REFERENCE RESOLUTION BY CONTEXTUAL REASONING FROM GRAMMATICAL REPRESENTATIONS WITH GETA_RUN

Dario BIANCHI, Rodolfo DELMONTE, Emanuele PIANTA, Sandra SARTORI

Università di Venezia

1. The Foundations

The topic of this paper is the theoretical foundations and the results of a system for text analysis and understanding called GETA_RUN, developed at the University of Venice, Laboratory of Computational Linguistics, Department of Linguistics and Language Teaching Theory. The main tenet of the theory supporting the construction of the system is that it is possible to reduce access to domain world knowledge by means of *contextual reasoning*, i.e. reasoning triggered independently by contextual or linguistic features of the text.

It is sensible to assume that when understanding a text a human reader or listener does make use of his encyclopaedia parsimoniously. Contextual reasoning is the only way in which a system for Natural Language Understanding should tap external knowledge of the domain. In other words, a system should be allowed to perform an inference on the basis of domain world knowledge when needed and only then. In this way, the system could simulate the actual human behaviour in that the access to extralinguistic knowledge is triggered by contextual factors independently present in the text and detected by the system itself.

It is also our view that humans understand texts only whenever all the relevant information is supplied and available. Descriptive and narrative texts are usually self-explanatory - not so, literary texts - in order to allow even naive readers to grasp their meaning. Note that we are not here dealing with spoken dialogues, where a lot of what is meant can be left unsaid or must be implicitly understood.

In the best current systems for natural language understanding (see the Proceedings of the ANLP'92, and the tutorial on Fully Implemented Natural Language Understanding System), linguistic components are kept separate from knowledge representation, and work which could otherwise be done directly by linguistic analysis is duplicated by the inferential mechanism. Linguistic representation is usually mapped onto a logical representation which is in turn fed onto the knowledge representation of the domain in order to understand and validate a given utterance or query.

Thus the domain world model must be priorly built, usually in view of a given task the system is set out to perform. However, it is clear that this modelling is domain and task limited and no generality whatsoever is achieved from it. In some of these systems, the main issue is how to make the two realms interact as soon as possible in order to take advantage of the inferential mechanism to reduce ambiguities present in the text or to allow for reasoning on linguistic data, which otherwise couldn't be understandable.

University of Venice
Working Papers in Linguistics
1993 vol.3,1
We assume that an integration between linguistic information and knowledge of the world can be carried out at all levels of linguistic description and that contextual reasoning can be thus performed on the fly rather than sequentially. This does not imply that external knowledge of the world is useless and should not be provided at all: it simply means that access to this knowledge must be filtered out by the analysis of the linguistic content of surface linguistic forms and the abstract representations of the utterances making up the text.

As we said, the task we are faced with when trying to simulate human understanding of texts is to scientifically isolate the contexts in which external knowledge of the world should be made available to the system, as well as providing the tools to deal with this task adequately. There is a description of our task which deserves quoting, and is taken from P. Bosch contribution to a book by Herzog & Rollinger(eds.), **Text Understanding in LILOG**, which we take to be the best example of the attempt to come to terms with the problem at hand. In his paper, the author makes the point of what he takes to be the main problem to be tackled: i.e. identifying in a text "inferentially unstable" concepts which are to be kept distinct from "inferentially stable" ones. The latter should be analysed solely on the basis of linguistic description, while the former should tap external linguistic knowledge of the world. Before entering into a comment of this issue, we would like to quote from his Conclusions:

"The central point of this paper is to try to give a direction to work on the interaction of linguistic analysis and knowledge representation in knowledge-based NL Systems. I have tried to argue and to demonstrate that without a full linguistic analysis there is little hope that we shall ever have reasonably general and portable language modules in NL systems. It has also become clear, I hope, that this is not a trivial task but requires a decent amount of empirical research for many years to come. But the linguistic research required is not isolated research in pure linguistics, but close cooperation with work on knowledge representation and - although this is a point I have not argued for - psychological work on conceptual systems, is imperative.

The most difficult problem to overcome, I believe, is that the most generally held belief in the scientific community with respect to our problem is that the distinction between linguistic and conceptual facts is arbitrary and hence not a proper research question, but a matter of pragmatic decisions. It is this belief more than anything else that inhibits further progress of the kind Brachman found lacking."(p.257)

We regard our work as a contribution towards this final goal which we identify tout court with contextual reasoning, i.e. performing inferential processes on the basis of linguistic information while keeping under control the contribution of external knowledge in order to achieve understanding of a text.

2. The System GETA_RUN

GETA_RUN is a general multilingual text and reference understander which represents a linguistically based approach to text understanding and embodies a number of general strategies on how to implement linguistic principles in a running system. The system addresses two main issues: multilinguality, and the need to restrict access to
extralinguistic knowledge of the world by contextual reasoning, i.e. reasoning from linguistically available cues.

This idea implies that a text must be built in such a way as to enable the reader to understand it with the least conceptual effort. In particular, external knowledge sources should be tapped only in a parsimonious way and the interaction between linguistic processing and knowledge representation and processing should be activated by internal strictly linguistically motivated cues.

As to multilinguality, the user may switch from one language to another by simply unloading the current lexicon and loading the lexicon for the new language: at present only Italian and English are fully implemented. Work is under way for other Romance and Germanic languages, such as French and German.

2.1 Parser. The system is based on LFG theoretical framework and has a highly interconnected modular structure. Basic grammatical representation modules are the Lexicon, C-structure and F-structure which is internally represented as a graph. The Parser is a DCG which exploits the properties of Prolog as to its general parsing strategy - topperdown, depth-first - and makes backtracking naturally available.

The basic tenet of the system is embodied in the wellknown fact that all languages share a common core grammar and may vary at the periphery: internal differences are predicted by principles. The DCG grammar allows the specification of linguistic rules in a highly declarative mode: it works topperdown and by making a heavy use of linguistic knowledge may achieve an almost complete deterministic policy. Principles are scattered throughout the grammar so that they can be made operative as soon as a given rule is entered by the parser.

In particular, a rule may belong either to a set of languages, e.g. Romance or Germanic, or to a subset thereof, like English or Italian, thus becoming a peripheral rule. Rules are activated at startup and whenever a switch has been operated by the user, by means of logical flags appropriately inserted in the right hand side of the rule. No flags are required for rules belonging to the common core grammar.

Some such rules include the following ones: for languages like Italian and Spanish, a Subject NP may be an empty category, either a referential little pro or an expletive pronoun; Subject NPs may be freely inverted in postverbal position, i.e. preverbal NP is an empty category in these cases. For languages like Italian and French, PP or adverbial adjuncts may intervene between Verb and Object NP; adjectival modifiers may be taken to the right of their head Noun. For languages like English and German, tense and mood may be computed in CP internal position, when taking the auxiliary or the modal verb. English allows an empty Complementizer for finite complement and relative clauses, and negation requires do-support.

Syntactic and semantic information is accessed and used as soon as possible: in particular, both categorial and subcategorization information attached to predicates in the lexicon is extracted as soon as the main predicate is processed, be it adjective, noun or verb, and is used to subsequently restrict the number of possible structures to be built. Adjuncts are computed by semantic compatibility tests on the basis of selection restrictions of main predicates and adjuncts' heads.

Syntactic rules are built according to the latest chomskyan paradigm with CP-IP functional maximal projections; however, the general underlying theoretical framework is
cast into LFG theory. Thus, we build and process syntactic phenomena like wh-movement before building f-structure representations, where only anaphoric binding for pronominals takes place.

2.2 Quantifier Raising. Since we know that quantifiers and quantified NPs usually take scope at propositional level, we assume f-structure to be an adequate level of representation in which quantifiers scope can be computed. We partially follow in this Halvorsen’s proposals, which however require a further mapping from f-structures to σ-structures in order to do that. We proceed as follows: after assigning Q-Markers to quantifiers and quantified NPs and adding this information as attribute-value pair at f-structure, we perform Quantifier Raising by traversing f-structure until we reach a propositional node. At that level we deposit a Quantifier-Operator (Q-Op), in an attribute that has a list as its value. Once Q-Ops have been produced, we are in a position to assign quantifier scope. In case more than one Q-Op is present in the list, the algorithm simply reorders the operators according to their quantifying force, and or to grammatical function. Otherwise, a search downward is performed in the f-structure for other q-ops. When some q-marker is found another attribute-value pair is added at pred level indicating a quantified interpretation.

2.3 The Binding Module. The output of the grammatical modules is fed then onto the Binding Module (BM) which activates an algorithm for anaphoric binding in LFG terms using f-structures as domains and grammatical functions as entry points into the structure. Pronominals are internally decomposed into a feature matrix which is made visible to the Binding Algorithm (BA) and allows for the activation of different search strategies into f-structure domains. Antecedents for pronouns are ranked according to grammatical function, semantic role, inherent features and their position at f-structure. Special devices are required for empty pronouns contained in a subordinate clause which have an ambiguous context, i.e. there are two possible antecedents available in the main clause. Also split antecedents trigger special search strategies in order to evaluate the possible set of antecedents in the appropriate f-structural domain. Special care is paid to pronominals bound by quantifiers or quantified NPs. The output of the BA is then passed on to an Interpretation Module which operates locally in order to spot the presence of conditions for Specific or Arbitrary Reading for pronominal expressions.

Finally, this information is added into the original f-structure graph and then passed on to the Discourse Module (DM).

2.4 Anaphora Resolution. Anaphoric binding of free pronominals takes as input Discourse level information which is computed by a Module of Discourse Anaphora (MDA) and decides on the basis of semantic categories attached to predicates and arguments of predicates whether to bind a pronoun to the locally available antecedent or to the discourse level one.

Discourse Anaphora is computed by a system which is very close to the ones available in literature on the same topic, and presented by C.Sidner and B.Webber in their publications. Definite descriptions are computed by means of locally available information and, but only when required, by tapping external knowledge sources. However, a set of
default rules are activated in lack of such knowledge which work simply on the basis of grammatical and semantic information.

This level of representation works on the basis of a list of candidates or possible arguments of discourse which includes all external pronouns and referential expressions. The algorithm creates a Weighted List of Candidates Arguments of Discourse (WLCAD).

The MDA detects Main, Secondary, Potential and Expected TOPIC for each sentence by weighting the list of external pronominals and of referential expressions made available by the grammatical representation. After the first sentence, the MDA tries to take advantage of discourse level internal cohesion and coherence mechanisms in order to check whether a given topic is reintroduced in the following text as a pronoun, a nominal substitute, a deictic and so on. The MDA is a finite state automaton that works strictly on the basis of two adjacent sentences; it has available a set of six states Continue, Change, Resume a previous topic; other states are Shifting, Retaining, Continue_Analyze. Special inferential mechanisms are activated at this local level by the presence of nontematic functions, i.e. SUBJects or OBJects of copulative and other verbs which do not assign a semantic role to their arguments. The output of the MDA is a Rhetorical Structure with a list of all Topics in each utterance and a state.

2.5 Computing Information Structure. As a final level of representation the system builds Information Structure at clause level. Each utterance is decomposed into separate clauses to account for coordination and subordination. The output is a set of values associated to wellknown labels such as Clause Type, Point of View, Factivity, Change in State of the World, Relevance and Discourse Relations. Point of View may vary between Subjective, Subjective_Extensional, Subjective_Intensional, Objective; Change may be Null, Early, Culminated, Gradual and Setting; Relevance may be Foreground and Background. All these values are computed mainly on the basis of the semantic, aspectual and temporal features associated with the verbal predicate at f-structure.

Information structure is passed on to the Semantic Module and Discourse Relation Reasoning Module. In particular, the latter computes Discourse Relations on the basis of information structure, temporal interpretation and the model. The output of this module is a set of values, which include: Narration, Elaboration, Description, Parallel, Explanation, Cause, Result.

GETARUN is now able to activate a Supervisor which computes Relevance associated to each individual or set asserted in the DKBF by adding scores derived from the Information Structure and the Rhetorical Structure.

2.6 Semantic and Temporal Interpretation. Semantic interpretation is carried out in two phases: a static phase which takes as input the description realised by LF, turns it into a list of lists which contain relations, their arguments and their modifiers and adjuncts. In particular, a given predicate will be represented as a relation in the sense that it will constitute a fact or a situation, according to semantic conditions, and its arguments will consequently be assigned a similar representation. However, the dynamic interpreter alone will make available world identifiers for entities referred to by arguments of a given predicate and will link intensional objects to extensional entities existing in the world.

Spatial and Temporal Location of States and Events is also computed in order to provide the basic location indices on which reasoning in the following module will operate.
The result of the DM is thus fed on to the Dynamic Interpreter for the creation of a Dynamical Knowledge Base of Facts and Situations (DKBFS). Entities are represented as Classes, Individuals and Sets which are asserted into the DKBFS together with their cardinality and restrictions where present. Each fact or situation is numbered as infon and space and time anchoring locations are associated to it.

The Dynamic Interpreter is responsible for long distance textual coreferential mechanisms: local topics are checked for coreferentiality with previously asserted individuals or sets existing in the model world. They may be regarded to be so in case their cardinality and/or their restrictions coincide. Other possible results of the SM interpretation procedure are set inclusion: a set or individual is regarded to be included into a previously asserted set in case its cardinality and/or its restrictions and/or inherent semantic features allow it. Also, there can be cases of situations supporting facts in the sense of a given set being introduced in the world as an intensional object, supporting the existence of an extensional individual interpreted as a fact.

The SM is able to assign cardinality to plural NPs on the basis of the restriction defined by their possessor, in case it is a pronoun bound at sentence or text level. Recency of referential procedures accounts for preferred antecedent assignment in case of semantic compatibility to the closest individual/set available.

The main or central interpretive assumption purported by the system is then constituted by the ability to control inferential search to external knowledge of the world: this is done only in case a singular definite NP is introduced in discourse as the higher referential expression candidate for topichood, whenever no matching takes place with the previous topics.

2.7 The World or Discourse Model. Reasoning is carried out in the Discourse Model into several separate modules: Discourse Relations, Temporal Reasoning and Situation Semantics Representations. Discourse Relations are built for each proposition by means of semantic information associated to main predicates and by the output of the Temporal Reasoner. The latter, is a module embodying J.Allen's ideas and Reichenbach's tripartite structure of temporal representation. According to the input received, the system may assign an interpretation in terms of discourse relations by means of a mechanism of Temporal Anaphora and Focus which takes into account information related to semantic roles associated with arguments of the predicates and matches them at first with the ones available in the previous proposition: a temporal anaphora or a new temporal focus is thus the outcome both of semantic and temporal reasoning.

These information are eventually passed on the Semantic Module where a representation in terms of situation semantics is computed: each proposition is translated by the logical form algorithm into a fact or a situation according to semantic conditions, and is completed by a polarity, a spatio temporal location constituted by constants inherited from the previous modules. The final ontology is made up by locations, individual entities which may also be treated as classes and set with a given cardinality, facts and situations about entities which may be attributes or roles according to their semantics, basically constituted by the grammatical function and semantic role associated to the argument or adjunct by the previous parse.

The Knowledge Base thus coincides with the dynamically built Discourse Model where extensionality and intensionality is computed according to the actual value of the
main predicated in a given context. The External World Knowledge Base is tapped only when needed, and in particular whenever a singular definite NP is introduced with a special topicality in the discourse and no antecedent is be available in the Model.

The Discourse Model is consequently a database in which inheritance is easily computed at the end of the story. In addition, we created a mechanism of Relevance Score assignment which takes into account topicality and discourse relations in which a given entity was involved by computing a score with different weights. In this way, all entities appearing in the story are finally sorted in descending order according to their relevance, with the list of facts and situations each one shared.

Besides c-structure and f-structure, we use logical form to compute quantifier scope and build logical representations which are eventually passed on the Discourse Model, where semantic interpretation is carried out.

Temporal aspectual interpretation is carried out on the output of f-structure representation and computes a complete interval logic coindexation of events and states at clause level on the basis of Reichenbach's tripartite model for temporal calculus. This is then passed on to the semantic module to serve for the understanding of Temporal Relations, which are cast on the basis of J.Allen's system for temporal logic notation.

3. How the System works

In processing a narrative text, a listener is developing a model of at least two things: 1. the entities under discussion, along with their properties and relationships to one another, and 2. the events and situations under discussion, along with their relationships to one another (e.g. consequential relations, simple ordering relations, elaboration relations, etc.). This representation can be called the Discourse Model.

NPs may evoke entities into the listener's DM corresponding to individuals, classes, sets, subsets of previously asserted sets or classes, relations, roles etc. An NP which evokes a discourse entity also specifies it, where "specify" means refer in a model, as opposed to refer in the outside world (Sidner 1983). Evoking an entity may cause its description in the model in case no such entity already exists; or it may cause the description of some additional property not already present in the DM associated to that entity. In case none of these two possibilities obtain, we say that the evoked entity cospecifies a previously asserted entity. Two possibilities remain: the entity has been defined a Topic by the Discourse Component and the identifier associated to the entity will be assigned to that Topic. Valid Topics are Expected, Main, Secondary and the first Potential Topic on the stack. This applies to all singular definite NPs. As for indefinite NPs the stack of Potential Topics can be scanned and whenever an indefinite singular NPs is found it is asserted as a fact or sit according to semantic conditions. Definite NPs, be they plural or singular, placed lower in the stack of Potential Topics, will not be described in the model as sets, they only appear as facts related to classes. Discourse relevance is taken as a determining factor in splitting up relevant entities from irrelevant ones. In other words, a definite NP be it singular or plural, is filtered by the Discourse Module in order to be regarded as a relevant possible entity.

The process of anaphora resolution in our system is split up in two separate procedures: the one takes into account pronominal reference and the other nominal
reference by means of definite NPs. In fact, the first procedure is only activated whenever pronouns left free by the module of anaphoric binding at sentence level are spotted by the algorithm. The second procedure on the other hand is always activated and is responsible also of the setting up of the current discourse state.

### 3.1 States and Domains

#### 3.1.1 Coherence as a Finite State Machine

In their paper on the Centering Approach, Brennan et al. (1987) present a formalization to modeling attentional structure in discourse as a means for capturing coherence at text level. Their approach embodies a set of rules and constraints that should reflect the relationships existing between what the discourse is about and linguistic choices made by discourse participants as to their local relevance in a given discourse segment. They also establish a typology of transitions from one utterance to the next which should describe the way in which utterances are linked together in a coherent local segment of discourse. As the authors note, if a speaker has a number of propositions to express, one very simple way to do this coherently is to express all the propositions about a given entity before introducing a related entity and then perhaps shifting the center to this new entity.

Even though their paper is an excellent presentation of the problem at hand, and the algorithm seems perfectly wrought out, we assume - as the authors themselves admit - that it is only in a preliminary stage in the coverage of real texts. Also, it is clearly a non efficient tool, since the work the algorithm does every time a Center must be established is redundant. There seems to be a lack of confidence in the possibility that real texts should actually behave the way their approach suggests.

The algorithm is thus deficitary in two ways: it is too strict and too limited in scope. This might depend on the fact that the authors did not try it on an extended number of texts, perhaps because they believed that the same approach could be simply adopted as it stands to cover other more complex cases.

Our approach to the general problem of reference resolution is guided by the following economy principle:

"Try a matching on the basis of grammatically encoded information; then explore the Model and/or the Discourse Domain in search of a suitable entity, and try an inference."

The Finite State Machine we use at intersentential level, is not so strict as the one proposed by Brennan et al. In particular, we also build up a Weighted List of Topics (WLT) which however does rely on Grammatical Functions, Semantic Roles and Selectional Restrictions information associated with each referring expression of the current utterance. The first Ref_ex in the rank list is always assigned a preferential status, however we also keep an eye to the second ranked Ref_ex. The rest of the stack might also be used to search for the presence of some Ref_ex which corefers with one of the previous Topics: this is done in order to ascertain whether some cohesion is present.

The representation we assign to each state of discourse is then more articulated: we use four specific Topic labels, Expected, Main, Secondary, Potential. The latter is in fact a
stack, and the previous ones are individual slots. Whenever a Ref_ex becomes Main Topic it usually reappears as Secondary Topic in the following text even though it is not included in the list of Ref_ex for that utterance. This is done to establish the persistence of a Main Topic in the discourse and for possible recovery by grammatical rather than inferential means. Matching procedures are much quicker and much easier than inferential processes on the Model Knowledge Base.

In case a pronoun has been used to corefer, a description of the head it has been bound to should be present in the Model. In case a nominal head has been used, this might or might not be present in the Model.

As we said, f-structures are augmented by the output of the algorithm for anaphoric control, i.e. attribute-value pairs indicating the relation intervening between an antecedent and a pronominal expression which has been bound at sentence level, by means of indices.

When an f-structure is passed to discourse modules, the head of the antecedent as well as the head of a functional and a syntactic controller are substituted into the head of the controlled element. In addition, since we want only referentially free elements to appear at discourse level, controlled or bound elements are discarded.

The discourse module builds up a list of these elements which will eventually contain nominal heads and free pronominals: each of these elements is characterized by a vector of functional and semantic information which is used by the discourse module to weight them, and assign them to a rank list.

The output is the weighted list of topics, where referential elements are ranked according to their relevance in the current sentence. In the first sentence of a text the current Expected and Potential Topics are chosen from referring expressions available according to rank.

In the following text, the Main and Secondary Topics are asserted, usually by reinforcing the Expected or the Potential Topic. From now on, however, the Model of discourse is available as well as the Domain of Consciousness: these may be accessed to guide the choice of the Main Topic from the currently available Weighted List of Topics or from the Model or else by picking the Subject of Consciousness.

Another important criterion for the setting up of a Main Topic is the state assigned to the previous segment of discourse: we use the following ones,

CHANGE, CONTINUE, RESUME, RETAINING, SHIFTING, CONTINUE_ANALYSE

Whereas a Change state indicates that a new entity is being referred to by some nominal or pronominal expression and is the current Expected Topic, a Continue state indicates that the same entity has been asserted as Main Topic. The remaining states are Resume, Retaining, Shifting and Continue_Analyse. Shifting is used whenever there is a grammatically marked focalized constituent: in other words, a presentational construction has been used in the text, either by Locative Inversion or a There-sentence or any other available means offered by the specific language. This causes the focalized constituent to be set as Main Topic. As to Continue_Analyse, we use it to indicate the fact that a previously described entity which was a set, is now being reintroduced as a subset.

Resume is a state that is invoked any time the first referring expression on the Weighted List of Topics is not available in the adjacent portion of text and must be
recovered from the Model. The entity is set as Expected Topic. Finally, a Retaining state indicates that the current utterance is promoting one of the Topics present in the adjacent portion of text as possible new topic of discourse: in this case, the first referring expression on the WLT does not coincide with the previous Main Topic. It will be set as Expected Topic.

3.1.2 Pronominal Anaphora. Pronominal binding at discourse level may be achieved in three different ways: either by simply matching functional features between the pronoun/s and the possible antecedents available in the adjacent discourse segment. Let’s dub this first mode as grammatical binding at discourse level. In this case, the system activates a strategy for choosing the most adequate antecedent which is based on a number of cues automatically set up by the general algorithm. Discourse states provide a first cue: in case a Continue is present, a Main Topic should be available; in case a Change is present in the previous stretch of discourse, an Expected Topic is available and so on. In case there is only one pronoun, a first match is tried with the adjacent most relevant possible antecedents. However functional features may direct the choice to other previously asserted Topics, like Secondary or Potential Topic. When two pronouns are present, the Weighted List of current referring expressions is used as the main cue for finding possible antecedents. Actually all other cues, i.e. discourse state and previous Topics relevance should conspire in directing the algorithm to the best choice. And this is what happens in our texts.

However, grammatical binding is not always possible, either because of some inconsistency due to inferential processes or because of domain restrictions. The second mode is then the one related to the presence of a pronoun and a definite NP which could modify the expectations built up by the general algorithm. In particular, in case there are two important characters, it is important to keep under control definite NPs in order to infer some general property related to one of the character and then leave the pronoun for the usual matching devices with the other character. In case no such inference takes place, inconsistencies may arise because it might be the case that the actual Main Topic is not the adequate antecedent for pronominal reference. This information can only be captured once a definite NP is checked by inferential means to be a property belonging to the Main Topic. In this case, binding the pronoun and coreferring the definite NP to the same entity would result in a clash, which might be dubbed as a case of Obviation or of Disjoint Reference.

Finally the last mode takes advantage of Discourse Domain: this notion is related to the need of segmenting the text into Objective and Subjective Domains. This subdivision is taken advantage of whenever a Subjective Domain is established and some character is assigned as Subject of Consciousness. This usually happens whenever the character is already asserted as the Main Topic of a given discourse segment. In case a SC is present, the weighted list of referring expressions for the current sentence is only used to introduce Expected or Potential Topics: the Main Topic on the contrary is always fixed to the current SC. This is also applicable to cases of “atmosphere statements” in which no new entity is being introduced, but there is the need to maintain the SC as Main Topic, so that it would be available to any pronoun or definite NP intervening in the subsequent stretches of text (see Text 4. below).
3.1.3 Nominal Anaphora. Before entering this module, the algorithm substitutes the head of the pronoun with the head of its antecedent. In this way, all free pronouns are eliminated from the current list of referring expressions. In fact there are cases in which some entity is introduced in the discourse as an indefinite pronoun and no antecedent may be found in the adjacent stretch of discourse. These cases are dealt with in the Model, by asserting the indefinite entity as an entity belonging to the class defined by semantic selectional restrictions associated to that NP by the verb.

In all other cases, we always deal with nominal heads of some kind, be they proper or common names. The module has two main tasks: one of coupling the first or the second ref_exs in the weighted rank list with any of the previously asserted Topics; the other of asserting the new state of discourse.

The state of discourse guides the selection of matching procedures between current referring expressions and previous Topics. In case the coupling or matching does not succeed because there is no relation between the adjacent discourse segment and the current sentence, the state of discourse should be a shifting. Or else the Discourse Domain should provide a Subject of Consciousness to introduce as Main Topic to mark the fact that the current sentence is a subjective sentence. In this case the first referring expression in the list will be asserted as Expected Topic, in case the following text might be shifting to this newly introduced entity.

Special procedures are always invoked every time the current sentence carries functional information which evaluates one argument as focalized: this always causes a Shifting to occur and a double operation to be computed. First, the focalized argument is asserted as Main Topic, second the previous Main Topic is asserted as Secondary Topic to mark its possible persistence in the text.

This module accesses the World Model of the previous text whenever needed. In particular there are two possible situation in which this is compulsory: whenever the current ref_ex is a proper name which however corefers with a property already associated to that name in the previous text. In this case the algorithm should compute this as a Continue and a reassertion of the same Main Topic.

A subcase is represented by the situation in which an entity is simply reintroduced with its proper name but it was not present in the adjacent portion of text. The entity is simply recovered from the model and asserted as Main Topic. Another subcase is constituted by the case when the current ref_ex is a proper name but the previous Main or Expected Topic was a plural common noun, a property already associated with the current name.

3.1.4 Some Examples. Consider briefly Text 1: when utterance 3 is processed, a subject of consciousness is established with the pronoun “she”, which is resolved in the Model as coreferring to Mary. Utterance 4. has a plural pronoun and no antecedent available in the adjacent text segment. The Model is invoked in order to search for two fully specified individuals, John and Mary, which are assigned as antecedent of “them”; this in turn binds “you” in the following complement clause. Utterance 5 is a case of Implicit Subjective Domain whose content is ascribable to the Subject of Consciousness(S-C), and is evaluation of a previous assertion: “it” is bound to a previous relation. Also Utterance 6. is a case of Implicit Subjective Domain where the LDBR is bound to the S-C. Also consider the binding of “his” in utterance 7. which has no local antecedent, but is computed
on the basis of “recency” criteria: Steve is the most recent proper name, or fully specified individual available as a singular entity.

Now consider Text 2. where a S-C is established in utterance 2, and is a proper name, “John”. In the following utterance, the system computes an Implicit Subjective Domain, a nonfactive assertion, a Continue with John as Main Topic, and the same S-C. In this way, when utterance 4. is reached, the pronoun “he” is bound to John rather than to Richard. Utterance 5. is again an Implicit-Subjective Domain so that when the following utterance is processed, the LDBR “himself” is adequately bound to John.

Text 1.
1. John gave Mary a rose.
2. She took it and put it in her hair.
3. She knew that she had been given a present, something precious.
4. When Steve faced them saying: “Are, you, enjoying, yourselves?”
5. It was horrible! It was shocking!
7. She felt only hostility and his determination to ruin that wonderful moment.
8. John smiled and went away embarrassed.

Text 2.
1. The three friends went all outside.
2. As they were walking in the garden, John said to himself “Sara will marry that man“, without any resentment.
3. Richard would marry her.
4. He felt strongly all this.
5. She was the right person for a man like Richard.
6. As for himself, he was absurd.
7. His demands upon Sara were absurd.
8. She would have accepted him still if he had been less absurd.
9. Richard began to sing.

Text 3.
1. Mary picked up the phone and called Jason.
2. Her husband, she thought, would have considered such a move as untruthful and utterly base.
3. Perhaps there was something in herself that could not help but do the wrong thing at the wrong time.
4. Jason answered immediately.

Text 4.
1. John went into a restaurant.
2. There was a table in the corner.
3. The waiter took the order.
4. The air was nice and clean.
5. He took the cup and drank his coffee.

3.2 Subjective/Objective Domains and the Subject of Consciousness. J. Wiebe and W. Rapaport in their paper (1988) present a computational theory for recognizing discourse passages which are told from the perspective of a character. In what follows, the two authors go on defining what they mean by perspective in narrative simply by using Ann Banfield's (1982) - which we also assume to be a landmark for literary text analysis - categorization of the sentences of narration into subjective and objective sentences. And finally they assume that “our task of recognizing the current perspective is, therefore, to recognize subjective sentences and the subjective characters to whom they are attributed” (131).

However, we take Banfield's categorization of sentences and the task of establishing the current perspective as a non sequitur. This is very much so if we consider the further fact that W & R make the claim that whenever a subjective context is independently established by Banfield’s categorization, we are dealing with the character’s beliefs, seen that what is being told in the narrative reflects his perspective.

Let’s consider more closely Banfield’s categorization: a subjective context is established by subjective sentences which may either portray the character’s thoughts (represented thought) or present a scene as a character perceives it (represented perception). In particular, such verbs as “hear, see, realize, know, think, wonder, remember, want” are regarded as markers of subjectivity.

In our system, these verbs are classified into separate categories:

a. Mental activity verbs
   think, wonder
b. Stative, Presuppositional and Factive verbs
   see, hear, remember, realise, know
c. Intensional verbs
   want, desire, wish, expect
d. Subjective
   consider, believe

It is our opinion that only a. c. and d. verbs may attribute a sentence to a subjective domain; on the contrary, stative and factive verbs only depict objective facts: as a consequence, the object of any such verbs is always regarded as a state or a fact in the world, and should be so understood by the reader, contrary to what the two authors assume.

Another important fact is the lack of consideration for structural contribution to the categorization of sentences in a subjective or objective domain. If we take a verb like “say” or “tell” they are classified as Reportive verbs: however, in case the addressee is a reflexive pronoun as in one of our examples, the context is pre-subjective. “John said to himself that...” is understood as the indication that something objective is happening, the saying relation, which however preludes to a subjective context. In other words, in this situation,
the system sets up an initial boundary for subjective domain, as discussed in some detail here below.

3.2.1 Long-Distance-Bound Reflexive Pronouns. Anne Zribi-Hertz was the first generative linguist to assume that the binding of Long-Distance-Bound (LDB) Reflexive Pronouns in English, that is, those pronouns that search for their antecedents outside the sentence domain, must be viewed from the level of Discourse Grammar. She introduces concepts such as Subject-of-Consciousness and Domain-of-Point-of-View to demonstrate that it is necessary to rely on semantic discourse concepts in order to explain the behaviour of LDB anaphoric expressions.

This means that e-comand does not rule over such occurrences of pronominal expressions and works only at sentence level.

Her analysis represents a step forward to the understanding and conceptualizing of discourse grammar and, at the same time, this kind of approach confirms that an exclusively syntactical sentence-internal approach is insufficient and inefficient.

Sells(1989) reaches almost the same conclusion as Zribi-Hertz but inside a different theory: he introduces concepts such as SELF, PIVOT, SOURCE, that is, few discourse roles which affect the distribution of LDB pronouns.

He further shows that LDB pronouns should be read as logophoric: their antecedent is always an entity whose thoughts and feelings are represented in the clause of the pronoun.

3.2.2 Subject-of-Consciousness and Domain-of-Point-of-View. Subject-of-Consciousness is a semantic property assigned to a referent whose thoughts and feelings are represented by a portion of discourse.

Authors like Virginia Woolf and James Joyce are masters in the use of this subjective fiction, where subjective means that facts are depicted as filtered by a conscience, that of the Subject-of-Consciousness, contrary to the objective narrative where no consciousness intervenes.

The Subject-of-Consciousness is a concept already pertaining to literature that has to be formalized because it is a factor that plays a fundamental role in discourse and in particular it helps in resolving the mechanism or binding of LDB anaphoric expressions: the Subject-of-Consciousness results as the antecedent of such LDB anaphoric expressions. As a result, to be able to identify it, is paramount for the interpretation of anaphoric processes.

Before giving our formal definition of SC, we need to define the concept of Domain-of-Point-of-View (DPV), since these two are interrelated and interdependent concepts.

A DPV may be formally defined as a portion of discourse which is the grammatical expression of one and only one narrative point-of-view; the point-of-view being the perspective used by the author to describe the facts, real or hypothetical, in a text: it may be the author himself, the main character, other characters, or there may be no point-of-view at all.

We have identified at least four different kinds of Domain-of-Point-of-View not just two as it was suggested by Zribi-Hertz’s analysis; each has its own particular structure, its features and functions.
We also built some tests that help identifying the different kinds of DPVs.

3.2.3 Objective Domain-of-Point-of-View. The Objective DPV (OD) is taken to be the default domain, that is, the domain in which reality is what a written text focusses on. In an objective DPV, reality is depicted as not being expression of any point-of-view, or rather, an objective DPV does not fall under the scope of any Subject-of-Consciousness.

In English, in an objective DPV usually the aspectual (or semantic) category of the verb in the main clause is a process, that is, in an OD the Temporal Focus moves.

We even noted that normally (not always) whenever we pass from a Subjective DPV to an Objective one, the old SC is re-established through the use of a proper name.

3.2.4 Pre-subjective Domain-of-Point-of-View. A Pre-Subjective DPV (PSD) is still an Objective DPV in the sense that there is still no SC, but the presence of a particular formal mark acts like a door which introduces necessarily as its object an Explicit Subjective Domain.

The formal mark that characterizes a a Pre-Subjective DPV is represented by the category of the verb: in a PSD the verb may be a psych verb, a verb of mental activity, or an emotional verb.

In a Pre-Subjective Domain the Subject-of-Consciousness of the depending Explicit Subjective Domain is instantiated through the use of a pronominal expression (i.e. it is the thinker, perceiver, or senser present in the PSD) and it is identified through few semantic roles as well as on basis of syntactic information:

pre_subjective_cat(subjective, [experiencer, actor, theme_nonaff]).
pre_subjective_cat(presuppositional, [actor]).
pre_subjective_cat(emotional, [experiencer, actor, theme_emot]).
pre_subjective_cat(reportive, [actor]).

Let us look at one example to clarify this point:
Ex 1 [psd Mary felt that] [esd she was unable to say "no".]

Here, the main clause is a PSD due to the presence of 'felt' which is an emotional verb; therefore, it necessarily introduces an Explicit Subjective Domain whose characteristics will be discussed below.

Note that in the PSD the SC is established ("Mary"), that is the "Senser", and that in the ESD it is present in the form of a personal pronoun.

The following sentence is part of a wider discourse that we have analyzed. The output of the computational analysis is shown:

Ex 2: She knew that she had been given a present, something precious.

RHETORICAL STRUCTURE:
stato(3, retaining)
topic(3, main, id2)
topic(3, secondary, id11)
CLAUSE IDENTIFIER: 3-n1
CLAUSE TYPE: main/prop
FACTUALITY: factive
CHANGE IN THE WORLD: null
RELEVANCE: background
DISCOURSE RELATION: description
DISCOURSE DOMAIN: pre_subjective(3-n1, sn3, she)
SUBJECT OF CONSCIOUSNESS: none

CLAUSE IDENTIFIER: 3-n18
CLAUSE TYPE: fcomp/prop
FACTUALITY: factive
CHANGE IN THE WORLD: null
RELEVANCE: background
DISCOURSE RELATION: narration
DISCOURSE DOMAIN: explicit_subjective
SUBJECT OF CONSCIOUSNESS: she/sn3 from 3-n1

3.2.5 Explicit Subjective Domain-of-Point-of-View. The preceding example shows what happens as a rule: there is a strict correlation between the Pre-Subjective DPV and the Explicit Subjective DPV (ESD), that is, the ESD always follows a pre-subjective DPV and the SC of the ESD is always instantiated in the PSD.

Direct speech is always treated as an objective domain, rather than as an Explicit Subjective DPV: when we are in presence of a direct report it is evaluated objectively by the reader not as something reported by the SC but as something viewed from an external position.

3.2.6 Implicit Subjective Domain-of-Point-of-View. The Implicit Subjective DPV (ISD) is more complex, but more interesting for our purposes: it expresses the thoughts and feelings of the current SC which is not syntactically present and remains implicit. The important fact is that all the pronouns (especially LDB reflexive pronouns) found in an ISD all refer back to the current SC.

First, we have realized that the Subject-of-Consciousness of an ISD is always the SC of the last preceding ESD. Thus, we always require the preceding DPV to be an Explicit Subjective one.

Furthermore, in order to individuate ISDs we have formulated some tests that allow us to single out this kind of DPVs, that is, we have found some formal marks which are sufficient conditions for Implicit Subjectivity.

If a verb is a state or indicates existence, this confirms Subjectivity. Secondly, we have seen that exclamations, questions in indirect report and the presence of modals (i.e., would, could, should, must, may) or intensional verbs are all manifestations of the current Subject-of-Consciousness and then always indicate Implicit Subjective DPVs.

As also reported in W & R’s paper, few adverbs seem to be 'subjective', as they express the will of a character and his personal judgement: for instance, 'of course', 'perhaps', 'literally', 'obviously' and many others.
In conclusion, we have individuated four different DPVs which establish the SCs, but there may be other factors at work, semantic or syntactic, which will allow us to deepen our knowledge of discourse segmentations into Domains-of-Point-of-View.

3.2.7 Domain Boundaries. Recognizing the structure and characteristics of each Domain-of-Point-of-View is paramount for individuating the Domain Boundaries: since a piece of literature is a sequence alternating Subjective and Objective DPVs, we have to establish where they begin and where they end.

For our purposes, what is relevant is the capability to detect the Subjective Boundaries: we have assumed that a Subjective Boundary possesses a formal subjective mark such as the presence of an Explicit Subjective DPV, and may potentially open an Implicit Subjective DPV without formal marks. In this last case, the presence of a modal, or exclamative sentence acts as a Subjective Boundary.

In sum, we have established that Implicit-Subjective, Explicit-Subjective and Pre-Subjective+ Explicit-Subjective DPVs may be Subjective Boundaries as easily seen below:

\[
\text{disc_domain}(., ., ., \text{CatSem}, ., ., \text{PrecDom}, \text{implicit_subjective}, \\
\text{SubjConsc})
\]

:\-

\text{stative_cat}(	ext{CatSem}), \\
\text{subjective_boundary}(	ext{PrecDom}), \\
\text{subject_of_consciousness}(	ext{SubjConsc}).

\text{subjective_boundary}(\text{implicit_subjective}). \\
\text{subjective_boundary}(\text{explicit_subjective}). \\
\text{subjective_boundary}(\text{pre_subjective}(., ., .)+\text{explicit_subjective}).

3.2.8 An example. The following example is a brief discourse where the LDB reflexive pronoun "herself", being in an Implicit Subjective DPV, is computed as coreferent with the Subject-of-Consciousness previously established with the procedure we have described in this paper.

(D1)

[mary, picked, up, the, phone, and, called, jason]

CLAUSE IDENTIFIER: 1-n4
CLAUSE TYPE: coord/prop
FACTUALITY: factive
CHANGE IN THE WORLD: null
RELEVANCE: background
DISCOURSE RELATION: narration
DISCOURSE DOMAIN: objective
SUBJECT OF CONSCIOUSNESS: none
This first sentence consists of two coordinate clauses; the DPV is objective, there being no subjective markers, this is further confirmed by the presence of two non-stative verbs.

\[(D2)\]

\[
\text{[her, husband, ', she, thought, ', would, have, considered, such, a, move, as, untruthful, and, utterly, base]}
\]

This construction is an indirect report where the main clause is a Pre-Subjective DPV due to the presence of a mental activity verb: the reportive clause is an Explicit Subjective DPV whose SC is 'she', that is, the 'senser/thinker' of the verb in the main clause. The ESD is a Subjective Boundary.

\[(D3)\]

\[
\text{[perhaps, there, was, something, in, herself, that, could, not, help, but, do, the, wrong, thing, at, the, wrong, time]}
\]
Reference Resolution

CLAUSE TYPE: main/prop
FACTUALITY: factive
CHANGE IN THE WORLD: null
RELEVANCE: background
DISCOURSE RELATION: description
DISCOURSE DOMAIN: implicit_subjective
SUBJECT OF CONSCIOUSNESS: she/sm23 from 2-n2

Being the preceding DPV a Subjective Boundary, the presence of a stative verb in this there- sentence confirms that this is an Implicit Subjective DPV. Other indicators of subjectivity are the possibility adverb “perhaps”, the presence of a predicative complement which is an “assertive indefinite pronoun” leaving the reference unspecified. The reflexive pronoun "herself" is successfully bound to the Subject-of-Consciousness established in the last ESD.

4. Definite and Indefinite NPs. Anaphors is a general term for a range of expressions that are context-dependent in that they either specify entities in an evolving model of the discourse that the listener is constructing, or they depend on other entities in the discourse model. They are called Discourse Anaphors in Webber’s paper on Tense (1988).

The dependency of a DA on a discourse entity may result from the ontology of the specified entity, as well as from discourse structure and its focusing effect.

It has been argued extensively in the literature that definite NPs are exclusively used to corefer or cospecify entities already in the DM, whereas indefinite NPs can be used to introduce new entities in the world. However, we shall stress the need to ensure that both definite and indefinite NPs can be used to corefer or cospecify entities and relations in the DM. Besides, we also should note that both kinds of NPs can be used to introduce new entities in the world.

4.1 Indefinite NPs. Consider first indefinite NPs: they may cospecify some previously asserted relation or they may introduce some generic property which is already inferable from the model:

1. Mary picked up the phone and called Jason.
2. Her husband would have considered such a move as base.

The NP “a move” is an event noun cospecifying the calling event, a relation previously asserted in the DM.

3. They appointed John managing director.
4. This was a position he had been longing for for ages.

The NP “a position” is an activity noun cospecifying the role property “managing director” asserted in the previous discourse segment and associated to John.

5. Richard would marry Sara.
6. He felt strongly all this.
7. She was the right person for a man like Richard.

The NP “a man” is a generic noun which corefers with Richard.

8. Once upon a time there were three little pigs.
9. They decided to build a little house each.

The NP “a little house” is asserted as a set and not a fact and a cardinality is assigned to this set due to its being in the scope of the distributive quantifier “each”. In turn, the floating quantifier is computed as an open adjunct controlled by the subject “They” which corefers with the antecedent NP “three little pigs” in the previous segment of discourse. In another version of the same story, the little house become a straw hut and then a little house again, as shown in the following excerpt:

10. As they reached a nice wood, they decided to build each a comfortable little house.
11. Timmy didn’t like working at all, so he thought to build quickly a straw hut.
12. Soon the little house was ready

The NP “a straw hut” is understood to be one of the little houses already introduced in the previous discourse segments. However, knowledge of the world is called for in order to make the appropriate inference. Also notice that the following sentence uses a definite NP to corefer with the previous indefinite NP, and here again the same problem arises: the hut becomes a little house. For sure, we want the DM to be consistent and to understand that there is only one little house under discussion, which however is a kind of house, a hut.

Finally the most interesting case: an indefinite NP which is understood as a generic property of an entity introduced in the model by the same relation.

13. John gave Mary a rose.
14. She took it and put it in her hair.
15. She knew she had been given a present, something precious.

The NP “a present” cospecify the NP “a rose” which has been previously introduced in the model by a giving relation. The dependency is contextually determined by the presence of an entity Mary which is assigned the same semantic role in both sentences: she is always the Patient argument of the predicate GIVE. The system can also recover the Agent argument which has been omitted and is represented in f-structure as a lexically-bound existential quantifier “exist”. At the end of the computations, we know that “a rose” is also “a present” which has been given to Mary by John. In order for the inferential mechanism to draw the relevant inference, tense is used: the past perfect can be regarded as a presuppositional tense, i.e. a tense that indicates that some fact or event took place in the previous portion of text. This can be assumed also by the presence of a presuppositional verb “know” that governs the sentential complement clause in which the giving relation is used.
4.2 *Definite NPs and Pronouns.* We shall start by quoting an example from Webber(1988, her 2.)

a. Wendy gave each boy a T-shirt.
b. They had a beautiful logo on the front.

Here we see that the plural pronoun is used to corefer with an indefinite NP which is singular, but is computed as being in the scope of a distributive quantifier, “each” which binds the cardinality of the indefinite to that of the noun “boy”, thus turning a singular to a plural NP, a set and not an individual.

Another interesting case is represented by her example 4. where a plural pronoun is made to corefer to a class related to a singular NP in the previous discourse by inferential processes,

c. The dachshund down the block bit me yesterday.
d. They are really vicious beasts.

The pronoun is computed as the controller of the property “beasts” which is in the open complement of the verb “be”; this in turn should be taken as a generic property associated to the definite NP “the dachshund” which the sentence evokes. Clearly this case requires the presence of an ISA Relation in the ontology of the entities, constituting the extra-linguistic knowledge required to understand the relations intervening in the text at hand.

We now comment her example 5. where a definite NP is introduced as a new entity associated to an existing discourse entity:

e. A bus came round the corner.
f. I signalled to the driver to stop.

The definite NP “the driver” is here understood as an entity associated with the “bus” mentioned in the previous sentence. In our system this can be achieved by an inference on the main location of the text. The driver is related to the bus, since the scenario was set before by mentioning the location in which the driver could be inferred as being a part of, or better a role linked to that frame. Better examples of this problem are constituted by the texts provided in the paper by Garrod and Sanford(1988), where the scenario effect is more marked:

16. John went into a restaurant.
17. There was a table in the corner.
18. The waiter took the order.
19. The atmosphere was warm and pleasant.

The scenario is set in the beginning of the text, either by a title “At the restaurant”, or by explicitly mentioning the restaurant as an Oblique argument of a going relation where the main Topic is also introduced. When we get the definite NP “the waiter”, an inclusion
relation inference has already being fired by the location "the corner" which is understood as being a part of the main location, the restaurant. Thus, "the waiter" can now be computed as being in an inclusion relation with "the restaurant", being a role pertaining to that scenario or frame. Another important side effect of the analysis is that the Main Topic, John, is not discarded in favour of the new entity "the waiter", but persists in the Discourse Module. Thus, in case the text continues with what the authors call "psychological atmosphere statements", we may impute the subjective judgment on the Main Topic, John, rather than on a possible Expected Topic, the waiter.

More complex examples can be derived from our texts, in which more than one character is introduced in the text. In particular, the story of the three little pigs has the most intricate plot we found, since we are given at the same time six possible entities to be used for further processing. This is the beginning for an abridged version:

Segment A.

Once upon a time there were three little pigs who lived happily in the countryside. But in the same place lived a wicked wolf who fed precisely on plump and tender pigs. The little pigs therefore decided to build a small house each, to protect themselves from the wolf. The oldest one, Jimmy who was wise, worked hard and built his house with solid bricks and cement. The other two, Timmy and Tommy, who were lazy settled the matter hastily and built their houses with straw and pieces of wood. The lazy pigs spent their days playing and singing a song that said, "Who is afraid of the big bad wolf?"

And these are the entities already present at this point of the computation:

i. a set of three little pigs
ii. a set of three little houses
iii. an individual, member of the set of little pigs, whose name is Jimmy
iv. a subset made of two little pigs, always members of the set of three little pigs, whose names are Timmy and Tommy
v. a single little house, member of the set of three little houses, made of bricks and cement, owned by Jimmy
vi. a subset of two little houses, included in the set of little houses, made of straw and little pieces of wood, owned by the subset of two little pigs, named Timmy and Tommy

As the text proceeds, coreference is activated by evoking one of the entities either by a property, the fact of being brothers or being pigs, or by explicitly indicating cardinality.

Segment B.

And one day, lo and behold, the wolf appeared suddenly behind their backs. "Help! Help!", shouted the pigs and started running as fast as they could to escape the terrible wolf. He was already licking his lips thinking of such an inviting and tasty meal. The little pigs eventually managed to reach their small house and shut themselves in, barring the door.

Consider the deictic singular NP "such an inviting and tasty meal", which requires knowledge of the world to compute the metaphor. We should also note that the sentence is a case of idiomatic expression computed by the parser at grammatical level. Reference to houses is introduced in the following text, however, there is a singular little house associated to the two lazy pigs, rather than a plural one: it is clear to us that we are still
Reference Resolution

talking about a set of two houses, and this achieved by means of the possessor's cardinality. However the following text, reverts this decision, in that it makes us infer that the two little pigs are now both in the same house.

Segment C.

In the meantime the wolf was thinking a way of getting into the house. He began to observe the house very carefully and noticed it was not very solid. He huffed and puffed a couple of times and the house fell down completely. Frightened out of their wits, the two little pigs ran at breakneck speed towards their brother's house.

Problems at this point arise in the reasoning mechanism since the only singular little house available is the one built by the wise brother, Jimmy, which is picked up for coreference. This is clearly wrong, because this is not the house that gets destroyed since it is still there in the last sentence. In order to prevent a failure in the reasoning process, we simply allow reference to a little house to be inferrable as belonging to the set of two little pigs under discussion, on the assumption that it is the closest one and is available for coreference in the previous portion of text. The inferential mechanism is always driven by a recency checker which estimates which is the closest topic being asserted: on that basis an inference is fired, and in case it is successful that topic is taken as being coreferent.

Other cases of definite NPs to be inferred from the ontology or extra-linguistic knowledge are listed in Webber's 1988 paper on CL, and are the following:

from shared culture, e.g. "the government"; the unique representative of a class, e.g. "the duck-billed platypus"; an entire class or set, e.g. "the stars"; or a functionally defined entity, e.g. "the largest tomato in Scotland".

5. Possession Relations.

The presence of a possessor is a property which adds some specificity to the head noun. In particular, it could allow to identify a certain object in the world, by the fact of its belonging to a certain possessor. In the text we analysed, the possessor is used frequently to tell different object included in a set apart, on the basis of the identity of the possessor. In particular, if we consider Segment A. again, where the text introduces a plural reference to house with the possessive pronoun "their", there are two possible inferences:

a. either the two little pigs built collectively a set of houses with two members;
b. or they built it disjointly and there is no set, contrary to the linguistic form used in the text, "their houses".

In fact, if we keep ourselves to the distributive reading induced by the presence of "each" in the previous portion of text, we are led to the conclusion that there should be three separate entities of the class "house" in the world.

However this is not actually what the text does, since I take the subsequent reference to a single house for Jimmy and a plural set for Timmy and Tommy to imply that we should consider the initial set as made up of two subsets and not of three individual entities. This is the reason of the transformation of a set with two members into a single entity, which can be used to corefer to the plural NP "their houses" in Segment A. with the NP "their small house" in Segment B. without any further specification. In turn, this latter NP
is coreferred in Segment C. simply as “the house” being the more prominent house locally available or last mentioned due to recency effects.

In the same segment we note then that in order to distinguish the reference to this now singular “house” from the other little house present in the model, this latter is evoked with a possessor’s specification, “their brother’s house”.

To recover the identifier of that house we make the system infer a property belonging to the class “brother” and search for a fully specified individual - someone who has been given a name - and possesses a house which has been introduced as a single entity. In case every little house were introduced both as set and as single individual there would be no straightforward way to draw the necessary inference. Or at least the inferential process required to recover that individual little house would have been by far more complicated.

In particular, one would have been obliged to compute first the individual entities who possess the property of being “brother”; then, by subtraction, calculate from the possessors’ identity - or their names - the single brother the genitive is now indicating. At this point the identity of the little house could be established.

We might regard the procedure by which we reached the conclusion to generalise the description of a set to the description of a single entity as a linguistically driven sloppy reading. There are empirical reasons that drove our decision: if we look at the longest version of the story of the three little pigs, we discover that it justifies our position. This is the relevant portion of text:

Segment A.
This is the story of three little pigs who went around the world seeking their fortune. Their names were Timmy, flute player, Tommy, violinist, Jimmy, great worker. As they reached a nice wood, they decided to build each a comfortable little house. Timmy didn't like working at all, so he thought to build quickly a straw hut. Soon the little house was ready and Timmy decided then to go and see what his little brothers were doing. At first he met Tommy the violinist. Also he, himself, did not have much wish to toil, so he was building a simple little house with sticks of wood.

Segment B.
Very soon, also the house of wood was ready. Like that of straw, it was not very resistant. But the two little pigs lazy had managed to finish their work in a short time and now they could enjoy themselves freely. While Timmy was playing the flute, Tommy accompanied him with his violin and together they were having a lot of fun.

Segment C.
Then, tired to make merry, they decided to go and see what their little brother was doing. They started walking and soon they reached Jimmy. The clever little pig was building a little house. But since Jimmy was farsighted and did not fear working hard, he built it with bricks and cement.

As we can see, the story starts by introducing the set of little pigs; then it assigns them names, and establishes thus their individuality. Also the little houses are introduced as sets, again by the use of a distributive quantifier. However, as the story continues, we see that each little pig builds its own little house separately. The identity of each little house is now preserved only by its specific property: in particular in Segment B. we see that in
order to cospecify each of the two little house previously introduced in text as single
entities, the linguistic form used is "the house of wood", even though this was the house
now under discussion. The following sentence, uses a pronoun "it" to corefer to the most
recently mentioned little house and a deixic with a property, "that of straw" to corefer to
the other singular little house already existing in the model. The specification of some
distinctive property is now required simply because in the model there are two entities of
the same class "house", both singular in number, i.e. existing as individuals, which are
being further specified as "not very resistant". This property, in force of the linguistic form
used in the text, is thus distributed over the two single entities.

As the story proceeds, in Segment C. we see that coreference with the little brother
Jimmy is achieved as in the previous version of the story, by the use of a possessive
pronoun. In this case, we might deem that the identity of the individual brother is recovered
by subtraction: first, the identity of the possessor is recovered, i.e. an inference is required
from the property "brother" to that of their names; then, a singular brother is searched for,
whose name does not match with the ones already assigned to the possessor. However, we
might also consider the possibility that the use of "brother" here is simply due to the need to
assert this property and to extend it to the set of little pigs. In fact, the following sentence
makes it clear to the reader that the topic is now set to Jimmy.

An important thing to notice, at this point, is the fact that the introduction of another
single entity belonging to the class "house" requires the use of an indefinite NP: another
possibility could have been the use of a possessive, "his little house". In both cases, the
system understands that it should add a new individual of the class "house" to the model
, since Jimmy does not yet own a little house. The first mention to a set of three houses is
computed as an intension, or a sit, and not as a fact. However, we take the use of an
indefinite NP a much simpler way to achieve the same goal. In this case, the factivity of
the proposition in which the NP is used, requires the indefinite NP to be regarded as a new
type, which however is included in the set of the three little houses mentioned at the
beginning of the story.

6. Proper Names

Differently from what happens with definite or indefinite NPs, notably the fact that
both types can be used either extensionally or intensionally to denote some entity, in the
case of proper names we know that we are always dealing with rigid designator of the same
individual in all possible worlds, as Kripke defined them. A proper name fixes the
reference to an individual in that it designates unambiguously that individual for any further
reference in the text.

However, the mechanism by which a proper name is used in a text is something that
deserves further scrutiny. In our texts, there are at least two ways to associate proper
names with individual entities. The first and more canonical method is the one that is
represented by a copulative construction, as in,

Example 1.
This is the story of three little pigs who went around the world seeking their fortune. Their names were Timmy, flute player, Tommy, violinist, Jimmy, great worker.

Another method, is the one represented by a predicative adjunct, as in,

Example 2.
The oldest one, Jimmy who was wise, worked hard and built his house with solid bricks and cement. The other two, Timmy and Tommy, who were lazy settled the matter hastily and built their houses with straw and pieces of wood.

These might be considered as more or less direct methods for name association with a property already existing in the world. A more subtle and indirect way of obtaining the same result is shown in the following example,

Example 3.
The three friends went all outside. As they were walking in the garden, John said to himself “Sara will marry that man”, without any resentment. Richard would marry her.

In this latter example, the association is implicitly achieved by the semantic import of the structural organisation of the utterance. A pronoun is used in the subordinate clause to corefer to the property “friends” in the previous text; then, the subordinator indicates coincidence of temporal relation between the main and the subordinate clause. The main clause, in turn makes the pronoun explicit and introduces proper names as prominent characters. The inference we are naturally drawing at this point is that the Subject pronoun “they” and the Subject of the main clause point at the same individual. However since the pronoun corefer with a set with cardinality 3, we are allowed to make further inferences: we assume that also “Sara” is the name of one of the entity denoted by the set of friends. The same applies to the appearance of “Richard” in the following sentence.

How are these assumption and inferences caused: I take the association rule for proper names to be formulated as follows:

**Rule for Proper Names Association**

Whenever a Proper Name appears in the world, check in the model if there is already some entity associated with that name;

In case the search fails, check whether there is an explicit (direct/indirect) association link with some entity in the current clause;

Else, check whether there is an implicit indirect association link with some entity in the local context.

In other words, we always require Proper Names to be associated with some previously asserted property in the local context. However, narratives show that a proper
name could be simply introduced as such, and be a new individual in the world, as for instance in,

Example 4.
Mary picked up the phone and called Jason.

or as in,

Example 5.
John went into a restaurant.

In these last two example, proper names are used to introduce some new entity in the world and the property assumed is simply a generic class specification in terms of selectional restrictions associated to that NP as argument of a given predicate.

Names are used freely in the following text to recover coreference to a given individual. In particular, they may be used in place of a pronoun, when the text would make its use ambiguous. In the first version of the story of the little pigs, names are introduced in the story, however no other mention is needed in the following text to recover the corresponding individuals. This is simply due to the fact that they are well distinguished as being either a set of three little pigs, a set of two little pigs, and a single individual: thus, the use of names becomes redundant.

However, the second version of the same story introduces the three little pigs at first as a set with cardinality three; but as the story progresses, each of the three little pig is introduced separately by its name. This is possible because priorly there has been an explicit association of names to each member of the set of three. At a certain point of the story, it would seem that talking of a single little pig induces ambiguity, however this is not so, as shown by Segment C, which we report here below.

Segment C.
Then, tired to make merry, they decided to go and see what their little brother was doing. They started walking and soon they reached Jimmy. The clever little pig was building a little house. But since Jimmy was farsighted and did not fear working hard, he built it with bricks and cement.

Rather than using the name for the first occurrence of a reference to Jimmy, we see that the text uses "their little brother", which is clearly less individuating as the three little pigs are all brothers: but the use of the possessive makes coreference clear. In the following sentence, we see however that a name is used: is this required or is it redundant? We take the use of the proper name to be not cognitively but textually required, since the use of a common name like "the little pig" would sound unnatural. Also, note that the use of a pronoun is impossible, since "their little brother" has not been established as a current Topic.

7. Discourse Structure Representation
We take this level of representation to be composed of relations of various kinds which interact with the domain and the state of discourse in order to check for their semantic consistency. However, differently from other approaches, in our case Discourse Structure is simply the result of all previous computations: we simply let previous semantic descriptions interact with a simple algorithm that takes care of structure in terms of UP and DOWN nodes by means of carefully organized PUSH and POP actions. We have been inspired by L.Polanyi’s (1988) proposal, who suggests that a structural representation should be composed of Subordination and Coordination relations between the clauses that make up the text under analysis. In our case, we see that a stretch of text or a discourse segment characterised by a set of Coordinates corresponds to a Topic Chain in which for instance, some properties of a Participant in the discourse are described.

In Polanyi’s model there are four possible parses intervening between two adjacent clauses A and B at any point in the analysis:

1. A is coordinated with B.
2. B is embedded relative to A.
3. B is subordinate relative to A.
4. A is superordinate to B.

Subordinate and Coordinate Nodes are created by the algorithm as the analysis proceeds. In order to embed some clause under a Subordinate node a PUSH action is executed; to exit the stack a POP action should be performed. Coordinate nodes are usually lists of clauses at the same level of attachment. As Polanyi notes, clauses attached under the same mother are accessible and pronominalization should be expected to hold in the Topic Chain by any daughter or rightmost node. On the contrary, whenever a POP to a higher level structure obtains, we should expect pronominalization in a Topic Chain to be barred. However, in case two characters are present, a Subordinate node could indicate the local shift from one to the other of the two characters and this should be marked off by the explicit mention of some property of the entity in focus. A POP from this level could be still performed by some pronoun, provided that the other entity is coreferred by the explicit mention of a property. The same result is achieved by our algorithm of discourse structure which receives as input Discourse States and Topic Structure, as well as Discourse Relations and Temporal Relations. However, a more fine-grained description, clearly would require more local computation which could be directed at the assessment of the semantic congruence of discourse segments as they are produced independently by our algorithm. As Polanyi comments,

"How semantic congruence is ascertained is an important issue. This process of semantic analysis is a world-knowledge and inference-driven semantic matching process making use of extra-linguistic knowledge, the meaning of the words and the structures encountered to perform an analytic and possible matching operation on the semantic values encoded in the semantic frames associated with the various nodes." (ibid.617)
Since we assume that extra-linguistic knowledge should be brought to bear independently by the system only when needed, we take semantic congruence to be just a contextually driven process.

Semantic relations in any given text are the main task to be faced when building structural representation. These relations are described by Discourse Relations and Temporal Relations. Consequently we shall label the nodes of the DSR with this information, as well as with facts or situations containing relations in which Topics are involved. We must point out that we use Discourse Relations as local markers of congruence in adjacency: no attempt whatsoever is made to build a higher structure that encompasses more clauses together. In addition, Discourse Relations are computed from aspectual, semantic category, temporal and syntactic information: differently from what both Polany and Mann & Thompson (1987) assume we note that there is no intervention of conceptual level reasoning. As to Coordination, Polanyi says that semantic congruence is obtained by a set of Generalised Union operations on clauses which express propositions conveying the values one function (the unifying property) has for a series of alternative arguments (the coordinated properties). In our case, once Discourse Segments are built, local reasoning could be invoked in order to ascertain whether Generalised Union Operations could be performed. In her example, a coordination obtains between the following sentences:

a. John is a very good athlete.
b. He can run a four-minute mile.
c. He throws a mean hardball, too.

in which the unifying property is the one expressed by the first sentence, and the following sentences are instances of this property: they must be in an ISA Relation with one another proceeding to the right. Other conditions are represented by the fact that they should all refer to the same Participant, and they should express the "most restrictive relevant natural set", in other words, they should form a sequence from the more general to the more specific property. This is the only example she discusses in her paper, and we don't know how her theory would work on more complex cases, such as the ones we present in this paper. However, the details of the underlying mechanism are left for further research!

In our system, information on the structure of discourse can be gathered from an extensive number of sources. In particular, the Module for the Resolution of Anaphora at Discourse Level, is itself a local finite state machine that parses the text, at the level of utterances. These indications should be consistent with the DSR as proposed in what follows. Also relevant to the issue under discussion is the Rhetorical Structure Representation, where we indicate the list of Topics present in a given utterance as well as the Discourse Domain, be it Objective or Subjective, together with the Subject of Consciousness in case there is one. Finally the Main Spatial Location and the Main Temporal Location are used to assign indeces to entities in the world: they should be consistent with other discourse markers.
We could say, that every time text progression is marked by the presence of a number of clauses related to the same Main Topic, these clauses should be regarded as a Segment or an Episode at the level of Semantic Relations affecting a given Entity in the text.

Every time a New Entity is introduced in the text, some interruption is brought to bear on the DRS, and this should be captured by an upward movement from one structural level to another.

It would seem that a POP action should take place every time a new Topic is added in prominent position, and it was not included in the previous list of Topics; in addition, the new attachment level is determined by the congruence of the current Clause as to the participants in the main Semantic_Relation, to the level in which these are present as Topics. One of the prediction that the model enables us to do, is that in case Pronominalization occurs, it will affect all the Topics visible at a certain level. A POP action will cause an UP node to be produced and this in turn will indicate that a nominal head has been used to introduce or reintroduce a given Topic.

On the contrary, a PUSH action takes place every time the previous Topic, be it Expected or Main, is asserted as Secondary and there is a persistence of the same previous Topics.

The existence of a Subjective Domain with a Subject of Consciousness requires the permanence at a certain level of Coordination.

In conclusion, a Discourse Structure is a set of Segments or stretches of Discourse or Text which are marked off by nodes: the following nodes are generated by our algorithm,

1. ROOT - to mark the beginning of a story
2. UP - to mark a break in the current Segment and a movement upward
3. SAME_LEVEL - to mark a subordination or coordination of a clause to the current Segment
4. DOWN - to mark an embedding movement in the current Segment

The definition of Root is self-explanatory, and we will not comment on it. As to the UP node, it occurs whenever there is an interruption in the current Segment: this might be caused either by a return to a previous Topic by means of a nominal expression which denotes some property of the Topic, or by the appearance of a new Topic. In the former case, the algorithm will indicate clause and utterance number of the attachment node; in the latter case, the UP node will simply be attached to the root.

The Same_Level node is used to set off segments of discourse. They are so regarded by the algorithm on the basis of two main rhetorical strategies:

1. a sequence of clauses can be analysed as a Segment because of its underlying Domain, and it must be a Subjective Domain. In this case, the Discourse Focus does not move forward and is stuck to the clause setting the beginning of the Segment. Discourse Relations may either be Descriptions, Elaborations, or Explanations;

2. a sequence of clauses is analysed as a Segment because it has the same Main Topic. In this case, the Discourse Focus is moved forward and the story progresses by
Enumerating a number of properties related to the same entity. Discourse Relations may be Narrations or any other previously mentioned Relation.

The second strategy is simply a Default strategy, and is clearly inherent in the first one. The Down node is used to mark off the beginning of a possible Segment of discourse, or simply a movement backward of the story where the main topic is however the same as the current one. In the latter case, a past perfect could be used to trigger the appropriate Temporal Relation, a BEFORE relation, and the related Discourse Relation, an Elaboration relation.

As a result, any reasoning based on cognitive means should build upon the structural representation as it is independently worked out by our algorithm, by adding further internal structure. We assume that conceptual reasoning implied by Rhetorical Structure Theory as proposed by Mann and Thompson, or by Polanyi could not possibly disrupt our structural representation, which is mainly Topic based and relies on local semantic relations rather than on global relations.

We include below some examples of Discourse Structures as computed by our algorithm.

```
TEXT 1.
root:new(1-1)
clause:1-1
topics:{expected:id1:john}
main_fact:give({id1:john, id2:rose, id3:mary}, 1)
temp_rel:overlap
disc_rel:narration
disc_dom:objective

  down:down(1-1)
  clause:2-2
topics:{secondary:id3:rose, expected:id2:mary}
main_fact:take({id2:mary, id3:rose}, 1)
temp_rel:after
disc_rel:narration
disc_dom:objective

  same_level:from(2-2)
  clause:2-3
topics:{secondary:id3:rose, expected:id2:mary}
main_fact:put({id2:mary, id3:rose, id8:hair}, 1)
temp_rel:after
disc_rel:narration
disc_dom:objective

  same_level:level(2-3)
  clause:3-4
```
topics: [main:id2:mary, secondary:id1:exist]
main_fact: know([id2:mary, id12:give], 1)
temp_rel: during
disc_rel: description
disc_dom: objective

down: down(3-4)
clause: 3-5
topics: [main:id2:mary, secondary:id1:exist]
main_fact: give([id2:mary, id11:present, id1:john], 1)
temp_rel: before
disc_rel: elaboration
disc_dom: subjective

same_level: from(3-5)
clause: 4-6
topics: [main:id2:mary, secondary:id17: [mary, john], expected:id16:steve]
main_fact: face([id16:steve, id17: [mary, john]], 1)
temp_rel: after
disc_rel: narration
disc_dom: objective

same_level: level(4-6)
clause: 4-7
topics: [main:id2:mary, secondary:id17: [mary, john], expected:id16:steve]
main_fact: say([id16:steve, id20:enjoy], 1)
temp_rel: after
disc_rel: narration
disc_dom: objective

same_level: from(4-7)
clause: 4-8
topics: [main:id2:mary, secondary:id17: [mary, john], expected:id16:steve]
main_fact: enjoy([id17: [mary, john], id17: [mary, john]], 1)
temp_rel: during
disc_rel: description
disc_dom: subjective

down: down(4-8)
clause: 5-9
topics: [main:id2:mary, secondary:id16:steve]
main_fact: be([infon110:shocking], 1)
temp_rel: during
disc_rel: evaluation
disc_dom: subjective
same_level:level(5-9)
clause:5-10
topics:[main:id2:mary, secondary:id16:steve]
main_fact:be([infon110:shocking], 1)
temp_rel:during
disc_rel:evaluation
disc_dom:subjective

same_level:level(5-10)
clause:6-11
topics:[main:id2:mary]
main_fact:be([infon135:herself], 0)
temp_rel:during
disc_rel:explanation
disc_dom:subjective

same_level:level(6-11)
clause:7-12
topics:[main:id2:mary, secondary:id16:steve]
main_fact:feel([id2:mary, infon144:[determination, hostility]], 1)
temp_rel:during
disc_rel:explanation
disc_dom:subjective

up:to(1-1)
clause:8-13
topics:[secondary:id2:mary, expected:id1:john]
main_fact:smile([id1:john], 1)
temp_rel:after
disc_rel:narration
disc(dom:objective

same_level:from(8-13)
clause:8-14
topics:[secondary:id2:mary, expected:id1:john]
main_fact:go([id1:john, id34:away], 1)
temp_rel:after
disc_rel:narration
disc_dom:objective

TEXT 2.
root:new(1-1)
clause:1-1
topics:[expected:id2:friend]
main_fact:go([id2:friend, id4:outside, infon15:all], 1)
temp_rel: overlap
disc_rel: narration
disc_dom: objective

same_level: level(1-1)
clause: 2-2
topics: [main:id2:friend, secondary:id9:john]
main_fact: say([id9:john, id9:john, id12:marry], 1)
temp_rel: after
disc_rel: narration
disc_dom: subjective

same_level: level(2-2)
clause: 2-3
topics: [main:id2:friend, secondary:id9:john]
main_fact: walk([id2:friend, id17:garden], 1)
temp_rel: during
disc_rel: parallel
disc_dom: objective

same_level: level(2-3)
clause: 2-4
topics: [main:id2:friend, secondary:id9:john]
main_fact: marry([id10:sara, id11:man], 1)
temp_rel: during
disc_rel: evaluation
disc_dom: subjective

same_level: level(2-4)
clause: 3-5
topics: [secondary:id2:friend, expected:id10:sara]
main_fact: marry([id2:friend, id10:sara], 1)
temp_rel: during
disc_rel: evaluation
disc_dom: subjective

same_level: level(3-5)
clause: 4-6
topics: [main:id9:john, secondary:id10:sara]
main_fact: feel([id9:john, id12:this], 1)
temp_rel: during
disc_rel: explanation
disc_dom: subjective

same_level: level(4-6)
clause: 5-7
Reference Resolution

topics:[main:id9:john, expected:id10:sara]
main_fact:be([infon91:person], 1)
temp_rel: during
disc_rel: explanation
disc_dom: subjective

same_level: level(5-7)
clause: 6-8
topics:[main:id9:john, secondary:id10:sara]
main_fact: be([infon117: absurd], 1)
temp_rel: during
disc_rel: explanation
disc_dom: subjective

same_level: level(6-8)
clause: 7-9
topics:[main:id9:john, secondary:id10:sara]
main_fact: be([infon130: absurd], 1)
temp_rel: during
disc_rel: explanation
disc_dom: subjective

upto(5-7)
clause: 8-10
topics:[main:id10:sara, secondary:id9:john]
main_fact: accept([id10:sara, id9:john], 1)
temp_rel: during
disc_rel: evaluation
disc_dom: subjective

downto(8-10)
clause: 8-11
topics:[main:id10:sara, secondary:id9:john]
main_fact: be([infon146: absurd], 1)
temp_rel: before
disc_rel: description
disc_dom: subjective

upto(1-1)
clause: 9-12
main_fact: begin([id20: richard, id37: sing], 1)
temp_rel: after
disc_rel: inception
disc_dom: objective
TEXT 3.
root: new(1-1)
clause: 1-1
topics: [expected:id1:mary]
main_fact: pick_up([id1:mary, id3:phone], 1)
temp_rel: overlap
disc_rel: narration
disc_dom: objective

same_level: from(1-1)
clause: 1-2
topics: [expected:id1:mary]
main_fact: call([id1:mary, id2:jason], 1)
temp_rel: after
disc_rel: narration
disc_dom: objective

same_level: level(1-2)
clause: 2-3
topics: [main:id1:mary, secondary:id8:husband]
main_fact: think([id1:mary, id10:consider], 1)
temp_rel: during
disc_rel: elaboration
disc_dom: objective

same_level: level(2-3)
clause: 2-4
topics: [main:id1:mary, secondary:id8:husband]
main_fact: consider([id8:husband, info: [base, untruthful]], 1)
temp_rel: after
disc_rel: evaluation
disc_dom: subjective

down: down(2-4)
clause: 3-5
topics: [main:id1:mary, secondary:id14:something]
main_fact: there_be([id14:[oggetto]], 1)
temp_rel: during
disc_rel: explanation
disc_dom: subjective

up: to(1-1)
clause: 4-6
topics: [secondary:id1:mary, expected:id2:jason]
main_fact: answer([id2:jason], 1)
8. References

* The system is an enlarged and updated version of a previous system with the same acronym presented at the 3rd Conference on ANLP, Systems Demonstrations, Trento(It), April 1992.

Dalrymple M. (1990), Syntactic Constraints on Anaphoric Binding, Ph. D. Dissertation, Stanford University.


Delmonte R., Bianchi D. (1991), Binding Pronominals with an LFG-Based Parser, II° IWTP91, Cancun.


