

A Hierarchical Account of Referential Accessibility

Nancy IDE

Department of Computer Science
Vassar College
Poughkeepsie, New York 12604-0520 USA
ide@cs.vassar.edu

Dan CRISTEA

Department of Computer Science
University “Al. I. Cuza”
Iasi, Romania
dcristea@infoiasi.ro

Abstract

In this paper, we outline a theory of referential accessibility called Veins Theory (VT). We show how VT addresses the problem of “left satellites”, currently a problem for stack-based models, and show that VT can be used to significantly reduce the search space for antecedents. We also show that VT provides a better model for determining domains of referential accessibility, and discuss how VT can be used to address various issues of structural ambiguity.

Introduction

In this paper, we outline a theory of referential accessibility called Veins Theory (VT). We compare VT to stack-based models based on Grosz and Sidner’s (1986) focus spaces, and show how VT addresses the problem of “left satellites”, i.e., subordinate discourse segments that appear prior to their nuclei (dominating segments) in the linear text. Left-satellites pose a problem for stack-based models, which remove subordinate segments from the stack before pushing a nuclear or dominating segment, thus rendering them inaccessible. The percentage of such cases is typically small, which may account for the fact that their treatment has been largely overlooked in the literature, but the phenomenon nonetheless persists in most texts. We also show how VT can be used to address various issues of structural ambiguity.

1 Veins Theory

Veins Theory (VT) extends and formalizes the relation between discourse structure and reference proposed by Fox (1987). VT identifies “veins” over discourse structure trees that are built according to the requirements put forth in Rhetorical Structure Theory (RST) (Mann and Thompson, 1987). RST structures

are represented as binary trees, with no loss of information. Veins are computed based on the RST-specific distinction between nuclei and satellites; therefore, RST relations labeling nodes in the tree are ignored. Terminal nodes in the tree represent *discourse units* and non-terminal nodes represent *discourse relations*.

The fundamental intuition underlying VT is that the distinction between nuclei and satellites constrains the range of referents to which anaphors can be resolved; in other words, the nucleus-satellite distinction induces a *domain of referential accessibility* (DRA) for each referential expression. More precisely, for each anaphor a in a discourse unit u , VT hypothesizes that a can be resolved by examining referential expressions that were used in a subset of the discourse units that precede u ; this subset is called the DRA of u . For any elementary unit u in a text, the corresponding DRA is computed automatically from the text’s RST tree in two steps:

1. *Heads* for each node are computed bottom-up over the rhetorical representation tree. Heads of elementary discourse units are the units themselves. Heads of internal nodes, i.e., discourse spans, are computed by taking the union of the heads of the immediate child nodes that are nuclei. For example, for the text in Figure 1,¹ with the rhetorical structure shown in Figure 2,² the head of span [5,7] is unit 5. Note that the head of span [6,7] is the list <6,7> because both immediate children are nuclei.
2. Using the results of step 1, *Vein expressions* are computed top-down for

¹ Figure 1 highlights two co-referential equivalence classes: referential expressions surrounded by boxes refer to “Mr. Casey”; those surrounded by ellipses refer to “Genetic Therapy Inc.”.

² The rhetorical structure is represented using the conventions proposed by Mann and Thompson (1988).

each node in the tree, using the following functions:

- *mark*(*x*), which returns each symbol in a string of symbols *x* marked with parentheses.
- *seq*(*x*,*y*), which concatenates the labels in *x* with the labels in *y*, left-to-right.
- *simpl*(*x*), which eliminates all marked symbols from *x*, if they exist.

The vein of the root is its head. Veins of child nodes are computed recursively, as follows:

- for each nuclear node whose parent has vein *v*, if the node has a left non-nuclear sibling with head *h*, then the vein expression is *seq*(*mark*(*h*), *v*); otherwise *v*.
- for each non-nuclear node with head *h* whose parent node has vein *v*, if the node is the left child of its parent, then *seq*(*h*,*v*); otherwise, *seq*(*h*, *simpl*(*v*)).

One of the conjectures of VT is that the vein expression of a unit (terminal node), which includes a chain of discourse units that contain that unit itself, provides an “abstract” or summary of the discourse fragment that contains that unit. Because it is an internally coherent piece of discourse, all referential expressions (REs) in the unit preferentially find their referees within that sub-text. Referees that do not appear in the DRA are possible, but are more difficult to process, both computationally and cognitively (see Section 2.2). This conjecture expresses the intuition that potential referees of the REs of a unit depend on the nuclearity of previous units: both a satellite and a nucleus can access a previous nuclear node, a nucleus can only access another left nuclear node or its own left satellite, and the interposition of a nucleus after a satellite blocks the accessibility of the satellite for any nodes lower in the hierarchy.

1. Michael D. Casey, a top Johnson & Johnson manager, moved to Genetic Therapy Inc., a small biotechnology concern here,
2. to become its president and chief operating officer

3. Mr. Casey, 46, years old, was president of J&J's McNeil Pharmaceutical subsidiary,
4. which was merged with another J&J unit, Ortho Pharmaceutical Corp., this year in a cost-cutting move.
5. Mr. Casey succeeds M. James Barrett, 50, as president of Genetic Therapy
6. Mr. Barrett remains chief executive officer
7. and becomes chairman.
8. Mr. Casey said
9. he made the move to the smaller company
10. because he saw health care moving toward technologies like the company's gene therapy products.
11. I believe that the field is emerging and is prepared to break loose,
12. he said.

Figure 1: MUC corpus text fragment

The DRA of a unit *u* is given by the units in the vein that precede *u*. For example, for the text and RST tree in Figures 1 and 2, the vein expression of unit 3, which contains units 1 and 3, suggests that anaphors from unit 3 should be resolved only to referential expressions in units 1 and 3. Because unit 2 is a satellite to unit 1, it is considered to be “blocked” to referential links from unit 3. In contrast, the DRA of unit 9, consisting of units 1, 8, and 9, reflects the intuition that anaphors from unit 9 can be resolved only to referential expressions from unit 1, which is the most important unit in span [1,7] and to unit 8, a satellite that immediately precedes unit 9. Figure 2 shows the heads and veins of all internal nodes in the rhetorical representation. In general, co-referential relations (such as the identity relation) induce equivalence classes over the set of referential expressions in a text. When hierarchical adjacency is considered, an anaphor may be resolved to a referent that is not the closest in a linear interpretation of a text. However, because referential expressions are organized in equivalence classes, it is sufficient that an anaphor is resolved to *some* member of the set. This is consistent with the distinction between “direct” and “indirect” references discussed in (Cristea, *et al.*, 1998).

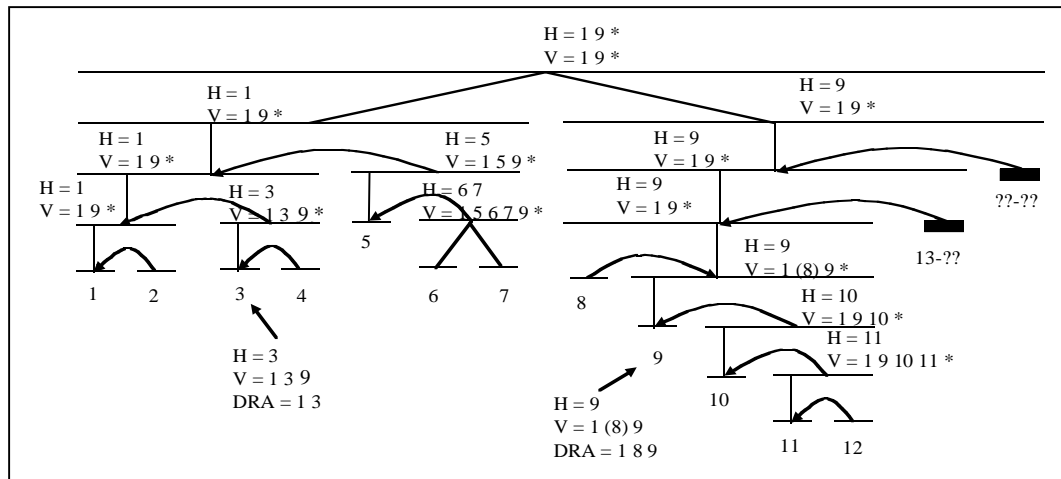


Figure 2: RST analysis of the text in Figure 1

2 VT and Stack-based Models

Veins Theory claims that references from a given unit are possible only in its DRA, i.e., that discourse structure constrains the areas of the text over which references can be resolved. In previous work, we compared the potential of hierarchical and linear models of discourse--i.e., approaches that enumerate potential antecedents in an undifferentiated window of text linearly preceding the anaphor under scrutiny--to correctly establish co-referential links in texts, and hence, their potential to correctly resolve anaphors (Cristea, *et al.*, 2000). Our results showed that by exploiting the hierarchical discourse structure of texts, one can increase the potential of natural language systems to correctly determine co-referential links, which is a requirement for correctly resolving anaphors. In general, the potential to correctly determine co-referential links was greater for VT than for linear models when one looks back 4 elementary discourse units. When looking back more than four units, the linear model was equally effective.

Here, we compare VT to stack-based models of discourse structure based on Grosz and Sidner's (1986) (G&S) focus spaces (e.g., Hahn and Strübe, 1997; Azzam, *et al.*, 1998). In these approaches, discourse segments are pushed on the stack as they are encountered in a linear traversal of the text. Before a dominating segment is pushed, subordinate segments that precede it are popped from the stack. Antecedents for REs appearing in the segment on the top of the stack are sought in discourse segments in the stack below it. Therefore, in cases where a subordinate segment *a* precedes a dominating segment *b*, a reference to an entity in

a by an RE in b is not resolvable. Special provision could be made in order to handle such cases—e.g., subsequently pushing a on top of b —but this would violate the overall strategy of resolving REs appearing in segments currently on the top of the stack.

The special status given to left satellites in VT addresses this problem. For example, one RST analysis³ of (1) proposed by Moser and Moore (1996) is given in

Figure 3. Moser and Moore note that the relation of an RST nucleus to its satellite is analogous to the *dominates* relation proposed by G&S (see also Marcu, 2000). As a subordinate segment preceding the segment that dominates it, the satellite is popped from the stack before the dominant segment (the nucleus) is pushed in the stack-based model, and therefore it is not included among the discourse segments that are searched to resolve co-references.⁴ Similarly, the text in (2), taken from the MUC annotated corpus (Marcu, et al., 1999), was assigned the RST structure in Figure 4, which presents the same problem for the stack-based approach: the referent for *this* in C2 is to *the Clinton program* in A2, but because it is a subordinate segment, it is no longer on the stack when C2 is processed.

³ Note that the left satellite problem arises as a result of gross discourse segmentation rather than unit size, and so our argument applies equally to analyses based on discourse theories other than RST (e.g., Centering Theory) in which unit size may be determined differently.

⁴ Moser and Moore (1996) also propose an *informational* RST structure for the same text, in which a « volitional-cause » relation holds between the nucleus *a* and the satellite *b*, thus providing for *a* to be on the stack when *b* is processed.

- (1) A1. George Bush supports big business.
B1. He's sure to veto House Bill 1711.

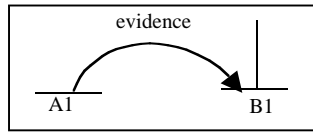


Figure 3: RST analysis of (1)

- (2) A2. Some of the executives also signed letters on behalf of the Clinton program.
B2. Nearly all of them praised the president for his efforts to pare the deficit.
C2. This is not necessarily the package I would design,
D2. said Martin Marietta's Mr. Augustine.
E2. But we have to attack the deficit.

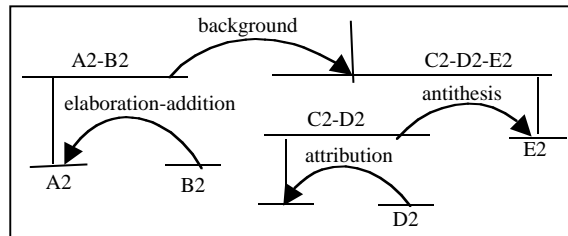


Figure 4: RST analysis of (2)

2.1 Validation

To validate our claim, we examined 23 newspaper texts with widely varying lengths (mean length = 408 words, standard deviation 376). The texts were annotated manually for co-reference relations of identity (Hirschman and Chinchor, 1997). The co-reference relations define equivalence relations on the set of all marked references in a text. The texts were also annotated manually with discourse structures built in the style of Mann and Thompson (1988). Each analysis yielded an average of 52 elementary discourse units. Details of the annotation process are given in (Marcu *et al.*, 1999).

Six percent of all co-references in the corpus are to left satellites. If only co-references pointing outside the unit in which they appear (inter-unit references) are considered, the rate increases to 7.76%. Among these cases, two possibilities exist: either the reference is unresolvable using the stack-based method because the unit in which the referent appears has been popped from the stack, or the stack-based algorithm finds a correct referent in an earlier unit that is still on the stack. Twenty-two percent (2.38% of all co-referring expressions in the corpus) of the

referents that VT finds in left satellites fall into the first category. For example, in text fragment (3), taken from the MUC corpus, the co-referential equivalence class for the pronoun *he* in C3 includes *Saloman Brothers analyst Jeff Canin* in B3 and *he* in A3. The RST analysis of this fragment in Figure 5 shows that both A3 and B3 are left satellites. A stack-based approach does not find either antecedent for *he* in C3, since both A3 and B3 are popped from the stack before C3 is processed.

- (3) A3. Although the results were a little lighter than the 49 cents a share he hoped for,
B3. Salomon Brothers analyst Jeff Canin said
C3. he was pleased with Sun's gross margins for the quarter.

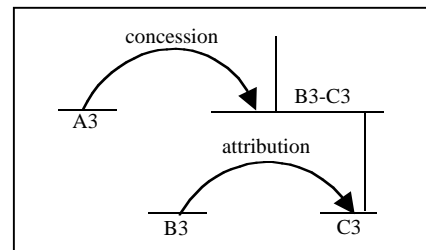


Figure 5: RST analysis of (3)

For the second case, where stack-based approaches find a co-referent (although not the most recent antecedent) elsewhere in the stack, it makes sense to compare the *effort* required by the two models to establish correct co-referential links. That is, we assume that from a computational perspective (and, presumably a psycholinguistic one as well), the closer an antecedent is to the referential expression to be resolved, the better. We have shown elsewhere (Cristea *et al.*, 2000) that when VT is compared to linear models, VT requires significantly less effort for DRAs of any size.

We use a similar strategy here to compute the effort required by VT and stack-based models. DRAs for both models are treated as ordered lists. For example, text fragment (4) reflects the set of units on the stack at a given point in processing one of the MUC texts; units D4 and E4, in brackets, are left satellites and therefore not available using the stack-based model. To determine the correct antecedent of *Mr. Clinton* in F4, it is necessary to search back through 3 units (C4, B4, A4) to find the co-referent *President Clinton*. In contrast, using VT, we search back only one unit to find *Clinton* in unit E4.

- (4) A4. A group of top corporate executives urged Congress to pass President Clinton's deficit-reduction plan,
- B4. declaring that it is superior to the only apparent alternative: more gridlock.
- C4. Some of the executives who attended yesterday's session weren't a surprise.
- [D4. Tenneco Inc. Chairman Michael Walsh, for instance, is a staunch Democrat who provided an early endorsement for Mr. Clinton during the presidential campaign.
- E4. Xerox Corp.'s Chairman Paul Allaire was one of the few top corporate chief executive officers who contributed money to the Clinton campaign.]
- F4. And others, such as Atlantic Richfield Co. Chairman Lodwick M. Cook and Zenith Electronics Corp. Chairman Jerry Pearlman, have also previously voiced their approval of Mr. Clinton's economic strategy.

We computed the effort $e(M, a, \text{DRA}_k, u)$ of a model M to determine correct co-referential links with respect to a referential expression a in unit u , given a DRA of size k . The effort is given by the number of units between u and the first unit in DRA_k that contains a co-referential expression of a . Obviously, for a given k and a given referent a , it is possible that no co-reference exists in the units of the corresponding DRA_k ; in these cases, we consider the effort to be equal to k . The effort $e(M, C, k)$ of a model M to determine correct co-referential links for all referential expressions in a corpus of texts C using DRAs of size k is computed as the sum of the efforts $e(M, a, \text{DRA}_k(u))$ of all referential expressions a .

We computed the effort for all referential expressions for which VT finds a co-referent in a left satellite, and for which the stack-based model finds *some* co-referent elsewhere in the text (i.e., we do not include cases where VT finds a co-referent and the stack-based model does not). Figure 6 shows the effort for VT and stack-based models for each k from 1 to 12, which represents the largest DRA required for either model to find a co-referent. For small k the effort required to establish co-referential links is similar for both models, because both can establish only a limited number of links. However, as k increases, the effort computed over the entire corpus diverges, with VT performing consistently better than the stack-based model.

In some cases, the stack-based model performs better than VT, in particular for small k . This occurs when VT searches back through n

adjacent left satellites, where $n > 1$, to find a co-reference, but a co-referent is found using the stack-based method by searching back m non-left satellite units, where $m < n$. This would be the case if, for instance, VT did not find a co-referent for *Mr. Clinton* in unit F4 from text (4) in the left satellite segment (therefore searching back 3 units to find a co-referent), but the stack-based model found a co-referent in C4 (1 unit away since the left satellites are not on the stack).

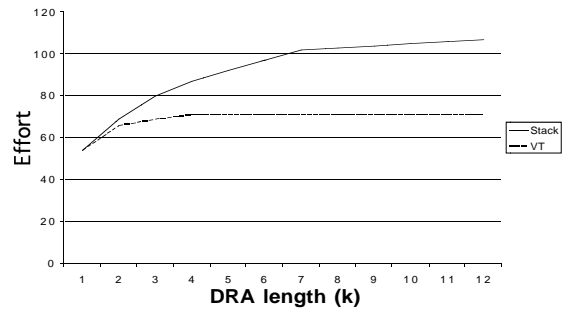


Figure 6: Effort required by VT and stack-based models

In our corpus, 15% of the co-references found in left satellites by VT required less effort using the stack-based method. All of the cases in which the stack-based model performed better are for small k ($k < 4$), and the average difference in distance (in units) is 1.25. In contrast, the average difference in distance in cases where VT out-performs the stack-based model is 3.8 units. This makes intuitive sense: because VT's DRAs will contain a co-referent found by the stack-based model, even if units within a left satellite intervene, the difference in distance can be no larger than the number of units in the left satellite segment. However, in cases where the stack-based model "misses" a co-referent contained in a left satellite, the maximum possible distance required to find a co-referent is the number of units in the DRA for the unit in question, computed over *all* segments currently on the stack, which is potentially much greater. At $k=4$, VT can determine all the co-referents in our corpus, whereas the stack-based model requires DRAs of up to 12 units to resolve them all. This accounts for the marked divergence in effort shown in Figure 6 as k increases. Although our corpus is necessarily small, these results suggest that VT has the potential to significantly reduce the search space required to establish co-referential links.

2.2 Exceptions

We have also examined the exceptions, i.e., co-referential links that VT and stack-based models cannot determine correctly. Because of the equivalence of the stack contents for left-balanced discourse trees, there is no case in which the stack-based model finds a referent where VT does not. There is, however, a number of referring expressions for which neither VT nor the stack-based model finds a co-referent. In the corpus of MUC texts we consider, 12.3% of inter-unit references fall into this category, or 9.3% of the references in the corpus if we include intra-unit references.

Table 1 provides a summary of the types of referring expressions for which co-referents are not found in our corpus—i.e., no antecedent exists, or the antecedent appears outside the DRA.⁵ We show the percentage of REs in our corpus for which VT (and the stack-based model as well, since all units in the DRA computed according to VT are in the DRA computed using the stack-based model) fails to find an antecedent, and the percentage of REs for which VT finds a co-referent (in a left satellite) but the stack-based model does not.

| Type of RE | VT | Stack-based |
|--------------|-------|-------------|
| pragmatic | 56.3% | 0.0% |
| proper nouns | 22.7% | 26.1% |
| common nouns | 16.0% | 39.1% |
| pronouns | 5.0% | 34.8% |

Table 1: Exceptions for VT and stack-based models, by type

We consider four types of REs:

- (1) Pragmatic references, which refer to entities that can be assumed part of general knowledge, such as *the Senate*, *the key* in the phrase *lock them up and throw away the key*, or *our* in the phrase *our streets*.
- (2) Proper nouns, such as *Mr. Gerstner* or *Senator Biden*.
- (3) Common nouns, such as *the steelmaker*, *the proceeds*, or *the top job*.
- (4) Pronouns

Following (Gundel, *et al.*, 1993), we consider that the *evoking power* of each of these types of REs decreases as we move down the list. That is,

⁵ Our calculations are made based on the RST analysis of the MUC data, in which we detected a small number of structural errors. Therefore, the values given here are not absolute but rather provide an indication of the relative distribution of RE types.

pragmatic references are easily understood without an antecedent; proper nouns and noun phrases less so, and are typically resolved by inference over the context. On the other hand, pronouns have very poor evoking power, and therefore a message emitter employs them only when s/he is certain that the structure of the discourse allows for easy recuperation of the antecedent in the message receiver's memory.⁶ Except for the cases where a pronoun can be understood without an antecedent (e.g., *our* in *our streets*), it is virtually impossible to use a pronoun to refer to an antecedent that is outside the DRA.

The alignment of the evoking power of referential expressions with the percentage of exceptions for both models shows that the predictions made by VT relative to DRAs are fundamentally correct—that is, their prevalence corresponds directly to their respective evoking powers. On the other hand, the almost equal distribution of exceptions over RE types for the stack-based model shows that it is less reliable for determining DRAs.

Note that in all VT exceptions for pronouns, the RST *attribution* relation is involved. Text fragment (5) and the corresponding RST tree (Figure 7) show the typical case:

- (5) A5. A spokesman for the company said,
 B5. Mr. Bartlett's promotion reflects the current emphasis at Mary Kay on international expansion.
 C5. Mr. Bartlett will be involved in developing the international expansion strategy,
 D5. he said

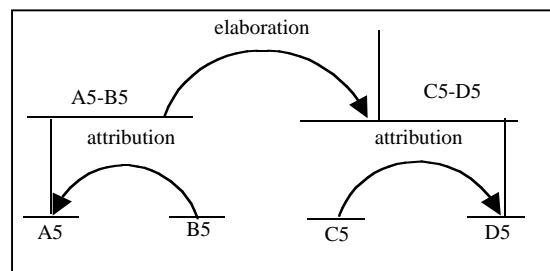


Figure 7: RST analysis of (5)

The antecedent for *he* in D5 is *a spokesman for the company* in A5, which, due to the nuclear-satellite relations, is inaccessible on the vein. Our results suggest that annotation of attributive

⁶ Ideally, a psycho-linguistic study of reading times to verify the claim that referees outside the DRA are more difficult to process would be in order.

relations needs to be refined, possibly by treating *X said* and the attributed quotation as a single unit. If this were done, the vein expression would allow appropriate access.

2.3 Summary

In sum, VT provides a more natural account of referential accessibility than the stack-based model. In cases where the discourse structure is not left-polarized, at least one satellite precedes its nucleus in the discourse and is therefore its left sibling in the binary discourse tree. The vein definition formalizes the intuition that in a sequence of units *a b c*, where *a* and *c* are satellites of *b*, *b* can refer to entities in *a* (its left satellite), but the subsequent right satellite, *c*, cannot refer to *a* due to the interposition of nuclear unit *b*--or, if such a reference exists, it is harder to process. In stack-based approaches to referentiality, such configurations pose problems: because *b* dominates *a*, in order to resolve potential references from *b* to *a*, *b* must appear below *a* on the stack even though it is processed after *a*. Even if the processing difficulties are overcome, this situation leads to the postulation of cataphoric references when a satellite precedes its nucleus, which is counter-intuitive.

3 VT and Structural Ambiguity

The fact that VT considers only the nuclear-satellite distinction and ignores rhetorical labeling has practical ramifications for anaphora resolution systems that rely on discourse structure to determine the DRA for a given RE. (Marcu, et al., 1999) show that over a corpus of texts drawn from MUC newspaper texts, the *Wall Street Journal* corpus, and the Brown Corpus, reliable agreement among annotators is consistently obtained for discourse segmentation and assignment of nuclear-satellite status, while agreement on rhetorical labeling was less reliable (statistically significant for only the MUC texts). This means that even when there exist differences in rhetorical labeling, vein expressions can be computed and used to determine DRAs.

VT also has ramifications for evaluating the viability of different structural representations for a given text, at least for the purposes of reference resolution. Like syntactic parsing, discourse parsing typically yields several interpretations, and one of the *a priori* tasks for further analysis of the parsed texts is to choose

one from among potentially several alternative structures. Marcu (1996) showed that using only rhetorical relations, as many as five different structures can be identified for some texts. Considering intention-based relations can yield even more alternatives. For anaphora resolution, the choice of one structure over another may have significant impact. For example, an RST tree for (6) using rhetorical relations is given in Figure 9; Figure 10 shows another RST tree for the same text, using intention-based relations. If we compute the vein expressions for both representations, we see that the vein for segment C6 in the intentional representation is <A6 B6 C6>, whereas in the rhetorical representation, the vein is <(B6), C6>. That is, under the constraints imposed by VT, *John* is not available as a referent for *he* in C6 in the rhetorical version, although *John* is clearly the appropriate antecedent. Interestingly, the intention-based analysis is skewed to the right and thus is a "better" representation according to the criteria outlined in (Marcu, 1996); it also eliminates the left-satellite that was shown to pose problems for stack-based approaches. It is therefore likely that in this case, the intention-based analysis is "better" for the purposes of anaphora resolution.

- (6) A6. Tell John to bring the car home by 5.
 B6. That way I can get to the store before it closes.
 C6. Then he can finish the bookshelves tonight.

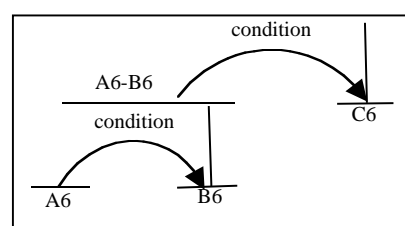


Figure 8: RST tree for text (6), using rhetorical relations

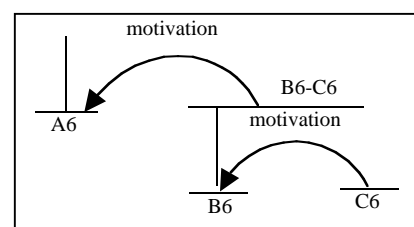


Figure 9: RST tree for text (6), using intention-based relations

Conclusion

Veins Theory is based on established notions of discourse structure: hierarchical organization, as in the stack-based model and RST's tree structures, and dominance or nuclear/satellite relations between discourse segments. As such, VT captures and formalizes intuitions about discourse structure that run through the current literature. VT also explicitly recognizes the special status of the left satellite for discourse structure, which has not been adequately addressed in previous work.

In this paper we have shown how VT addresses the left satellite problem, and how VT can be used to address various issues of structural ambiguity. VT predicts that references not resolved in the DRA of the unit in which it appears are more difficult to process, both computationally and cognitively; by looking at cases where VT fails we determine that this claim is justified. By comparing the types of referring expressions for which VT and the stack-based model fail, we also show that VT provides a better model for determining DRAs.

Acknowledgements

We thank Daniel Marcu for providing us with the RST annotated MUC corpus, and Valentin Tablan for developing part of the software that enabled us to process the data.

References

- Azzam S., K. Humphreys and R. Gaizauskas (1998). Evaluating a Focus-Based Approach to Anaphora Resolution. *Proceedings of COLING-ACL'98*, 74-78.
- Cristea, D., Ide, N., and Romary, L. (1998). Veins Theory: A Model of Global Discourse Cohesion and Coherence. In *Proceedings of the 36th Annual Meeting of the ACL and of the 17th COLING (COLING/ACL'98)*, 281-285.
- Cristea, D., Ide, N., Marcu, D., and Tablan, V. (2000). An Empirical Investigation of the Relation Between Discourse Structure and Co-Reference, in *Proceedings of Coling 2000*, 208-214.
- Fox, B. (1987). *Discourse Structure and Anaphora. Written and Conversational English*. No 48 in Cambridge Studies in Linguistics, Cambridge University Press.
- Grosz, B. and Sidner, C. (1986). Attention, Intention and the Structure of Discourse. *Computational Linguistics*, 12, 175-204.
- Gundel, J., N. Hedberg, and R. Zacharski (1993). Cognitive Status and the Form of Referring Expressions. *Language*, 69:274-307.
- Hahn U. and M. Strübe (1997). Centering in-the-large: Computing referential discourse segments. *Proceedings of ACL-EACL'97*, 104-111.
- Hirschman L. and N. Chinchor (1997). MUC-7 Co-reference Task Definition.
- Mann, W.C., Thompson S.A. (1988). Rhetorical structure theory: A theory of text organization, *Text*, 8:3, 243-281.
- Marcu D., E. Amorrtortu, and M. Romera (1999). Experiments in Constructing a Corpus of Discourse Trees. *Proceedings of the ACL'99 Workshop on Standards and Tools for Discourse Tagging*.
- Marcu, D. (2000). Extending a Formal and Computational Model of Rhetorical Structure Theory with Intentional Structures à la Grosz and Sidner. *Proceedings of the 18th International Conference on Computational Linguistics*, 523-29.
- Marcu, D. (1999). A Formal and Computational Synthesis of Grosz and Sidner's and Mann and Thompson's theories. *Workshop on Levels of Representation in Discourse*, 101-108.
- Marcu, D. (1996). Building Up Rhetorical Structure Trees. *The Proceedings of the Thirteenth National Conference on Artificial Intelligence*, vol. 2, 1069-1074.
- Mitkov, R. (1997). Factors in anaphora resolution: they are not the only things that matter. A case study based on two different approaches. *Proceedings of the ACL/EACL'97 Workshop on Operational Factors in Practical, Robust Anaphora Resolution*, 14-21.
- Moser, M. and J. Moore. (1996). Towards a Synthesis of Two Accounts of Discourse Structure. *Computational Linguistics*, 18(4): 537-544.
- Sidner, C. (1981). Focusing on the Interpretation of Pronouns. *Computational Linguistics*, 7:217-231.