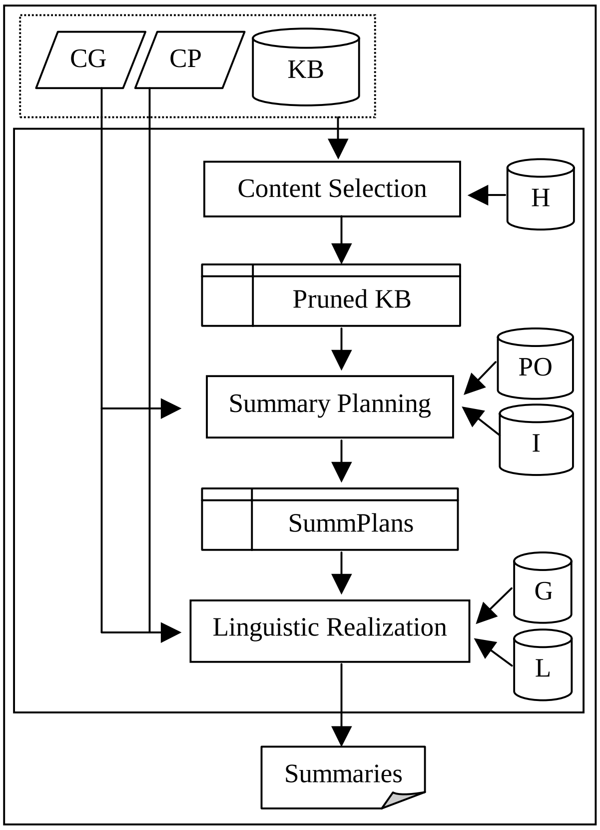


Figure 1: The DMSumm Architecture

Content selection comprises two tasks: a) reproducing both CP and CG; b) pruning the KB (using heuristics – H), which has been hand-structured[1]. Summary planning builds upon the

[1] For details of this process, see Rino and Scott (1994) and Rino (1996a).

mapping of semantic and intentional (I) relations onto rhetorical ones. For this, plan operators (POs) are used on a goal-driven basis (Maybury, 1992; Moore and Paris, 1993), which are the means to structure SummPlans. Linguistic realization is thus carried out, resulting in the surface summary itself. Since we have not yet made automatic this process, the natural language (NL) under focus here is not relevant. However, it is important to notice that planning is totally language-independent (Rino, 1996a, 1996b; Rino and Scott, 1996). So, plugging linguistic repositories such as a specific NL grammar (G) and lexicon (L) is totally circumstantial. This issue will not be addressed here.

In what follows, we present the main features of each discourse representation language (Section 2) and fully describe the summary planning process (Section 3). Some examples illustrate the DMSumm reasoning at this level of summary generation (Section 4). Important conclusions are shown in Section 5.

2. The DM Representational Levels

We highlight below the main features of each representational level: semantic, intentional and rhetorical.

2.1. The Problem-Solution

In order to define her DM, Rino has carried out a linguistic analysis of scientific texts written in English by native English speakers. Those texts, referring to the Physics domain, have driven to the Problem-Solution, hereafter P-S, paradigm (Winter, 1976; Jordan, 1980). Usual justifications for such a conclusion are that scientific discourse, in general, is structured in a well-defined logical sequence, whose identifiable super-components may well be represented by the logical sequence *Situation-Problem-Solution-Results-Evaluation-Conclusion*.
Such data had become the basis for the DMSumm input KB: at its super-structural level, there is a sort of logical sequence; at its microstructure level, information units are semantically related. Although semantic relations resemble RST (Rhetorical Structure Theory) ones, they are strictly subject-based, for they convey literal meaning and are model-dependent, in that they hold between two propositions that refer back to the P-S components.

Although the P-S model has been drawn on scientific texts, it can be systematically applied to varied genres and domains (Hoey, 1983; Jordan, 1980; Rino; 1996b). For this reason, DMSumm applies to a wide range of domains and genres. As an example of a P-S analysis, consider the following '*Fall Cushioning*' text (Hoey, 1983, p. 68), whose text segments are labeled for reference. According to Hoey, this text conveys a technological/informative genre. The KB shown in Figure 2 gives one possible interpretation for it. This can be considered the full source information for DMSumm. Underlined information refer to P-S super-components; intermediate nodes in italic refer to semantic relations that introduce the KB microstructure; the leaves refer to content information, tagged according to possible P-S components.

"Fall Cushioning"
1. Helicopters are very convenient for dropping freight by parachute, but this system has its problems.
2. Somehow the landing impact has to be cushioned to give a soft landing.
3. The movement to be absorbed depends on the weight and the speed at which the charge falls.
4. Unfortunately most normal spring systems bounce the load as it lands, sometimes turning it over.
5a. To avoid this, Bertin, developer of the aerotrain, has come up with an air-cushion system
5b. which assures a safe and soft landing.
6. It comprises a platform on which the freight is loaded with, underneath, a series of balloons supported by air cushions.
7. These are fed from compressed air cylinders equipped with an altimeter valve which opens when the load is just over six feet from the ground.
8. The platform then becomes a hovercraft, with the balloons reducing the deceleration as it touches down.
9. Trials have been carried out with freight-dropping at rates from 19 feet to 42 feet per second.
10. The charge weighed about one and half tons, but the system can handle up to eight tons.
11. At low altitudes freight can be dropped without a parachute.

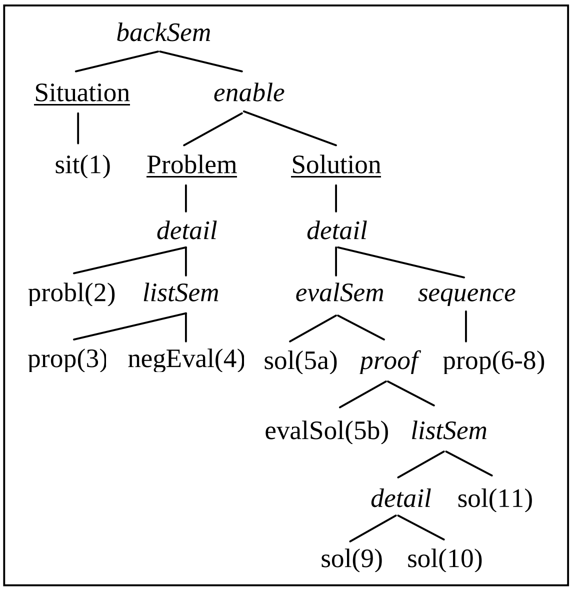


Figure 2: The "Fall Cushioning" KB

2.2. The intentional model

Intentions in Rino's DM are founded on Grosz and Sidner Discourse Theory (GSDT, 1986) in that discourse segments are inter-related by means of their contribution to each other and to the discourse as a whole. According to GSDT, any contribution is expressed through the DSeg underlying intentions, defined according to four basic intentional relations: *dominates* (DOM), *satisfaction-precedes* (SP), *supports* (SUP), and *generates* (GEN). Rino has specified further such a set, introducing the *symmetry* (SYMM) relationship, which is defined according to the following implication: DOM(X,Y), DOM(Y,X) ➔ SYMM(X,Y).

Rino's assumption for summary generation is that a rhetorical setting conveying discourse can be found based on the combination of intentions and information expressed according to the P-S components. So, intentions represent the key to summary planning, since they are responsible for voting or vetoing information segments to be rhetorically inter-related. Intentional relations are, therefore, defined according to the KB content. For the "Fall Cushioning" text, e.g., some of the possible intentional relations could be given by means of the corresponding P-S tags, as follows (*probl*: problem; *sol*: solution):

- DOM(sol,probl): conveying a problem contributes to understand a solution statement.
- SP(probl,sol): stating a problem helps presenting its corresponding solution.

Two points must be emphasized: a) intentions referring to information units can vary, in spite of the KB model being kept invariant. This is due to flexible writer's intentions during discourse production; b) although intentions have been addressed by other researchers (e.g., Moore and Paris, 1993), the approach proposed here differs from those in that it aims at an open-domain environment (although constrained to the P-S paradigm) and it conveys a broader discourse organization on intentional basis, specially addressing the writer's tailoring of his/her discourse.

2.3. The Rhetorical Model

Rino's DM relies on the assumptions underlying the Rhetorical Structure Theory, or RST (Mann and Thompson, 1987), notably: a) discourse segments in a RST tree are (supposedly) guaranteed to be coherently inter-related, once adequate RST relations have been chosen to organize discourse (Hobbs, 1985; Mann and Thompson, 1987); b) nuclearity can correspond to relevance, in that an RST nucleus conveys more relevant information than its corresponding satellite (Marcu, 1996a); c) summarization choices can be made on nuclearity basis (Marcu, 1998a, 1998b; Sparck-Jones, 1993), given the preceding assumptions. Summary planning is, thus, supposed to deliver RST SummPlans, which allow for a quite reasonable and clear correspondence with signaling surface choices. Such features allow DMSumm to be strongly staked on summary planning.

Finally, one of the most convincing reasons for having adopted RST is that it has so far been well explored for generation purposes and it interestingly allows for goal-driven planning on the basis of plan operators that a) indicate discourse segments contributing to the overall discourse; b) may be modularly defined, in that DSegs contributions are specified with respect to their neighborhood in a discourse structure; c) for the former reasons, any SummPlan can be generated incrementally. Such potentialities make evident that much can be done for summary production purposes when considering RST plans. In particular, we can address our basic constraints, namely, gist preservation and communicative goal satisfaction. To satisfy those, DMSumm is assured to keep the CP (which is assumed to convey the gist) in the leftmost nuclear position in a SummPlan (this is also suggested by Marcu, 1996a).

3. Summary Planning

In Rino's model, gist is supposed to be punctual and it is actually specified by the writer, as the CP to be conveyed. In turn, this is a leaf of both the original and the pruned KB. This guarantees that it will be preserved after selecting content. At this stage, the CG role is to indicate the kind of information that must be focused upon by the pruned KB. By considering such criteria, Rino's proposal corroborates Sparck Jones (1993) and is reinforced by Jordan (2000), in that gist dependency is considered in AS.

Rino suggests only the CGs *describe*, *report*, and *discuss* as the main ones, but these are further refined in order to observe the contributing discourse GSDT setting. Her contributing communicative goals systemic network delineates diverse chains of POs.

It is worth noticing that keeping the input message unchanged reveals the assumption that we are dealing with the same source text, under its very same unique interpretation. However, for the same source text, different interpretations are possible, given, e.g., readers' subjective expectations or background knowledge. For this reason, Rino allows for variations of CG and CP, keeping the KB unchanged. Illustrations on this are given in her thesis (1996a) and are taken as base-examples in DMSumm. In allowing for both varying CPs and CGs at the input message level and varying planning strategies to satisfy CGs, a wide multiplicity of summaries is made possible. This, in turn, brings about assessment issues that will be explored later in this article.

Possible mappings of intentions and semantic relations onto RST relations are shown in Table 1 (Rino defined only those 10 cases, based on her linguistic analysis). The table indicates how constraints at the informational (thus, semantic) and intentional levels must be handled in order to generate a discourse (thus, rhetorical) structure. For example, the first case indicates that, if an *enable* relation between information segments X and Y (X *enables* Y) in the KB and *SP* (*X SP* Y) and *DOM* (*Y DOM* X) relations hold at the intentional level, then either a *means* (X is the means to Y) or a *purpose* (Y is the goal to be achieved through X) relation can hold at the rhetorical level. Only one of them will, thus, be inserted in the SummPlan. Figure 3 shows a PO that operates on such a mapping for describing any concept X.

Table 1: Mapping of intentions and semantics onto rhetoric

Semantics	Rhetoric	Intentions
enable(Y,X)	purpose1(X,Y) means(Y,X)	X sp Y Y dom X
rationale(X,Y)	purpose2(X,Y) justify1(X,Y)	Y sup X X dom Y ¬ X sp Y
proof(X,Y)	evidence(X,Y) justify2(X,Y)	Y sp X X dom Y Y sup X
cause(Y,X)	nonvolresult(Y,X) nonvolcause(X,Y)	Y sp X X dom Y Y gen X
listSem(X,Y)	list(X,Y) contrast(X,Y)	X symm Y
attribute(X,Y) detail(X,Y) exemplify(X,Y)	elaborate(X,Y)	Y sup X
evalSem(X,Y)	evaluate(X,Y)	X dom Y
reason(X,Y)	explain(X,Y)	Y gen X
sequence(X,Y)	sequence(X,Y)	X sp Y
backSem(X,Y)	background(X,Y)	¬ X sp Y Y sup X

Name	describe-by-means
Header	describe(X)
Effect	know-about(reader,X)
Constraints	not know-about(reader,X), isa(X,Y,enable), X sp Y, Y dom X
Nucleus	describe(X)
Satellite	know(reader,means(Y,X))

Figure 3: A PO describing a concept through the *means* RST relation

Only content selection and summary planning have so far been made fully automatic. Linguistic realization has been carried out by hand, on the RST relations interpreting basis. For the time being, this is fine, since we are concerned with the DMSumm deep generation level and, mainly, with the consistency of both, the POs specification and their corresponding application.

Currently, the PO repository consists of ca. 50 POs, which have been designed on a case studies basis involving three KBs and six different combinations of CGs and CPs, resulting in 181 SummPlans. Table 2 shows the resulting data according to varied input messages.

Table 2: DMSumm case-based performance

	CG	PC	SummPlans
KB1	*describe*	*result*	**123**
	describe	*problem*	**2**
	discuss	*new problem*	**3**
KB2	*report*	*solution*	**45**
KB3	*report*	*situation*	**4**
	report	*method*	**4**
		Total	**181**

Two distinct approaches in applying POs were undertaken here, considering the way they were first specified: PO groups were tailored to each case study, i.e., their specification was fully example-based. Focusing upon all the examples, the full PO repository was obtained. Summary planning can thus be input-driven or unconstrained. In the former case, only those strategies (or PO groups) that were customized to the corresponding case are considered to be applicable. In the latter, any strategy allowed by the full PO repository may be indistinctly applied to any input. The SummPlans shown in Table 2, for instance, were obtained through the unconstrained application of PO groups. Applying specific strategies to the KB of Figure 2 (Section 2), e.g., we obtained 20 SummPlans; applying non-customized strategies, 45 SummPlans were generated instead. By carrying out a subjective evaluation, all of those were considered coherent. Besides assessing the coherence of the SummPlans, we can also assess their complexity, mirrored by their structural depth in Table 3. For KB1.1 in the table, for example, we shall read the following summary generation history: 1 SummPlan with depth 1; 3 SummPlans with depth 2; and so on. Although complexity is definitely not dependent solely on the structural features, such illustrations give a good idea about the complexity related to both summary planning and the subsequent linguistic realization.

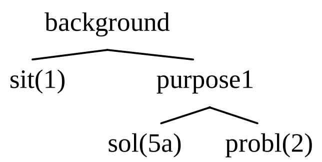
Table 3: SummPlans' Depth

	SummPlans' Depth						
	1	*2*	*3*	*4*	*5*	*6*	*7*
KB1.1	**1**	**3**	**14**	**35**	**46**	**21**	**3**
KB1.2	**1**	**1**	**0**	**0**	**0**	**0**	**0**
KB1.3	**1**	**0**	**0**	**0**	**0**	**2**	**0**
KB2	**1**	**3**	**5**	**15**	**16**	**5**	**0**
KB3.1	**1**	**1**	**0**	**2**	**0**	**0**	**0**
KB3.2	**1**	**0**	**2**	**1**	**0**	**0**	**0**
Total	6	8	21	53	62	28	3

4. Illustrating summary planning

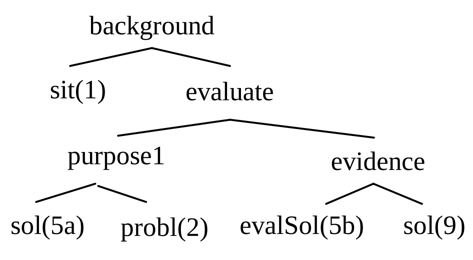
We present below an account on summary planning, considering the 'Fall Cushioning' text as source. Let us assume that we want to report the *adopted solution*, i.e., our CP is *sol(5a)* and our CG is *report*, in our representation language.

In applying diverse summary planning strategies, the SummPlans were automatically generated. Two of them are shown in figures 4 and 5 below. The corresponding summaries have been hand-generated.



Helicopters are very convenient for dropping freight by parachute, but this system has its problems. The landing impact has to be cushioned to give a soft landing. To avoid this, Bertin, developer of the aerotrain, has come up with an air-cushion system.

Figure 4: SummPlan 1



Helicopters are very convenient for dropping freight by parachute, but this system has its problems. The landing impact has to be cushioned to give a soft landing. To avoid this, Bertin, developer of the aerotrain, has come up with an air-cushion system, which assures a safe and soft landing. Trials have been carried out with freight-dropping at rates from 19 feet to 42 feet per second.

Figure 5: SummPlan 2

Corroborating the CP assumption and, also, Marcu's (1996a) claim that the most important, or the most salient, information of a text is the most nuclear leaf of a rhetorical tree, the CP is kept in the leftmost leaf of any illustrated SummPlan. Moreover, the CP is guaranteed to always be a SummPlan explicit component, as claimed by Rino. Although we describe here solely summary planning, it is worth noticing that summarizing choices occur in any of the 3-step summary generator process: a) any information may be selected; b) POs allow for specific decisions to be taken incrementally; and c) the same thing can be said in several ways (Souza and Scott; 1990).

5. Assessing DMSumm potentialities

We have worked with the discourse scenario presented in Table 1 and some case studies have been addressed. Although this is quite limited, an initial analysis has shown that all SummPlans consistently and correspondingly convey the CP and satisfy the CG. Moreover, they were considered coherent, when we map RST relations onto surface choices.

Even though Marcu (1996b) pinpoints weaknesses of approaches based on rhetorical relations, our DMSumm model still improves on gist preservation and goal satisfaction, when compared to other approaches. This is due to the fact that a) it is domain and genre-independent, although it has been originated on a specific discourse basis; b) intentionality is combined with information in order to choose a particular rhetorical organization. This is based upon both the original content, along with their related context dependencies, and a pre-established contribution setting between discourse segments, founded on the GSDT and the P-S models.

The main bottlenecks of our DMSumm, concerning the proposed deep approach, refer back to its complexity: a) interpretation has not yet been addressed, and this evidences its fragility, if we aim at having a full text summarizer in the future; b) RST structuring brings to our system all its inherent problems, already widely explored by other researchers. However, most of similar proposed models have proven that it is still worthwhile considering such foundations.

The results obtained so far are promising in that they convey coherent structures that can be well realized. They provide evidence that our proposal is theoretically sound. However, substantial further analysis must be carried out after having the full DMSumm system providing real texts. We are now comparing our results with those obtained from statistical summarizers. We will also submit the summaries to human evaluators, in experiments to assess in detail their thematic progression and their coherence and cohesive patterns.

Acknowledgments

This work is funded by FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo (Proc. 99/11954-6).

6. References

Black, W.J. and Johnson, F.C. (1988). A Practical Evaluation of Two Rule-Based Automatic Abstraction Techniques. In: *Expert Systems for Information Management*, Vol. 1, No. 3. Department of Computation. University of Manchester Institute of Science and Technology.

Cawsey, A. (1993). Planning Interactive Explanations. In: *Int. Journal of Man-Machine Studies*, Vol. 38, pp. 169-199.

Grosz, B. and Sidner, C. (1986). Attention, Intentions, and the Structure of Discourse. In: *Computational Linguistics*, Vol. 12, No. 3.

Hobbs, J.R. (1985). *On the Coherence and Structure of Discourse*. Technical Report CSLI-85-37, Center for Study of Language and Information, Stanford University.

Hoey, M. (1983). *On the Surface of Discourse*. George Allen & Unwin Ltd.

Hovy, E. (1988). *Generating Natural Language under Pragmatic Constraints*. Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey.

Jordan, M.P. (1980). Short Texts to Explain Problem-Solution Structures – and Vice Versa. In: *Instructional Science*, Vol. 9, pp. 221-252

Jordan, M,P. (2000). A Pragmatic/Structural Approach to Relevance. In: *Technostyle*, Vol. 16, No. 2, pp. 47-67.

Larocca Neto, J., Santos, A.D., Kaestner, A.A., Freitas, A.A. (2000). Generating Text Summaries through the Relative Importance of Topics. In: *Proceedings of the International Joint Conference IBERAMIA/SBIA*, Atibaia, SP.

Mann, W.C. and Thompson, S.A. (1987). *Rhetorical Structure Theory: A Theory of Text Organization*. Technical Report ISI/RS-87-190.

Marcu, D. (1996a). Building Up Rhetorical Structure Trees. In: *The Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Vol. 2, pp. 1069-1074. Portland, Oregon.

Marcu, D. (1996b). Distinguishing between Coherent and Incoherent Texts. In: *The Proceedings of the Student Conference on Computational Linguistics in Montreal*, pp. 136-143. Montreal, Canada.

Marcu, D. (1998a). Improving Summarization through Rhetorical Parsing Tuning. In: *The Sixth Workshop on Very Large Corpora*, pp. 206-215. Montreal, Canada.

Marcu, D. (1998b). To build text summaries of high quality, nuclearity is not sufficient. In: *The Working Notes of the AAAI-98 Spring Symposium on Intelligent Text Summarization*, pp. 1-8. Stanford, California.

Maybury, M.T. (1992). Communicative Acts for Explanation Generation. In: *Int. Journal of Man-Machine Studies 37*, pp. 135-172.

McKeown, K.R. (1985). *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press.

Moore, J.D. and Paris, C. (1993). Plannig Text for Advisory Dialogues: Capturing Intentional and Rhetorical Information. In: *Computational Linguistics*, Vol. 19, No. 4, pp. 651-694.

O'Donnell, M. (1997). Variable-Length On-Line Document Generation. In: *Proceedings of the 6th European Workshop on Natural Language Generation*, Gerhard-Mercator University, Duisburg, Germany.

Rino, L.H.M. (1996a). *Modelagem de Discurso para o Tratamento da Concisão e Preservação da Idéia Central na Geração de Textos*. Tese de Doutorado. IFSC-Usp. São Carlos - SP.

Rino, L.H.M. (1996b). A sumarização automática de textos em português. In: *Anais do II Encontro para o Processamento Computacional de Português Escrito e Falado*, pp. 109-119. Curitiba - PR.

Rino, L.H.M. and Scott, D. (1994). *Automatic generation of draft summaries: heuristics for content selection.* ITRI Techn. Report ITRI-94-8. University of Brighton, England.

Rino, L.H.M. and Scott, D. (1996). A Discourse Model for Gist Preservation. In: Dibio L. Borges and Celso A. A. Kaestner (eds.), Advances in Artificial Intelligence. *Lecture Notes in Artificial Intelligence*, Vol. 1159, pp. 131-140.

Scott, D.R. and Souza, C.S. (1990). Getting the Message Across in RST-Based Text Generation. In: R. Dale, C. Mellish and M. Zock (eds), *Current Research in Natural Language Generation*, pp. 47-73, London, Academic Press.

Sparck Jones, K. (1993). What might be in a summary? In: Knorz, Krause and Womser-Hacker (eds.). *Information Retrieval: Von der Modellierung zur Anwendung*, Vol. 93, pp. 9-26, Universitatsverlag Konstanz.

Winter, E.O. (1976). *Fundamentals of Information Structure*. Hatfield Polytechnic, Hertfodshire, England.