Motivations And Implications Of Veins Theory

Dan Cristea

Faculty of Computer Science of the "Alexandru Ioan Cuza" University of Iasi, Romania
Institute of Theoretical Computer Science, Iasi branch of the Romanian Academy

Abstract. The paper deals with the cohesion part of a model of global discourse interpretation, usually known as Veins Theory (VT). By taking from the Rhetorical Structure Theory the notions of nuclearity and relations, but ignoring the relations' names, VT computes from rhetorical structures strings of discourse units, called *veins*, from which domains of accessibility can be determined for each discourse unit. VT's constructs best fit with an incremental view on discourse processing. Linguistics observations that lead to the elaboration of the theory are presented. Cognitive aspects like short-term memory and on-line summarization are explained in terms of VT's constructs. Complementary remarks are made over anaphora and its resolution in relation with the interpretation of discourse.

1 Introduction

A discourse is different from a text, because a discourse is a text in the progress of reading or hearing in a human brain. So, a discourse exists only as a process and, as such, it has a dynamic nature. When the reading comes to an end, the discourse also finishes and only a representation of it remains in the reader's memory.

The main concerns of the studies dedicated to discourse have been on proposing a representation that best describes its structure and on understanding the relationship existing between structure and referentiality. In Atentional State Theory (AST) [12] the discourse is seen as having a recursive segmental structure residing in a tree-like representation, while the dynamic interpretation uses a stack model in which the references are allowed to occur from the top state elements towards the bottom. The Rhetorical Structure Theory (RST) [17] gives only a static representation while ignoring any concern on referentiality. Centering Theory (CT) [1], [11] uses the notion of segment from AST to propose a local theory of discourse coherence.

We will review in this paper the cohesion part of a model of global discourse interpretation, usually known as Veins Theory (VT), while also noticing some new facts about it. By taking from RST its notions of nuclearity and relations, but ignoring the relations' names, VT [7] reveals a "hidden" structure in the discourse tree, called *vein*, which enables to determine a *domain of evocative accessibility* (*dea*) for each discourse unit, as that space of the discourse where all anaphors belonging to that unit can find an antecedent. As such, the vein expression of a discourse unit gives the minimal span necessary to understand that particular unit in the context of the whole discourse. VT allows for an integrated explanation of the common points of AST,

Cristea D. (2005).
Motivations And Implications Of Veins Theory.
In *Proceedings of the 2nd International Workshop on Natural Language Understanding and Cognitive Science*, pages 32-44
DOI: 10.5220/0002565000320044

RST and CT, while also correcting some AST predictions relative to accessibility domains (the nucleus to nucleus references and references from nuclei to left satellites).

In the following section we present linguistic observations that lead to the formulation of VT. The basic definitions are revised in section 3. VT's claim on discourse cohesion is presented in section 4. The last section gives a synthesis of the theory, exhibits a cognitive argumentation, quickly reviews applications based on the findings of the theory and shows some possible future developments.

2 The Intuitions Underlying VT

The notion of vein was born by synthesizing observations on how references align within the representation of a discourse as a tree. Considering the hierarchical organization given by the tree structure and the *principle of compositionality* (conforming to which, a relation that holds between two spans also holds between the most salient units of those spans [19]), which allow long-distance sibling relations between discourse units, these observations are collected below (to simplify the wording, we will say that "a unit *A* refers to a unit *B*" when we mean "a *referential expression* (*re*) belonging to the unit *A* refers to a *discourse element* (*de*) introduced or referred in/from unit *B*"). In the examples of this section we will mark with numbered *u* – the units, and with *R* – the relations. An upper *n* or *s* at the shoulder of an expression indicates that the corresponding text span is a nucleus, respectively a satellite. The names of relations in our commentaries of the examples are taken from RST.

a). Right satellites or nuclei can refer to their left nuclear siblings: in combinations $u_1^n\ R\ u_2^s$, or $u_1^n\ R\ u_2^n$, u_2 can refer to u_1;

Ex. 1:

1. John left home without an umbrella
2. although he watched the TV morning forecast announcing rain.

The pronoun *he* in unit 2, a satellite of unit 1, refers to the entity [**John**], introduced by the referential expression *John* in the first unit.

b). A right nucleus can refer to a left satellite: in combinations $u_1^s\ R\ u_2^n$, u_2 can refer to u_1 as in:

Ex. 2:

1. Although John watched the TV morning forecast announcing rain,
2. he left home without an umbrella.

where *he* in 2, a nucleus, refers to [**John**] introduced in 1, a left satellite of it.

c). A right satellite of a nucleus *u* is not accessible from another, more distant, right sibling of *u*, nuclear or satellite: in combinations $(u_1^n\ R_1\ u_2^s)^n\ R_2\ u_3^n$ or $(u_1^n\ R_1\ u_2^s)^n\ R_2\ u_3^s$, u_3 can refer to u_1 but not to u_2.

Ex. 3:

1. John told Mary that he loves her.
2. He was never married
3. and lived until 40 with his mother.
4. She, on the contrary, was married twice.

The sequence 2-3-4 ELABORATES on 1. The sequence 2-3 is in a relation of CONTRAST (a paratactic relation) towards 4, while unit 3 ELABORATES on 2. The structure is therefore: $u_1^n\ R_1\ ((u_2^n\ R_2\ u_3^s)^n\ R_3\ u_4^n)^s$. For most readers, *she* in unit 4 must be [**Mary**], and not [**John's mother**], although [**John's mother**] is the most recent entity from the position of unit 4 in agreement in gender and number with the pronoun *she*. The reason why the reader prefers Mary instead of the mother is because s/he recognizes the unit 4 as being in a CONTRAST relation with unit 2 (evidenced by *on the contrary*), which makes the two units to be perceived as adjacent, and having the same status with respect to a common nucleus, unit 1. Their proximity however is not linear but hierarchical, on the structure. This makes unit 3 to be closed for reference from unit 4, and the pronoun *she* in 4 to find its antecedent in the common upper nucleus – unit 1.

d). A nucleus blocks the reference from a right to a left satellite: in combinations $(u_1^s\ R_1\ u_2^n)^n\ R_2\ u_3^s$, u_3 can refer to u_2 but not to u_1.

Ex. 4:

1. With one year before finishing his mandate as president of the company,
2. Mr. W. Ross has begun to bring about its bankruptcy.
**3. There were rumors that he has obtained it by fraud.*

In this example the reader is confused on who the referent of the pronoun *he* in unit 3 could actually be. 1 and 3 are both satellites of unit 2: 1 is in a CIRCUMSTANCE relation towards 2, while 3 is intended to give a BACKGROUND for 2, if *it* would be perceived as referring [**the mandate as president of Mr. Ross**]. However this coreferential link is found with difficulty, which lowers the understandability of the whole discourse. It can be repaired in two ways:

Ex. 5:

1. Mr. W. Ross has begun to bring about the bankruptcy of his company.
2. with one year before finishing his mandate as president.
3. There were rumors that he has obtained it by fraud.

In Ex. 5, unit 2, expressing the positioning in time of the action expressed by the main clause, unit 1, is a satellite of 1, and unit 3, reproducing a gossip occasioned by an element introduced in unit 2, is a satellite of 2. The reference *it*=[**Mr. Ross' mandate as president**] can be recuperated without difficulty. The motivation for the failing of Ex. 4 compared to the acceptance of Ex. 5 stays not in the linearly longer distance between the anaphor and antecedent in Ex. 4 than in Ex. 5, but in the fact that a nuclear unit is interposed between the unit of the anaphor and the unit of the antecedent in Ex. 4, contrary to Ex. 5 where this situation does not occur.

If the reference is eliminated, then the discourse is also repaired:

Ex. 6:

1. With one year before finishing his mandate as president of the company
2. Mr. W. Ross has begun to bring about its bankruptcy.
3. There were rumors that he has been elected by fraud.

3 VT's Basics

The fundamental intuition underlying the unified account on discourse structure and accessibility in VT is that the RST-specific distinction between nuclei and satellites constrains the range of referents to which anaphors can be resolved. In other words, the nucleus-satellite distinction, superimposed over a tree-like structure of discourse, induces for each anaphor a *dea*. More precisely, for each anaphor *x* in a discourse unit *u*, VT hypothesizes that *x* can be resolved by examining discourse entities from a subset of the discourse units that precede *u*. If the *x*'s antecedent belongs to a unit that resides beyond the *dea* of *u*, then the link anaphor-antecedent is found with difficulty or, in order to realize it, strong referential means should be surfaced (as for instance proper names).

The discourse structure assumptions in VT are, to a great extent, the same as in RST: a) the basic units of a discourse are non-overlapping spans of text, usually a clause of a sentence (expressing an event, or a situation); b) discourse structures are represented as trees. Unlike RST, in VT, without any loss of generality, the trees are considered binary; a similar representation is used by Marcu [19]; c) terminal nodes of the tree represent *elementary discourse units* (*edu*s) and non-terminal nodes represent *discourse relations.* Unlike RST, VT is not concerned with the type of relations among textual spans, but considers only the topological structure of the discourse; d) a polarity, established among the daughters of a relation, identifies at least one node as *nucleus,* considered essential for the writer's purpose; non-nuclear nodes, which include spans of text that increase understanding but are not essential to the writer's purpose, are called *satellites*.

Vein expressions defined over a discourse tree are sub-sequences of the sequence of units making up the discourse. To define vein expressions, the following notations are used:

- each terminal (leaf) node (discourse unit) has an attached label;
- *mark(*α*)* is a function that takes a string of symbols α and returns each symbol in α marked in some way (e.g., within brackets);
- *unmark(*α*)* is the reverse function of *mark()*. It removes all markings attached to symbols in the expression α. (e.g. *unmark(*α *.* *mark(*β*)* *.* γ*)* = α *.* β *.* γ);
- *simpl(x)* is a function that eliminates all marked symbols from its argument, if they exist, e.g. *simpl*(*mark*(α)) = $\varnothing$, the empty string, and *simpl*($\alpha \cdot$ *mark*(β) $\cdot$ γ)) = $\alpha \cdot \gamma$;
- *seq(*α*,*β*)* is a sequencing function that takes as input two non-intersecting strings of terminal node labels, α and β, and returns that permutation of α concatenated with β that is given by the left-to-right reading of the sequence of labels in α and β on the terminal frontier of the tree. The function maintains the markings, if they exist and *seq*($\varnothing$, α) = α; *seq*(α, *seq*(β)) = *seq*(*seq*(α), β) = *seq*(α, β);

- *H(n)* and *V(n)* are the notations for the head and vein expressions of a node *n*;
- *pref(u, α)* retains the prefix of the expression *α* up to and including the symbol *u*.

VT computes two expressions that are attached to all nodes of a discourse structure. The notion of *head* in VT is equivalent to that of Marcu's *promotion set* [19]. The intention in the head expression of a node of a discourse tree is to capture the sequence of the most important units in the span of text covered by the node. It is a sequence of unit labels as follows:

1. The head of a terminal node is its label.

2. The head of a non-terminal node is the concatenation of the heads of its nuclear daughters.

Note that the recursive definition of head induces a bottom-up computation over the tree structure.

The *vein* expression of a node is intended to give the sequence of *edu*s which are significant for summarizing, in the context of the whole text, the span of text covered by the node. In the vein expression of any node in the discourse structure, are included *edu*s belonging to the span covered by the node, possibly together with *edu*s outside the span. By synthesis, or summary, of a text span we understand a shorter text, which still renders the original idea of the text. Irrespective whether it is realized by paraphrasing or by concatenating sub-sequences of the original text [16], any summary should be comprehensible by itself (among other things, this means that it should contain all elements that allow the resolution of anaphors). When the span to be summarized is extracted from a larger span, in order for the summary to be comprehensible, it should contain also elements from outside the span, which belong therefore to the context. We have, in this case, the summary of a text span, *in the context* of a larger span. Let's note also that, in many respects, "summarizing" is equivalent to "understanding" because what we are usually left after the reading of a text is a synthesis of it.

In the following, the whole text is called *total context*. In Fig. 1, the nodes to which the definition currently applies are depicted in grey. They are simultaneously drawn with a rectangle and a circle in order to suggest that they can be either inner nodes or terminal nodes.

Once each node of the tree is marked for the head expression, vein expressions are computed top-down for each node in the tree:

1. The vein expression of the root is its head expression.

The vein expression of the root node, conforming to the intention associated to the vein expression of a node, should be made of the most significant *edu*s that are necessary to understand/summarize the span covered by the node (in this case – the whole text), in the total context. But, since the covered text span in this case is the whole text, this gives us the definition of the head expression of the root node.

2. For each nuclear node whose parent node has a vein *v*:

a) if the node does not have a left non-nuclear sibling, then its vein expression is *v* (see Fig. 1a);

b) otherwise, if the left non-nuclear sibling has the head *h*, then the vein expression of the nuclear node is *seq(mark(h), v)* (see Fig. 1b)**.**

The definitions say that in order to understand/summarize, in the total context, a nuclear span, a right satellite sibling can be ignored, while a left satellite is significant. When positioned at the right of a nuclear unit, a satellite can be ignored, since the same units are necessary to understand/summarize, in the total context, the nuclear span plus the satellite span, or only the nuclear span. When positioned at the left, a satellite helps to understand/summarize its right nucleus, but should be ignored for any other right satellite of this nucleus (case commented in Ex. 4). The marking function *mark* signals the contribution of this left satellite, in order that a subsequent removal is operated in the vein expression of a right satellite (see 3b below). On the contrary, twin nuclei cannot be understood/summarized one without the other, meaning that the same units are significant to understand/summarize each one of them as their union span.

3. For each non-nuclear node of head *h* whose parent node has a vein *v*:

a) if the node is the left daughter of its parent, then its vein expression is *seq(h,v)* (see Fig. 1c);

b) otherwise, the vein expression is *seq(h, simpl(v))* (see Fig. 1d).

The definitions express the fact that in the understanding/resuming, in the total context, of a satellite span, one should add to the units that contribute to the understanding/resuming of its parent node the most important units within the satellite span itself (given by the sequence of units in its own head expression). Let's note that the vein expression of the parent node of this satellite, with one exception, inherits only head expressions of nuclear nodes from its own ancestors, therefore the significant units belonging to the satellite own span cannot be there and must be included explicitly. The exception mentioned refers to exactly the case when a satellite is placed on the left side of the nucleus towards which this node is itself a satellite, and whose units have been recorded by markings. The *simpl* function will delete this influence (see an example in Fig. 2).

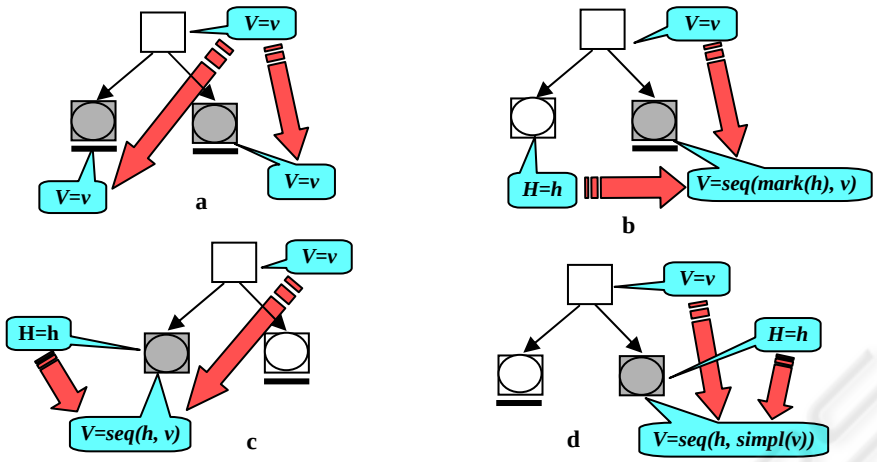

Fig. 1. Computing vein expressions. The node to which the computation applies is depicted in dark; nuclei are underlined.

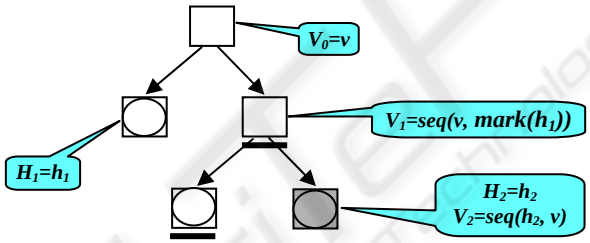

Fig. 2. Simplifications in the computation of the vein expression of a right satellite: V_2=seq(h_2, simpl(V_1))=seq(h_2, simpl(seq(v, mark(h_1)))) = seq(h_2, seq(v)) = seq(h_2, v).

4 The Relationship Between Discourse Structure And Referentiality

If we particularize the intuition behind the vein expression to a terminal node, we obtain: the vein expression of a terminal node *u* gives the sequence of *edu*s that are significant for understanding/summarizing *u* in the total context. Among other things, which we will not discuss in this paper, this means that, within the material indicated by the vein expression of an *edu*, antecedents of all anaphors belonging to that *edu* must be found. More precisely, seen: – the semantic nature of the anaphoric relation [13], – a representation of anaphoric relations in which *re*s of a textual layer are linked to representations of *de*s on a semantic layer, as the one proposed by Cristea and Dima [4] – and the common cognitive nature of anaphora and cataphora (as discussed in section 2), which allows for a unique directionality in the search for

antecedents, always towards the beginning of the text, we are lead to the definition of a domain of evocative referential accessibility (on short *domain of evocative accessibility – dea*):

$$dea(u) = pref(u, unmark(V(u)).$$

The definition of *dea* formalizes the first conjecture of VT (or the *cohesion conjecture*), which defines for any discourse unit a specific domain of accessibility computed in relation with the discourse structure: antecedents of the *re*s belonging to an *edu u* are mostly found among the *de*s anchored in the *edu*s which precede *u* in its vein expression, including *u* itself.

The first conjecture hypothesizes two types of anaphoric processes: *evocative* (or *imediate*) and *post-evocative* (or *inferential*). The evocative processes appear most frequently, are resolved quickly and can be realized at the surface by any referential material, including the most fragile, as empty subjects and pronouns. They give fluency to the text and make it cohesive. The post-evocative processes are less frequent, need a greater inferential load for their resolution and make use of strong referential material (as proper nouns).

If we transfer this classification to the anaphoric references involved in these processes, we will have evocative and post-evocative references (see Fig. 3). In the evocative references, the backward-looking chain of units anchoring *re*s that are referentially related intersects the *dea* of the anaphor's unit in at least one more unit than the anaphor's unit itself. In post-evocative references this double intersection is missing. In [3] and [6] the evocative references are further detailed in direct and indirect. In *direct references* the second intersecting unit (looking backward from the anaphor's unit) is the linearly most recent one, counting from the anaphor's unit, anchoring the same *de* as the one referred by the anaphor (in case of coreference), or a *de* that is anaphorically related to the anaphor's *de* (in case of functional reference). In *indirect references* the two backward looking chains intersect in a unit that is not linearly most recent from the anaphor's unit.

Sometimes an anaphor belonging to the post-evocative class can be understood without even having to make a connection to an antecedent. These are usually called pragmatic references or pseudo-references. The interpretation of *re*s in this class can be made based on knowledge that comes from outside the test, from common knowledge. Although the text contains at least one more *re* that realizes the same *de* as the anaphor, the coreferential expressions may not be represented identically in order for the text to be understood.

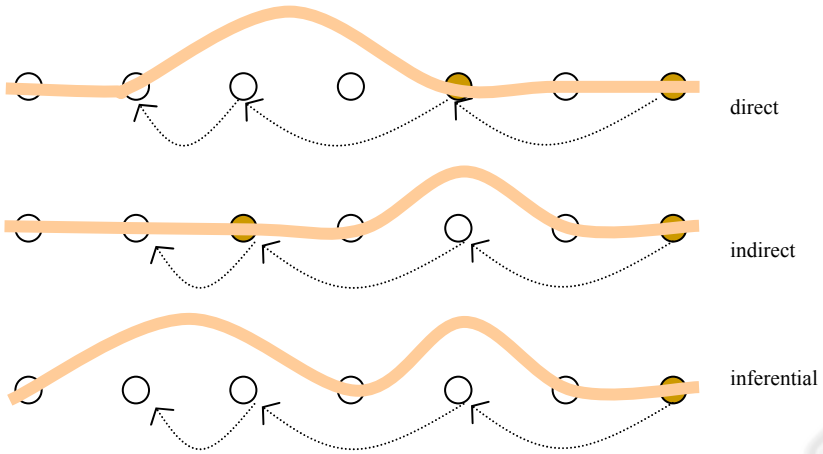

Fig. 3. Evocative and post-evocative references. Anaphoric chains are depicted by dotted-lines and *dea* chains by thick lines. The anaphor's unit is the last one to the right.

5 Discussions

The fundamental assumption underlying VT is that an inter-unit reference is possible *only if the two units are in a structural relation with one another,* even if they are distant from one another in the text stream. Furthermore, inter-unit references are *rather to nuclei than to satellites,* reflecting the intuition that nuclei assert the writer's main ideas and provide the main "threads" of the discourse [17]. This is shown in the computation of veins over (binary) left polarized discourse trees, where any reference from a nuclear unit must be to entities contained in linguistic expressions from the previous nuclei (although perhaps not *any* nucleus). On the other hand, satellites depend on their nuclei for their meaning and hence may refer to entities introduced within them.

Given the mapping of Grosz and Sidner's [12] stack-based model of discourse structure on RST structure trees outlined by Moser and Moore [21] and Marcu [18], the domains of referentiality defined for left-polarized trees using VT are consistent with those defined using the stack-based model. However, in cases where the discourse structure is not left-polarized, VT provides a more natural account of referential accessibility than the stack-based model. In non left-polarized trees, 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 the nuclear unit *B*. In stack-based approaches to referentiality, such configurations raise problems: as *B* dominates *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 "right" references of cataphora included in satellites that precede their nuclei, which is counter-intuitive.

Inferential references, as defined by VT, seem to minimize the importance of the domain of referential accessibility, because references can now "escape" from the domain. Does the domain of accessibility have any significance anymore? Is it an artificial invention or is it defended by a natural characteristic of the manner people process texts? We claim that there are two significantly distinct types of anaphora resolution processes: *evocative* (or *associative*) and *post-evocative* (or *inferential*).

The evocative resolution processes are based on associations, which are processes of pattern-matching on feature structures decorated with morpho-semantic attributes. They are performed between a feature structure projected by the anaphor *re* and a *de* that already exists in the *dea* of the unit the anaphor belongs to [4]. These are fast processes, direct ones being faster and more frequent than indirect ones. 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. Because co-referential expressions are organized in equivalence classes, it is sufficient if an anaphor is resolved to some member of the set. This is consistent with the distinction between *direct* and *indirect* references.

On the other hand, the post-evocative processes are inferential processes that are developed in memory, based on the knowledge accumulated by the preceding discourse, or based on the cultural knowledge the subject owns. We believe these inferences swing the semantic space in an order that is also dictated by the discourse structure. Eventually, the target entity can be found based on a pattern-matching process between the projected structure of the anaphor and the semantic representation of the antecedent. They are slow – computationally and cognitively (compel to more inference load), require more powerful referencing means (like proper nouns), and are less frequent.

An aspect not described in this paper is VT's account on discourse coherence [7]. Starting from *dea*s, the notion of segment in a hierarchical sense is introduced, which generalizes the classical notion of segment as employed in AST [12] and CT [11]. By this, VT generalizes CT from a local theory of coherence to a global one.

Empirical evidences on the VT's claims on cohesion and coherence have been reported in [6], [7] and [14] with experiments developed on corpora annotated to discourse structure and coreferentiality in English, French and Romanian. In particular, these studies reveal the following: in most cases the references are direct; in less cases the references are indirect; in very few cases the references are pragmatic; inferential references which are not pragmatic signal a hard-to-make inference or a failed discourse. Moreover, it can be proved that VT's assumptions regarding the cohesion are stable to the change of granularity (the limit below which material *edu*s are considered) from lower to upper.

A side effect of corpus research motivated by the evaluation of VT claims was the notice that there is a strong relationship between the different kinds of referential expressions and their distribution with respect to the three kinds of references put in evidence by VT. It was revealed an alignment between the evoking power and the percentage of different types of referential expressions that did not corresponded to a vein reference (inferential). Four types of inferential references have been discovered: pragmatic, proper nouns, common nouns and pronouns, which revealed to have descending frequencies, in this order. Pragmatic and proper nouns references are easily resolved, which makes their use much less restricted by the placement of an antecedent on a current *dea*. At the other pole, pronouns are very fragile evoking

means, and, as such, a message emitter employs them when s/he is certain that the current structure of the discourse allows for easy recuperation of the antecedent on the *dea* of the anaphor. The alignment of the evoking power of referential expressions with the percentage of exceptions of references outside the *dea*s shows that the predictions made by VT in the cohesion conjecture are correct. Practically, except for the cases when the pronoun can be understood without an antecedent, it becomes impossible to use a pronoun as an anaphor to refer an antecedent that is outside the *dea*.

Scholars dealing with the interpretation of discourse and reading in connection with the cognitive science [2], [15], [23], [26] generally, agree on three types of memory: immediate memory (IM), short term memory (STM) and long term memory (LTM). IM is a sensorial storage of information, which allows the retaining of traces from the last half second. STM keeps information for few seconds. According to Miller [20], the length of this memory seems to be of 7±2 signs (words, figures, letters – depending on the context), while others estimate this ”buffer” to an average of 13-15 words [22]. In [9] and [10] an incremental discourse parsing model is described in which the developing structure is updated with a new auxiliary tree after the reading of each sentence. The discourse tree becomes bigger and bigger as the text unfolds.

In the human memory, as well as in automatic discourse parsing systems, summarization processes must evolve in parallel with the building of the discourse structure. We believe that the STM should be linked to the *dea* of the last *edu* processed: either the last 7±2 *edu*s in this sequence, or the same number of event structures – as representations of *edu*s, or only words picked up from this buffer. When we replace the current unit u_n with the next unit u_{n+1}, actually we replace the STM $dea(u_n)$ with the STM $dea(u_{n+1})$. Sometimes this means a simple prolongation of the preceding *dea*, other times it means the deletion of the most distant in time unit and the inclusion of a new unit – the current *edu*. STM is therefore made of a chain of *edu*s (or of microstructures corresponding to *edu*s), which is projected from the dynamic evolving discourse structure. The alterations affecting the STM string reflect the updates of the sub-discourse in focus, while reading. When the interest has moved along another direction, the content of the current vein and, consequently, of the current *dea*, is updated too. The inclusion and deletion from STM of certain mini-structures, therefore these “recall” and “oblivion” processes, resemble the calling in attention of Walker’s [26] cash memory model. The recall processes are possible from the discourse structure that is kept in a summarized form in the LTM. Evocative anaphoric processes are thus developing in the STM, while post-evocative processes are of an inferential type, and necessitate greater inference load to recover *de*s from memory or evoke entities kept in the generic cultural sphere of the individual. We belief these processes evolve also on the developing discourse structure, but leaving the *dea* when the resolution failed there.

There are as many ways to read a text as there are *edu*s in it. These different readings are given by the *edu*s’ vein expressions. Each vein represents a summary of the text focused on the respective unit. When the reader is focused on a certain episode or entity mentioned by the text s/he can skip entire fragments and look for the manner in which the element of interest integrates in the whole discourse. Summaries focused on different events or entities can contain elements in common, while each of them has also specific elements, although strongly correlated to the main line of the

discourse. All these sub-discourses are coherent and, generally, there are no anaphoric references whose interpretation to necessitate elements outside the summary itself.

We believe that the processes of anaphora resolution and discourse structure building are interdependent to such a degree that discourse analysis should make use of them in tandem, and combine their partial results to acquire the best discourse tree. In the same way that anaphora resolution can benefit from the discourse structure, already solved anaphora can be used in determining the best structure, which in turn contributes to the resolution of further anaphora. The constraints evidenced act as forces that, in a well-understood discourse, give rise to a sort of state of equilibrium, resembling the minimum potential energy of a physical system. Humans have an innate cognitive mechanism that allows them to obtain naturally the most plausible interpretation of a text. When arrived there, they are invigorated by the reach of a "comfortable" mental state, which should be based on the maximal satisfaction of a constraints system. In [9], a model and an implementation that mimics this behavior are described. Scores contributed by the cohesion conjecture are combined with scores contributed by the coherence conjecture of VT (hierarchical generalization of CT) in order to obtain the most "fluid" possible discourse structure (maximum of cohesion and of coherence).

VT's account on the relationship between discourse structure and referentiality can be exploited in three ways:

- to constrain a simultaneous parsing and anaphora resolution process towards that interpretation that requires minimum inferential load in building the structure and in identifying the antecedents of referential expressions [3], [5], [8], [9];
- to correct discourse structure when referential links are known [25];
- to guide a process aimed at producing focused summaries [9], [10].

The notice that slightly modified texts can display the same vein structure (although not the same tree structure) can lead to the idea that veins could be seen as a kind of sub-specification representation [24], a direction which has not been investigated yet. Also, as trees annotated at discourse structure and veins can lead to almost instantaneous computation of focused summaries on any discourse entity or event mentioned in the text, it would be worth investigating an RDF representation of vein structures obtained by processes of automatic parsing, with interesting applications in Semantic Web.

References

1. Brennan,S.E.; Walker Friedman,M. and Pollard, C.J.: A centering approach to pronouns. *Proc. of the 25th Annual Meeting of ACL*, Stanford (1987) 155-162
2. Cornea, P.: Introduction in the Theory of Reading (in Romanian), Polirom Publishing House, Iaşi (1988)
3. Cristea, D.: An Incremental Discourse Parser Architecture. D. Christodoulakis (Ed.) *Proceedings of the Second International Conference - Natural Language Processing - NLP 2000*, Patras, Greece Lecture Notes in Artificial Intelligence 1835, Springer (2000)
4. Cristea, D. and Dima, G.E.: An Integrating Framework for Anaphora Resolution. *Information Science and Technology*, Romanian Academy Publishing House, Bucharest, vol. 4, no. 3 (2001)

5. Cristea, D., Dima, D.E., Postolache, O.D., Mitkov, R.: Handling complex anaphora resolution cases. *Proceedings of the Discourse Anaphora and Anaphor Resolution Colloquium*, Lisbon, Portugal (2002)
6. Cristea, D., Ide, N., Marcu, D., and Tablan, M.V.: Discourse Structure and Co-Reference: An Empirical Study. *Proceedings of the 18th International Conference on Computational Linguistics COLING'2000*, Saarbrueken (2000)
7. Cristea, D., Ide, N., and Romary, L.: Veins Theory: A Model of Global Discourse Cohesion and Coherence. *Proceedings of the 17th Coling and the 36th Annual Meeting of the ACL* (COLING-ACL'98). Montreal, Canada, (1998) 281–85
8. Cristea, D. Postolache, O.D., Dima, D.E., Barbu C.: AR-Engine – a framework for unrestricted co-reference resolution. *Proceedings of the LREC'2002*, Las Palmas, Spain (2002)
9. Cristea, D., Postolache, O., Pistol, I.: Summarisation through Discourse Structure. In *Proceedings of CiCling 2005*, Springer LNSC, vol. 3406 (2005)
10. Cristea, D., Postolache, O., Puşcaşu, G., Ghetu, L.: Local and global information exploited in producing summaries. *Proceedings of the International Symposium on Reference Resolution and Its Aplications to Question Answering and Summarization*, Venice, Italy, June (2003)
11. Grosz, B.J.; Joshi, A.K. and Weinstein, S.: Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 12(2), (1995) 203-225
12. Grosz, B.J., and Sidner, C.: Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3), (1986) 175–204.
13. Halliday, M.A.K. and Hassan, R.: Cohesion in English, Longman, London and New York (1976)
14. Ide, N., and Cristea, D.: A Hierarchical Account of Referential Accessibility. *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics*, ACL'2000, Hong Kong (2000)
15. Kintsch, W. and Van Dijk, T.A.: Comment on se rappelled et on résume des histories, *Langages*, 40 (1975)
16. Mani, I.: Automatic Summarization, Amsterdam, John Benjamins (2001)
17. Mann, W.C., and Thompson, S.A.: Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text* 8(3), (1988) 243–281
18. Marcu, D.: A formal and computational synthesis of Grosz and Sidner's and Mann and Thompson's theories. *Proceedings of the Workshop on Levels of Representation in Discourse*, Edinburgh (1999)
19. Marcu, D.: The theory and practice of discourse parsing and summarization, The MIT Press, Cambridge, Massachusetts (2000)
20. Miller, G.: The magical number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information, *The Psychological Review*, vol. 63, (1956) 81-97
21. Moser, M., and Moore, J.D.: Toward a synthesis of two accounts of discourse structure. *Computational Linguistics*, 22(3), (1996) 409–419
22. Richadeau, F.: La lisibilité. Langage-Typographie-Signes-Lecture, Paris (1969)
23. Schank, R. and Abelson, R.: Scripts, plans, goals and understanding, Hillsdale, N.J (1977)
24. Schilder, F. Robust Discourse Parsing Via Discourse Markers, Topicality and Position. *Natural Language Engineering* 1, (1), (2001) 1-22
25. Sereţan, V. and Cristea, D.: The Use of Referential Constraints in Structuring Discourse. *Proceedings of the LREC'2002*, Las Palmas, Spain (2002)
26. Walker, M.A.: Limited attention and discourse structure. *Computational Linguistics*, (1996) 22-2