COMPUTING DISCOURSE ANAPHORA FROM GRAMMATICAL REPRESENTATION

RODOLFO DELMONTE³
DARIO BIANCHI³

³Dipartimento di Fisica - Viale delle Scienze Università degli Studi - 43100 P A R M A (Italy)
³Istituto di Linguistica - Ca’ Garzoni Moro San Marco 3417- Università di Venezia 30124 - V E N E Z I A (Italy)

(this is the extended version of a paper to appear on Research in Humanities Computing,CUP,London, with the same title)

Abstract
The paper presents an implemented algorithm to compute anaphora resolution and coreference, relatively to pronominal and nominal expressions in discourse. The algorithm works on the output of other modules: the parser and its associated module for pronominal binding within the sentence(see Delmonte & Bianchi, 1991); the module for scope assignment to quantified expressions(see Bianchi & Delmonte, 1989a; Delmonte, 1990); the module for inferential processes based on KL-Prolog(Adorni et al.,1987). Our model is strongly syntactically based: in particular, access to the modules is guided by the grammatical basis of the computation, the two remaining modules being triggered by specific items: the presence of quantified indefinite expressions, of universal quantifiers or generic expressions for the module of scope assignment; the failure of some referring expression to be picked up as Main Topic, thus conflicting with evaluation procedures independently set up by the algorithm, as the need to set up an adequate plural antecedent for a plural pronoun or noun to corefer. As to inferential processes, these are only activated at a given state whenever the scoring procedures independently set up by the algorithm require a nominal expression to corefer in the text. They can also be triggered in a given state whenever the system is in an ambiguous context: one or two pronouns to be coreferred and no MAIN or SECONDARY TOPIC.

0. INTRODUCTION
When understanding or producing discourse or text, human beings must follow strictly some general principles underlying the distribution of information which include not only coherence but also what might be regarded as biologically set mental constraints. These constraints determine the amount of information which may be computed at a given time when reading or listening to texts. In particular, any given text introduces one or more TOPICS of discourse or themes and builds up on it a structurally and semantically coherent argumentation which consists both of generic and specific descriptions, personal or generic evaluations, expansions and so on. It is a fact, that a human being cannot possibly follow this textual processes for more than one or two TOPICS at a time, even though he may well temporarily store the information related to a previously discussed TOPIC in a memory storage.

In our algorithm, we activate only two TOPICS at a time which may either be an EXPECTED and a POTENTIAL TOPIC in case the text just starts being organized; or, a MAIN and a SECONDARY TOPIC, in case the text has already started. Pronominal and nominal expressions introduced at a given point are computed accordingly. They may either CONTINUE a Main or Secondary Topic or they may contribute to turn an Expected and/or a Potential Topic into a Main Topic. Secondary Topics may only arise whenever a given nominal expression has already been used as Main Topic and is then temporarily stored into memory.

1. PRONOMINALS
1.1 Accessibility and referring
Following M.Ariel(1988) we believe that referential expressions are processed in a certain way according to their inherent definition or classification in terms of referential features: these in turn determine when a referring expression must be processed, if an anaphor as soon as possible - and we compute anaphors at sentence level only, whether it can act as antecedent or not and whether it is dependent on its reference on other expressions or whether it is free. A pronoun is free in a certain domain, but a noun or a proper noun is free anywhere. The context taken into account is viewed in terms of accessibility of the referent to the addressee. In case a referential expression requires Knowledge of the World, this is less accessible than previous linguistic material, which alone can provide the higher degree of accessibility for a given referent. In turn, a referential expression must be
a noun or a proper noun in case of ambiguity. In a stretch of discourse, whenever a pronoun is not usable because it can cause incoherence, being ambiguously referable to one or the other antecedent, then a noun or proper noun must be introduced.

We might consider the following properties as relevant for a separation of Ps or pronominal elements into classes according to a feature system, where the 0 property is the basic one distinguishing Anaphors from Ps:

0. can corefer in the discourse
1. can be assigned arbitrary reading
2. can be bound to a quantifier or a quantified NP
3. can be used for contrast and emphasis
4. can take split antecedents

Italian differs from languages like English because it possesses a greater number of lexical items to express the properties listed above. For instance, English has only one kind of pronoun to express properties 2, 3, i.e. lexical pronouns; and another kind of pronominals which expresses property 1 - i.e. big Pro. Only non anaphoric pronominals possess property 4. In English, contrary to what expected, 1st and 2nd person reflexives can be bound in the discourse; 3rd person reflexives can be bound in the discourse according to Zribi-Hertz (1990) principles of the Subject of Consciousness as antecedent and Domain of Point of View as discourse relevant domain.

As to properties 2,3, whereas English merges them within a single class, Italian splits them into two different classes: one kind of pronominals expresses properties 3 - i.e. independent lexical pronouns. Another kind of pronominals expresses property 2 - i.e. little pros and clitics. In turn, clitics and little pros can be treated as syntactic variables and be bound at c-structure level: in this case they would be seen as part of a syntactic chain. Possessing a higher number of lexical items, Italian has as a counterpart a more restrictive system of referentiality: reflexives cannot be used logophorically. Thus, it would seem that a system for pronominal binding in English would necessarily require the computation of discourse structure, with a Subject of Consciousness and a Domain of Point of View. Italian, on the contrary allows modularity to the system, keeping sentence level separated from discourse level analysis.

1.2 The syntactic and semantic basis
The f-structure representation that we produce is richer than the one proposed in the literature even though it is strictly within the basic theoretical framework. In particular, we use a lexical form which records both the syntactic constituent, the grammatical function and the semantic role associated to a given argument of a predicate; in addition, the semantic role is further specified by a list of inherent features. It could be argued that some of this information is redundant, in view of the fact that an Agent may only be a SUBJect and in turn a SUBJ may only be an NP: however this is not always the case, as happens with SUBject infinitives. Besides, open functions, XCOMPs which encode the same information as small clauses in a configurational approach, are assigned a major constituent in the grammar rather than in the lexicon, on the basis of the principle that X ranges over lexical heads and that the parser instantiate the appropriate c-structure representation accordingly.

Another important difference is constituted by the attribute lex_form, which records the grammatical processes undergone by the lexical entry when the parser analyses the sentence. In particular, a transitive verb like "muovere"/move, may undergo "inchoativization" and the output of this lexical redundancy rule is recorded in the lex_form associated with the parse, where the underlying obj/theme_nonaff has been turned into the subj/theme_nonaff, as shown by the following example taken from the story which we will comment at length below.

5. [la,casa,non,si,mosse,di,un,solo,palmo]/the house did not move an inch...
net(p=21)
index:f7
pred:muovere
lex_form:[np/subj/theme_nonaff/[oggetto, strumento]]
mood:ind
tense:pass_rem
cat:risultato
subj/theme_nonaff:index:sn106
cat:[oggetto, luogo]
Another important feature present at the level of f-structure is the output of the binding module. In line with LFG and other generative theoretical frameworks, pronominal expressions, morphologically expressed or unexpressed are bound at various levels of representation: functionally controlled ones are bound at lexical or structural level by the parser which produces as output a list of annotated c-structures, where controlled PROs are associated with their antecedents. Arbitrary PROs on the contrary, cannot be assigned antecedents being generic in reference (see Delmonte and Bianchi 1991). Generic readings are also produced whenever a pronoun is in a chain with a quantified antecedent and tense is not specific: this is computed at sentence level by a module of interpretation.

A second set of pronominal expressions, those which can be bound sentence internally, are computed by a binding module which follows the lines traced in LFG for 'anaphoric control': in addition, since Italian possesses a much wider variety of pronouns than English, rules for deictic pronouns, independent pronouns or contrastive pronouns, as well as for possessive reflexive ones (also called long anaphora) are also specified in terms of Functional Structures.

The module for pronominal binding operates strictly sentence internally and also specifies which pronouns must be given external reference, these being the object of analysis of the present algorithm. Subordinate sentences are dealt with by the binding module, since f-command is still a viable tool for domain accessibility and allows to treat differently structurally different sentences: in particular, the subordinate may precede the main clause or it may follow it thus resulting in a different configuration for antecedents and pronouns as we discuss in Delmonte and Bianchi 1991. Coordinate clauses are dealt with by a special module which allows to dispense with f-command, since no particular restriction seems to result from this kind of configuration, apart from precedence. It is a well known fact that binding of pronouns by means of a quantified antecedent may take place as long as precedence is respected: this is particularly true in Italian; on the contrary English allows backward chains to be realized (see Carden 1982).

There are differences which distinguish a language like Italian where verbal agreement morphemes can be used referentially and are instantiated into an empty ‘pro’ or Null Subject: from a language like English where the same morphemes must be computed as non-referential and a NS can only occur in specific contexts, e.g. as subjects of a conjoined proposition. Moreover, English can use a personal pronoun deictically or contrastively provided it is strongly stressed whereas Italian possesses a different lexical variety for the same case. Besides, non-subject pronouns in a language like Italian can become clitics - enclitics or proclitics, and can be bound within their utterance or in the discourse according to structural constraints.

Broadly speaking, one could say that Italian is a language structurally underdetermined but referentially overdetermined in the sense that the syntactic structure of Italian is highly ambiguous whereas the referential processes set up both at sentence and at text level are very well determined. The contrary may apply to English, which is structurally overdetermined and referentially underdetermined. It is a well known fact that English always requires a SUBj ect to be lexically expressed in preverbal position whereas Italian and Romance languages...
do not require it. Rather, they have a principle of pro-drop which allows the SUBJect to be left lexically unexpressed or to be inverted in postverbal position. Consider the following example:
6. i. pro Noto che pro non era solida
   ii. He noted that it was not solid.

The Italian version in i. has two little pro's where the English version uses two different pronominal forms, a personal form 'he' pointing to a third person human antecedent and 'it' indicating a non human antecedent. No coreference would result between the two pronouns when the binding module is activated. However, the Italian sentence only makes available an empty category with third person specification, more features should be provided by the grammatical module. In particular, verb subcategorization would assign a human semantic feature to the first pro in force of the fact that a SUBJect for the verb NOTE has to be 'human'. However, the second pro is associated to the SUBJect of a copulative verb, BE, which only indirectly governs this function.

According to LFG, an indirectly governed function is deprived of semantic features and is interpreted in the predicate, which in our case is the adjectival SOLID. Our system would provide the little pro of semantic features percolating from the SUBJect of the predicate adjectival SOLID, and transmitting via Lexical Functional Control its features, Person Number and Gender to the controller, the non thematic SUBJect, which happens to be an empty category, deprived of a lexically expressed element - a variable in prolog. Only in this way, binding is prevented to apply and no coreference would result between the two empty categories. This is clearly shown in the f-structure representation for the sentence here below,

7. [esso, si, mise, ad, osservare, attentamente, la, casetta, che, non, era, davvero, molto, solida] he began to observe

net(p013)
main/prop:index:f18
  coord:index:f5
    pred:metteresi
      lex_form:[np/subject/actor/[umano, animato], vpinf/vcomp/prop/a/[subject=subject/actor, subj=x]]
    mood:ind
    tense:pass_rem
    cat:attivita
    subject/index:sm21
      cat:[animato]
      pred:esso
      pers:3
      gen:mas
      num:sing
      case:[nom]
      spec:def+
      tab_ref: [+ref, +pro, -ana, +me]
      antecedent: external
      interpretation: specific
  vcomp/prop/index:f11
    pred:osservare
      lex_form:[np/subject/experiencer/[umano, animato], np/object/theme_nonaff/[1_1]]
    mood:inf
    tense:pres
    cat:perceittivo
    subject/experiencer/index:sm45
      cat:[umano]
      pred:pro
      binders:sm21
      tab_ref: [+ref, +pro, +ana, -me]
      interpretation: specific
    object/theme_nonaff/index:sm43
      cat:[oggetto, luogo]
      pred:casa
      gen:fem
      num:sing
      spec:def+
      mods:cat:[oggetto: 1]
      pred:piccolo
      gen:fem
      num:sing
      tab_ref: [+ref, -pro, -ana, +class]
Coming now to pronominal expressions, Italian possesses a higher variety of lexical forms performing different functions. In line with Bresnan et al. (1985) and contrary to the proposal contained in Dalrymple (1990) we use functional features as lexically specified properties of individual anaphoric elements. These features both account for and translate lexical category, in this way directly triggering the binding algorithm. Features also serve to restrict the type of possible antecedents in terms of reference to the SUBJect; to set up a hierarchy for antecedenthood in which possible antecedents are ranked according to their associated grammatical function and thematic role; to unify morphological features checking for agreement in person and number, and selectional restrictions imposed by inherent semantic features; to tell apart quantifiers and quantified NPs which cannot be used as antecedents in backward pronominalization. A complete list of features is given below which we comment here briefly: all elements marked +ref can be treated as antecedents in the discourse; the second feature is definiteness which can take on three values ± and 0 which is introduced when no determiner precedes the noun or pronoun; the third feature is partitivity and is used to set quantified expressions apart from non quantified ones: +part is associated to all partitive quantifiers and quantified NPs, -part is used for universal quantifiers which in turn may be definite or indefinite. The fourth feature is ±pro and sets apart nominal expressions from
pronominal expressions; the fifth feature is \( \pm a_n \) and sets apart anaphoric elements which can be bound sentence internally, from non-anaphoric ones. Finally, the sixth feature varies from \( \pm c_l \) to \( \pm m_e \); now, where the one applies the other cannot apply. In particular \( \pm c_l \) is used to set apart common nouns which individuate a class from proper nouns which serve to name an individual; on the contrary, \( \pm m_e \) is only used for pronominal expressions and differentiates morphologically expressed pronouns from morphologically unexpressed ones. A seventh feature appears with anaphors to distinguish long distance anaphors which are subjective from the ones which are not.

1. Table of referential features and categories

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(+\text{ref}, \pm 0 \text{def}, \pm \text{part}, \pm \text{pro}, \pm \text{ana}, \pm \text{class}), 1</td>
<td>common noun</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 1</td>
<td>vbl</td>
</tr>
<tr>
<td>(+\text{ref}, \pm 0 \text{def}, \pm \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{class}), 0</td>
<td>proper noun</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 6</td>
<td>clitic</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 3</td>
<td>lexical pronoun</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 3</td>
<td>little pro</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}, \pm \text{subj}), 9</td>
<td>long reflexive</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 9</td>
<td>short reflexive</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 4</td>
<td>big PRO</td>
</tr>
<tr>
<td>(+\text{ref}, \pm 0 \text{def}, \pm \text{part}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 10</td>
<td>pro quantif</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 8</td>
<td>relative pro</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 8</td>
<td>interrogative pronoun</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}, \pm \text{subj}), 8</td>
<td>possessive anaphor</td>
</tr>
<tr>
<td>(+\text{ref}, \text{nil}, \text{nil}, \pm \text{pro}, \pm \text{ana}, \pm \text{me}), 7</td>
<td>possessive pronoun</td>
</tr>
</tbody>
</table>

2. THE ALGORITHM

To establish coreference for a pronominal expression, the main topic of discourse must be established. As for terminology, we use topic of discourse (see Bullwinkle 1977) rather than focus, an attribute which we keep for non-argument grammatical functions, which in line with LFG are derived by grammar rules at sentence level and are FOCUS TOPIC ADJunct MODifier. The algorithm for discourse analysis takes a sentence at a time and looks for antecedents to pronouns which have been assigned to an EXTERNAL referent by the binding algorithm at sentence level.

We proceed then, on the basis of the following principles:

8. there is only one possible TOPIC of DISCOURSE which may be assigned as the MAIN TOPIC for each sentence;

9. there may be another topic which is assigned as SECONDARY TOPIC and may be computed together with the MAIN for each sentence;

10. TOPICS may be changed according to the states of the discourse model: in this case we follow Brennan's (1986) approach to centering, even though we use a different set of states.

The algorithm starts by extracting a list of all referential expressions from the sentence under analysis; then, it proceeds by substituting all referentially bound NP heads with the head of their antecedents. Supposing now that we start from the first sentence of a text or discourse, we are now ready to establish the EXPECTED MAIN TOPIC of discourse from the referential expressions made available by this list. In order to do this, at first we look for specific structural information, like a Presentative sentence in which the SUBJECT is presented or introduced into the discourse by means of a specific structural configuration like a 'there' sentence or a locative inverted structure; in lack of this structure, we assign scores to the NPs included in the list and consequently choose from among the Weighted List the NP which has been assigned the highest score. To assign scores we proceed as follows:

11. we compute the distance of a given NP from the root of the graph of the sentence because we intend to privilege as possible TOPICS those NPs which are positioned at the higher levels;

12. we assign scores according to thematic or semantic roles and grammatical functions according to the following hierarchy:
As to ADJuncts, we also established a hierarchy since we take INSTRumental to be more relevant than MODal or TEMPoral ADJuncts. Note that AGent subsumes other semantic roles like POSSessor, SOUrc_e, INSTigator, etc.; GOAL subsumes ADDressee; THeme_Unaffected subsumes POSSEssion, INFORMATION, LOCation etc.

As to grammatical functions, the nuclear functions come first, i.e. SUBJECT OBJECT and OBJECT2, then we have OBLique and ADJuncts. Only referential arguments are imported at this level since non-referential ones cannot become the object of coreference (or co-specification, in Sidner's terms) in subsequent discourse. This subdivision is readily made available in LFG at the functional level, in particular because the theory distinguishes between open and closed functions (see Bresnan 1982).

13. After NPs have been weighted, we proceed by ordering them and then we sort NPs which have the same head erasing from the list occurrences of the same NP which have received a lower score. At this point, a semantic filter is applied, so that pruning deletes those referring expressions which have been assigned a score lower than a given threshold. The pruned list becomes the Weighted List, in which all referential expressions have been ordered according to their scores.

14. The Weighted List contains information as to the probability that a given NP may become the MAIN TOPIC of discourse. However, when computing the first sentence of discourse or text, only the EXPECTed Topic can be computed: the remaining NPs are assigned to OTHER Topics, a list from where a referential expression may be picked up in case the EXPECTed Topic is not reinforced. Two more slots are available as output of the computation: a SECONDARY Topic which is a TOPic scoring lower than the MAIN Topic, and has the highest probability to become a MAIN Topic in the following sentence in case a change of state in the main algorithm takes place.

15. After an EXPECTed Topic has been computed, a reinforcement is required in the form of a pronominal: only in this way a MAIN Topic may ensue.

Following Sidner's Expected Focus Algorithm (1983, 287), our algorithm chooses an expected topic which must be reinforced in the following sentence; however this topic may also be rejected. While the expected focus algorithm can always choose an expected focus, its choice may have to be rejected because the default position is overridden by other factors. Typically, this occurs when a pronoun, which does not co-specify with the expected focus, is used in the second sentence of the discourse, and no anaphors is used to co-specify with the expected focus. Whenever an anaphor is encountered, the current Main Topi_c, if already established, or else the Expected Topic is tested as co-referent with (or co-specifier for) the anaphoric expression. It has to satisfy syntactic agreement with the list of grammatical features, as well as selectional semantic restrictions represented both by inherent features, and thematic roles compatibility.

2.1 States of the Algorithm
The algorithm may assume four main states:

a. CONTINUING: indicating that all the TOPICS of the current sentence correspond to those already fixed for the previous sentence, and the individual/s referred to is/are exactly the same one/s;
b. RETAINING: this state precludes to a possible change of discourse TOPIC, when in the current sentence a new TOPIC has been introduced as SUBJ in the form of a lexical nominal expression;
c. SHIFTING: indicating an abrupt change of discourse TOPIC due to the spotting of a particular structural configuration, like an inverted subject or a 'there' sentence.
d. CHANGING: this is the state ensuing from a change of discourse TOPIC.

Minor states of the algorithm may arise with,

a.1 RESUME: whenever a Secondary Topic is resumed after a CHANGING has taken place and is introduced as Main Topic;
b.1 CHANGE_ANALYSE: whenever a deictic demonstrative pronoun or a nominal substitute is used which varies both the lexical head of the antecedent and establishes a set membership relation with it.

3. TOPIC SHIFT AND FOCUS
When a Main Topic is established new topics can be added or there can be a topic SHIFT. To add new topics it is sufficient to connect the subject of a sentence to a previously mentioned POTENTIAL TOPIC. Expected topics are all noun phrases used non-predicatively. A SHIFT or Topic Movement can take place only by means of overt syntactic structures, being instantiated by LFG non argument functions such as TOPIC and FOCUS for fronted constituents - i.e. for constituents which have been clefted, dislocated, extraposed, topicalized and so on. TOPIC and FOCUS in fact are discourse markers which can be used directly in our algorithm without any further elaboration. A FOCUS constituent is the trigger for TOPIC SHIFT and causes two things to happen: i. the previously established Main Topic is demoted to SECONDARY TOPIC, a role which will be explained below; ii. the FOCUS constituent is automatically raised to MAIN TOPIC without the need to wait for it to be reinforced.

The SECONDARY TOPIC is a repository for constituents which have been previously used as Main Topics and may be reintroduced in the following discourse. In case the shifted topic is not 'reinforced' in the following sentence and reference to the previous Main Topic is activated by means of SUBJECT noun phrase, the Secondary Topic may be restored to its previous role. The other important function that Secondary Topic fullfills in our system is the possibility of having more than one referential expression to corefer to, and bind, pronouns and anaphors. When this happens the Main and the Secondary Topic are the more plausible candidates. The stack containing Potential Topics is renewed with each new sentence analysed. Theoretically, the algorithm works according to the following,

**D. PRINCIPLES OF PRAGMATIC BINDING**

a. Only an OBJ ect can bind a TOPIC OF DISCOURSE in the following sentence;
b. Only a FOC us can be used to introduce a new TOPIC OF DISCOURSE;
c. Only SUBJECTs can bind the TOPIC OF DISCOURSE from the following sentence;
d. Only the TOPIC OF DISCOURSE can bind the SUBJECT in the following discourse;
e. Use a pronominal to bind.

We can use Bresnan's(1990) examples to show how these principles work:

16.i. I'm looking for my friend Rose.
   ii. *Among the guest of honour was sitting Rose.
   iii. Rose/She was sitting among the guest of honour.

17.i. Where's Rose?
   ii. *Among the guest of honour was sitting HER / *she/*her /*Rose
   iii. Rose/She was sitting among the guest of honour.

3.1 Neutralization

3.1.1 Deictic pronouns

A number of interesting phenomena can be covered by an adequate grammatical representation and they concern discourse bound referring elements like deictic pronouns, quantified expressions and nominal expressions with an indirectly governed function. Discourse bound pronouns can be divided up into different types (partially following Bosch 1981): anaphoric ones, 'which continue or sustain a previously established focus towards a specific item which he had oriented his attention to before'; deictic ones, 'which are a means for achieving the focussing... of the attention towards a specific item which is part of the respective deictic space'(ibid.,68).

The use of deictic pronouns is quite common in texts and ranges from first and second person personal subject pronouns, to deictic demonstrative pronouns like 'this' and 'that', which in Italian can be used to corefer both with a human referential antecedent and with a non referential sentential antecedent - differently from what happens in English (see Halliday and Hasan 1975, 63). Here are some examples taken from the story of 'The Three Little Pigs', made up of 25 sentences which we computed entirely. We discuss the first five sentences in their Italian version. Literal translation is provided also below. This is the English translation: '1.Once upon a time there were three brothers little pigs who lived happily in the countryside.2.But in the same place lived a wicked wolf who fed precisely on plump little pigs.3.The little pigs therefore decided to build a small house each, to protect themselves from the wolf.4.The oldest one, Jimmy, who was wise worked hard and built his house with solid bricks and cement.5.The others, Timmy and Tommy, who were lazy and idle settled the matter
hastily and built their house with hay and little pieces of wood. We show the f-structure as output of the binding module at sentence level, and then the output of the discourse module.

18. [ci, erano, una, volta, tre, fratelli, porcellini, che, vivevano, felici, nella, campagna] There were one time three piglets who lived happyplural/mas in the countryside

net(pos1)
index: f18
pred: essere
lex_form: [np/subj/thema_bound/[umanoi _]]
mood: ind
tense: pres
cat: esistenza
subj_foc/theme_bound: index: sn272
cat: [umanoi, animato, relazionale]
pred: fratello
gen: mas
num: plur
spec: def: 0
part: +
card: tre
mods: mod/predic: topic/topic_type:relative
index: sn274
cat: [umanoi, animato, relazionale]
pred: pro
pers: mas
gen: plur
num: 3
case: [nom]
binder: sn272
tab_ref: [+ ref, + pro, + ana, - me]
anteecedent: sn272
interpretation: specific
index: f21
pred: vivere
lex_form: [np/subj/actor/[animatol _], pp/obl/locative/in/[luogo]]
mood: ind
tense: imp
cat: esistenza
subj/actor: index: sn275
cat: [animalo _]
pred: pro
pers: 3
gen: mas
num: plur
case: [nom]
spec: def: +
binder: sn274
tab_ref: [+ ref, + pro, - ana, - me]
anteecedent: sn274
interpretation: specific
obl/locative: index: sn298
cat: [luogo]
pred: campagna
gen: fem
num: sing
spec: def: +
tab_ref: [+ ref, - pro, - ana, + class]
adj: s: adj/emozione/index: ssa3
cat: [emozione]
pred: felice
gen: mas
num: plur
subj/nlcsn/index: sn299
pred: vbl
binder: sn275
adj: s: noun_mod/index: sn273
cat: [edibile, animato]
pred: porcellino
and this is the output of the discourse module,

19. [nello,stesso,luogo,pero,,viveva,anche,un,terribile,lupo,che,si,nutriva,proprio,di,porcellini,grassi]In the same place however lived also a terrible wolf who himself feasted just of piglets plump_plur/mas

net(po2)

pred:vivere

index:sn82

pred:lupo

gen:mas

num:sing

spec:def:-

foc:anche

mood:ind

tense:imp

subj_foc/actor:index:sn82

cat:[ferocious, animato]

pred:lupo

gen:mas

num:sing

spec:def:-

foc:anche

mood:cat:[ferocious, anche]

pred:terribile

gen:mas

num:sing

adjs:adj/predic:topic:topic_type:relat

index:sn83

cat:[ferocious, animato]

pred:pro

pers:mas

gen:sing

num:3

case:nom

binder:sn82

tab_ref: [+ ref, + pro, + ana, + me]

antecedent:sn82

interpretation:specific

index:i13

pred:nutrire
and this is the output of the discourse module,

ref_ex(sn82, lupo, [+ ref, - def, nil, - pro, - ana, + class], nil, mas, sing, [animate,ferocious], subj_foc/actor)/3
ref_ex(sn81, luogo, [+ ref, + def, nil, - pro, - ana, + class], nil, mas, sing, [place], obl/locative)/28
ref_ex(sn303, porcellino, [+ ref, 0 def, nil, - pro, - ana, + class], nil, mas, plur, [animate,edible], obl/food)/60

MAIN TOPIC : ref_ex(_, lupo, _ nil, mas, sing, [animate,ferocious], _ _)  
SECONDARY TOPIC : ref_ex(_, porcellino, _ nil, mas, plur, [animate,edible], _ _)  
POTENTIAL TOPICS : ref_ex(sn81, luogo, [+ ref, + def, nil, - pro, - ana, + class], nil, mas, sing, [place], obl/locative) state(2, shifting)

Sentence three reintroduces both participants in the story with different roles and in different ways: the little pigs are reintroduced as SUBject, thus receiving a high score, by means of a demonstrative pronoun; on the contrary, the wolf is mentioned in an adjunct clause and receives a low score. As all deictic pronouns, the demonstrative
is computed both as a pronoun and as a nominal head, thus receiving the feature +class. In this case all functional features match.

20. [questi, allora, 'per, proteggersi, dal, lupo', 'decidero, di, costruirsi, ciascuno, 'na, cassetta] these_plur/mas then to protect themselves from the wolf, decided to build themselves each sing/mas one houselet

net_po3
index:f9
pred: decidere
lex_form: [np/subj/actor/[animato, umano], vpinf/vcomp/prop/di/[subj=subj/actor]]
mood: ind
tense: pass_rem
cat: soggettivo
subj/actor: index: sn13
  cat: [animato, umano]
  pred: questi
  pers: 3
  gen: mas
  num: plur
  case: [nom, acc]
  spec: def: +
  tab_ref: [+ ref, + pro, - ana, + class]
  antecedent: external
  interpretation: specific
vcomp/prop: index: finf2
pred: costruire
  lex_form: [np/subj/agent/[animato, umano], np/obj/theme_aff/[oggetto, luogo]]
mood: inf
tense: pres
cat: cambiamento
subj/agent: index: sn23
  cat: [animato, umano]
  pred: pPro
  binder: sn13
  tab_ref: [+ ref, + pro, + ana, - me]
  interpretation: specific
obj/theme_aff: index: sn21
  cat: [oggetto, luogo]
  pred: casa
  gen: fem
  num: sing
  spec: def: -
subj/poss: index: sn22
  cat: [oggetto, luogo]
  pred: pPro
case: gen
  spec: def: +
  binder: sn20
  tab_ref: [+ ref, + pro, + ana, - me]
  interpretation: specific
mods: cat: [oggetto, _]
  pred: piccolo
  gen: fem
  num: sing
  tab_ref: [+ ref, - pro, - ana, + class]
obj2/benefit: index: sn20
  cat: [umano, animato]
  pred: si
  pers: 3
  case: [dat]
  spec: def: +
  tab_ref: [- ref, - pro, + ana, + me, - subj]
  antecedent: sn23
  interpretation: specific
adjs: adj/quantitative: index: saa16
  cat: [quantif]
  pred: ciascuno
  gen: mas
  num: plur
  spec: def: +
and this is the output of the discourse module,

ref_ex(sn13, questii, [+ ref, - def, nil, + pro, - ana, + class], mas, plur, [animato, umano], subj/actor)/23
ref_ex(sn21, casa, [+ ref, - def, nil, - pro, - ana, + class], nil, fem, sing, [object/place], obj/theme_aff)/35
ref_ex(sn17, lupo, [+ ref, + def, nil, - pro, - ana, + class], nil, mas, sing, [animate,ferocious], obl/agent)/40

sn43 = porcellino

EXPECTED TOPIC : ref_ex(_, porcellino, - nil, mas, plur, [animate,edible], _/
SECONDARY TOPIC : ref_ex(_, lupo, - nil, mas, sing, [animate,ferocious], _/

POTENTIAL TOPICS : ref_ex(sn21, casa, [+ ref, - def, nil, - pro, - ana, + class], nil, fem, sing, [object/place], obj/theme_aff)
state(3, change)

However, in the following sentence, the nominal substitute ‘the oldest one’ no longer agrees in number with the antecedent, the set of three little pigs, only pointing to a subset made up of a singleton. In order to compute this difference, we simply neutralized the feature Number and let all other features match as before. In addition, since the individual being referred is not strictly speaking equal to the antecedent contained in the Main Topic, we use continue_analaysia.

21.[il,maggiore,'',jimmi,'',che,era,saggio,'',lavorava,di,buona,lena,e,costrui,la_sua,casetta,con,solidi,mattoni,e, cemento]/ The_sing/mas oldest_sing, jimmi, who was wise_sing/mas, worked of good_sing/fem will and built his_sing/fem houselet with solid_plur/mas bricks_plur/mas and cement_sing/mas
net(po4)
main/prop/index:f10
coord/index:f6
pred:lavorare
lex_form:[subj/agent/[umano, animato]]
mood: ind
tense: imp
cat: attivita
subj: agent: index: s1
cat: {umano, animato}
pred: maggiore
pers: 3
gen: mas
num: sing
case: [nom]
spec: def
adjs: adj/ theme: index: s2
cat: {animato}
pred: jinni
gen: mas
num: sing
spec: def
modal/predicate: topic: topic_type: relative
index: s3
cat: {animato}
pred: pPro
pers: mas
gen: sing
num: 3
case: [nom]
binder: s2
tab_ref: [+ ref, + pro, + ana, - me]
antecedent: s2
interpretation: specific
index: f4
pred: essere
lex_form: [np/subj/theme_bound/[_, _]]
mood: ind
tense: imp
cat: esistenza
subj: theme_bound: index: s4
cat: {animato, umano}
pred: pro
pers: 3
gen: mas
num: sing
case: [nom]
binder: s3	tab_ref: [+ ref, + pro, - ana, - me]
antecedent: s3
interpretation: specific
acomp/prop: index: saa2
cat: soggettivo
pred: saggio
gen: mas
num: sing
subj: prop: index: s5
pred: vbl
binder: s4
aspect: state
tab_ref: [+ ref, - pro, - ana, - class]
tab_ref: [+ ref, + pro, - ana, + class]
antecedent: external
interpretation: specific
adjs: adj/modal: sem_mark: buona_lena
obj/nil: index: s6
cat: {modo}
pred: lena
gen: fem
num: sing
spec: def
.tab_ref: [+ ref, - pro, - ana, + class]
aspect: activity
coord: index: f9
pred: costruire
and this is the output of the discourse module,
sn1 = porcellino

MAIN TOPIC : ref_ex(_, porcellino, _, 3, mas, sing, [animate, edible], _/agente)
POTENTIAL TOPICS : ref_ex(sn38, casa, [+ ref, + def, nil, - pro, - ana, + class], nil, fem, sing, [object, place], obj
/aff_theme)

ref_ex(sn2, jinumi, [+ ref, 0 def, nil, - pro, - ana, - class], nil, mas, sing, [animate, human], nadj/theme)
ref_ex(sn6, lena, [+ ref, 0 def, nil, - pro, - ana, + class], nil, fem, sing, [manner], adj/modal)
state(4, continue_analyse)

In this case, the nominal substitute 'other ones' has a plural Number but the same set is coreferred, extracting though a different subset from the previous one. We used the same strategy of Neutralization in order to let the Main Topic continue, and the state is CONTINUE_ANALYSE as before:

22. [gli, altri, 'i, timmy, e, tommy, 'i; pigri, e, oziosi, se, la, sbirgarono, in, frettà, 'i, costruendo, le, loro, cassette, con, la, paglia, e, con, pezzetti, di, legno]/The_plur/mas others_plur/mas, Timmy and Tommy, lazy_plur/mas and idle itself it_sing/fem dealt in hurry, building the_plur/fem their_plur houselets with the straw and with little pieces of wood net(po5)

index:f3
pred:sbrigarsi

lex_form:[np/subj/agent/[animato, umano], np/obj_theme/valutativo[/_ _]]
mood:ind
tense:pass_rem
cat:risultato

subj/agent:index:sn187

cat:[umano, animato]
pred:altri
pers:3
gen:mas
num:plur
case:[nom, acc]
spec:def+
mods:index:sa3
cat:[valutativo]
coordin:head:coordinant

gen:mas
num:plur
coord:index:sa1
cat:[valutativo]
pred:pigri

gen:mas
num:plur
coord:index:sa2
cat:[valutativo]
pred:oziosi

gen:mas
num:plur
adj:sn190

adj/the:main/index:sn188

cat:[umano, animato]
pred:timmy

gen:mas
num:sing
spec:def0

tab_ref:[+ ref, - pro, - ana, - class]

adj/the:indice:sn189

cat:[umano, animato]
pred:tommy

gen:mas
num:sing
spec:def0

tab_ref:[+ ref, - pro, - ana, - class]

antecedent:external
interpretation: specific

obj_avv/theme_nonaff:index:sn263
  cat:[]
  pred:la
  pers:3
  gen:fem
  num:sing
  case:[acc]
  tab_ref:[+ ref, + pro, + ana, + me]
  antecedent:external
  interpretation: specific

adjx:adj/modal:sem_mark:in
obj/nil:index:sn199
  cat:[modo]
  pred:freuta
  gen:fem
  num:sing
  spec:def:0
  tab_ref:[+ ref, - pro, - ana, + class]

adj:indice:figurand15
pred:costruire
  subcat:[np/subj/agent/[animato, umano], np/obj/theme_aff/[oggetto, luogo]]
  mod:ger
  tempo:pres
  cat:cambiamento
  subj/agent:index:sn419
    cat:[animato, umano]
    pred:pPro
    tab_ref:[+ ref, + pro, + ana, - me]
    antecedent:sn187
    interpretation: specific

obj/theme_aff:index:sn228
  cat:[oggetto, luogo]
  pred:cosa
  gen:fem
  num:plur
  spec:def:+
  subj/poss:index:sn347
    cat:[]
    pred:loro
    gen:85173
    num:85270
    spec:def:+
    tab_ref:[+ ref, + pro, - ana, + me]
    antecedent:
    interpretation: specific

modal:cat:[oggetto, _]
  pred:piccolo
  gen:fem
  num:plur
  tab_ref:[+ ref, - pro, - ana, + class]

adjx:adj/strumentale:sem_mark:con
  index:sn356
  cat:[oggetto, strumento]
  pred:paglia
  gen:fem
  num:sing
  spec:def:+
  tab_ref:[+ ref, - pro, - ana, + class]

adj/strumentale:sem_mark:con
  index:sn413
  cat:[oggetto, strumento]
  pred:pezzo
  gen:mas
  num:plur
  spec:def:0
  obj/theme:index:sn418
    cat:[oggetto, strumento]
    pred:legno
    gen:mas
18

num: sing
spec: def: 0
  tab_ref[+ ref, - pro, - ana, + class]
mods: cat: [oggetto]
  pred: piccolo
gen: mas
num: plur
  tab_ref[+ ref, - pro, - ana, + class]
aspect: accomp
aspect: achiev_tr
and this is the output of the discourse module,
ref_ex(sn187, altri, [+ ref, + def, nil, + pro, - ana, + class], 3, mas, plur, [animate, human, object1 _], subj/agent)/10
ref_ex(sn228, casa, [+ ref, + def, nil, - pro, - ana, + class], nil, fem, plur, [object, place], obj/theme_aff)/35
ref_ex(sn188, timmy, [+ ref, 0 def, nil, - pro, - ana, - class], nil, mas, sing, [animate, human], nadj/theme)/1050
ref_ex(sn189, Tommy, [+ ref, 0 def, nil, - pro, - ana, - class], nil, mas, sing, [animate, human], nadj/theme)/1050
ref_ex(sn199, fredda, [+ ref, 0 def, nil, - pro, - ana, + class], nil, fem, sing, [manner], adj/modal)/1070
sn522 = porcellino
MAIN TOPIC : ref_ex(_, porcellino, _, 3, mas, plur, [animate, edible], /agent)

POTENTIAL TOPICS : ref_ex(sn228, casa, [+ ref, + nil, - pro, - ana, + class], nil, fem, plur, [object, place], obj/aff_theme)
ref_ex(sn188, timmy, [+ ref, 0 def, nil, - pro, - ana, - class], nil, mas, sing, [animate, human], nadj/theme)
ref_ex(sn189, Tommy, [+ ref, 0 def, nil, - pro, - ana, - class], nil, mas, sing, [animate, human], nadj/theme)
ref_ex(sn199, fredda, [+ ref, 0 def, nil, - pro, - ana, + class], nil, fem, sing, [manner], adj/modal)
state(5, continue analyse)

Another interesting case of Neutralization is constituted by the use of Direct Speech and first and second person personal forms. This happens suddenly in the text: the algorithm should behave in such a way as to let the Main Topic and or Secondary Topic to continue, and this is what it does, by neutralizing the feature Person and letting the remaining features match. These are the three related sentences, and their translation: 15. Frightened out of their wits, the two little pigs ran at breakneck speed towards their brother's house. 16. Fast, little brother, open the door! The wolf is chasing us.' They got in just in time and pulled the bolt.
23. [spaventatissimi, i, due, porcellini, corsero, a, perdisato, verso, la, casetta, del, fratello] /Frightened at their most plur
/mas the two piglets ran at breathless (speed) towards the houselet of the brother
net(1015)
index: 9
pred: correre
lex_form: [subj/agent/ humano, animato]
mood: ind
tense: pass_rem
cat: attivita
subj/agent: index: sn10
  cat: [editable, animato]
  pred: porcellino
gen: mas
num: plur
spec: def: +
  part: +
card: due
  tab_ref[+ ref, - pro, - ana, + class]
adj: adj/modal: sem_mark:a
  obj/nl: index: sn83
cat: [modo]
  pred: perdisato
gen: mas
num: sing
spec: def: 0
  tab_ref[+ ref, - pro, - ana, + class]
adj/loca: sem_mark: verso
  obj/nl: index: sn138
cat: [oggetto, luogo]
pred: casa
gen: fem
24. ['presto', 'fratellino', 'aprici, il, abbiamo, il, lupo, alle, calcagna,']/Quick, brother + little, open_imper/ 2nd_pers+ us_plur! (We) pro have_present_plur/1st_pers the wolf at our heels. net(po16)
index:f1
dir_speech/prop:index:f4
pred/are vere
lex_form:[np/subj/experiencer/[umanos animato], np/obj/theme_bound/[l _], pcomp/locative/[luogo]]
mood:ind
tense:pres
cat:stato
subj/experiencer:index:sn4
cat: [umanos animato]
pred: pro
pers:1
gen:mas
num: plur
case:[nom]
spcc: def+
tab_ref:[+ ref, pro, - ana, - me]
precedent: external
interpretation: specific
obj/theme_bound:index:sn5
cat: [ferocious, animato]
pred: lupo
gen: mas
num: sing
spcc: def+.
and this is the output of the discourse module,

1. ref_ex(sn4, pro, [+ ref, + nil, + pro, - ana, - me], 1, mas, plur, [human, animate], subj/experiencer)20
2. ref_ex(sn5, lupo, [+ ref, + nil, - pro, - ana, + class], nil, mas, sing, [animate, ferocious], obj/theme_nonaff)25
3. ref_ex(sn16, calcagna, [+ ref, + nil, - pro, - ana, + class], nil, fem, plur, [object, place], pcomp/locative)38
4. ref_ex(sn2, fratello, [+ ref, 0, nil, - pro, - ana, + class], nil, mas, sing, [human, animate, relational], subj/experiencer)80

sn4 = porcellino

MAIN TOPIC : ref_ex(_, porcellino, _, _, mas, plur, [human, animate], subj/experiencer)

SECONDARY TOPIC : ref_ex(_, lupo, _, nil, mas, sing, [animate, ferocious], obj/theme_nonaff)

POTENTIAL TOPICS : ref_ex(sn16, calcagna, [+ ref, + nil, - pro, - ana, + class], nil, fem, plur, [object, place], pcomp/locative)
As can be noticed, personal forms should be made visible to the sentence binding module so that the clitic 'ci/us' could be adequately bound by the little pro 1st person pronoun of the following sentence before reaching the discourse module. Also, the bound subject of the imperative should be bound to the vocative 'fratellino' at sentence level. In the following sentence, 3rd person plural verbal agreement is used to continue the story.

25. [l'ecorso, appena, in, tempo, ad, entrare, e, a, tirare, il, chiaviestrello](They) pro made_past/plur/3rd_pers just in time to enter and to pull the bolt.

net(po17)
index:f3
pred:fare
lex_form:{np/subj/agent/[[umano, animato], pp/obl/temporal/in/[tempo]]}
mood:ind
tense:pass_rem
cat:risultato
subj/agent:index:sn1
   cat:[umano, animato]
   pred:pro
pers:3
gen:mas
num:plur
case:[nom]
spec:def+
tab_ref:[+ref, +pro, -ana, -me]
ancestors:sn1
interpretation:specific
obl/temporal:index:sn30
cat:[tempo]
pred:tempo
gen:mas
num:sing
spec:def:0
vcomp/prop:index:finf23
coord:index:finf24
pred:tirare
lex_form:{subj/agent/[[umano, animato], obj/theme AFF/[[oggetto, strumento]]}
mood:inf
tense:pres
cat:risultato
subj/agent/index:sn34
   cat:[umano, animato]
pred:pro
   tab_ref:[+ref, +pro, +ana, -me]
   ancestors:sn1
   interpretation:specific
   obj/theme AFF/index:sn32
      cat:[oggetto, strumento]
pred:chiaviestrello
gen:mas
num:sing
spec:def+
tab_ref:[+ref, -pro, -ana, +class]
aspect:achiev_tr
adjs:avv:pred_avv:appena
type:temp
duraz:punt
config:[tr<ctd]
aspect:achiev_tr

and this is the output of the discourse module.

ref_ex(sn1, pro, [+ ref, + nil, + pro, - ana, - mel], 3, mas, plur, [human, animate], subj/agent)/10
ref_ex(sn32, chiaviestrello, [+ ref, + nil, - pro, - ana, + class], nil, mas, sing, [object, instrument], obj/theme AFF)/1035
ref_ex(sn30, tempo, [+ ref, 0, nil, - pro, - ana, + class], nil, mas, sing, [time], adj/temporal)/1070
sn5 = porcellino
MAIN TOPIC : ref_ex(_, porcellino, _, _, mas, plur, [human, animate], subj/agent)
POTENTIAL TOPICS : ref_ex(_, lupu, _, nil, mas, sing, [animate, ferocious], obj/theme_bound)
ref_ex(sn32, chiavistello, [+ ref, + nil, - pro, - ana, + class], nil, mas, sing, [object, instrument], obj/aff_theme)
ref_ex(sn30, tempo, [+ ref, 0, nil, - pro, - ana, + class], nil, mas, sing, [time], adj/temporal)
state(17, continue)

4. NOMINAL EXPRESSIONS

4.1 Rejecting the Expected Topic and Activating Inferencing Mechanisms

As we said previously, we let the algorithm activate inferential mechanisms only when a state of CHANGE is detected in the previous sentence or whenever a nominal expression is used to cospecify a Main Topic while CHANGE is detected. No such triggering takes place for SHIFTING from a given Main Topic; nor for RETAINING, i.e. whenever a new topic is introduced as an indefinite expression or a proper noun. In a state of CONTINUE inferencing may be activated as long as the grammatical representation allows it: in other words, the nominal expression must be computed as an indirectly governed function by the parser: these are SUBJ or OBJ functions which are not directly governed by the main predicate of a sentence, but by a predicate function, such as the ACOMP contained in the lexical forms of verbs such as 'BELIEVE<SUBJ,ACOMP>OBJ>', or 'BE<ACOMP>SUBJ'. We compute the difference in government by associating a special semantic role, THEME_BOUND to the function.

The result of scoring is used to set up adequate conditions for triggering inferential processes both in presence of a pronouns and a nominal expressions. However, the behaviour of pronouns is only determined after grammatical constraints are satisfied at sentence level. As to what triggers an inference to be drawn Ehrlich(1981) states clearly the point: at first a relation between expressions must somehow be perceived before an inference is drawn, and this relation is clearly syntactic and semantic in nature. As she comments 'people do not draw inferences randomly to relate linguistic expressions', showing how in two examples people related 'bus' and 'vehicle' only when certain conditions would require it; the examples she uses are the following:

26i. A bus came roaring round the corner.
   ii. The vehicle nearly flattened a pedestrian.

27i. A bus came roaring round the corner.
   ii. It nearly smashed one vehicle.

In the first example an entity 'bus' is introduced in the discourse as the Topic and then reinforced in the following sentence using a class noun 'vehicle' which subsumes the reference of 'bus'. Criteria for relating the two referring expressions are the use of SUBJ function in the second sentence, together with definiteness, and inferencing. In example 17, once the pronoun 'it' is processed as anaphoric to the Expected Topic, it must be obviative with any other referring expression contained within the same f-structure, the sentence. In our system triggering material may derive both from sentence level analysis and from discourse level one. In case the Expected Topic is rejected as possible coreferent of the pronoun and another phrase is chosen, the rejected phrase is retained for possible re-introduction later in discourse. Rejection is possible when compatibility requirements and/or agreement tests are not met. In our system, whenever a common noun is used as SUBJECT of a sentence instead of a pronoun and it is different from the Main Topic it triggers rules for FILTERING and inference checking. In particular a subject which is a Bound Theme, as the subject of a copulative sentence, cannot be computed as an Expected Topic and is filtered. Inferencing is limited to a table lookup procedure for the encyclopaedic information associated to a certain lexical entry, in other words an IS_A relation.

Another case in which the algorithm calls for inference check is the presence of more than one candidate satisfying semantic and syntactic requirements to be picked up as antecedent of a pronoun. Our approach is based on the premises that once a possible Main Topic is chosen as antecedent of a pronoun, in case some incompatibility or some ambiguity exist, inferring processes are called for to confirm or reject an Expected Topic.

We shall now discuss a text we have analysed, which, differently from the examples found in the literature is taken directly from a newspaper. It deals with politics and there are three main topics: Avveduti who has been appointed secretary general by his father-in-law who is a senator and Trabucchi, a minister of trade. In the first text, we have an introductory sentence and a continuation in which a pronoun and a definite NP is used to corefer
to the previous Topics. Inferences must be drawn in order to establish both the antecedent for the pronoun and for the definite NP. However, in the case of the pronoun, in order to be able to trigger the inferential device, a nominal head must be made available and this is done via the grammatical representation. In particular, a predicate like APPONIT has an NP OBJECT which has Bound Theme as semantic role, to indicate that the OBJ is non thematic or indirectly governed, as shown by the complete lexical entry, 'APPOINT, trans:achievement, soc_institution,[SUBJ/Ag, OBJ/Th_bound, NCOMP/Prop]'. The open function is an open proposition which however makes available a nominal head to the OBJECT pronoun, and this nominal head in turn is used to trigger the inference. In the third sentence the continuation is computed only by means of pronominal expressions which are graded independently according to scoring and the adequate antecedents are thus picked up automatically.

TEXT N. 1

28. [a, avveduti, piaceva, parlare, del, suocero] / To Avveduti liked past/sing/3rd_pers talking about the father-in-law.
- ref_ex(sn8, avveduti, [ref, 0def, nil, - pro, - ana, - class], nil, mas, sing, [human], obj2/experiencer)/10
- ref_ex(sn9, suocero, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [human, relational], obl/disc_subj)/31

EXPECTED TOPIC : ref_ex(sn8, avveduti, [+ ref, 0 def, nil, - pro, - ana, - class], nil, mas, sing, [human], obj2/experiencer)
POTENTIAL TOPICS : ref_ex(sn9, suocero, [+ ref, + def, nil, - pro, - ana, + class], nil, mas, sing, [human, relational], obl/disc_subj)
state(1, change)

29. [il, senatore, lo, aveva, nominato, segretario, particolare] / The senator had appointed him secretary general
- ref_ex(sn17, senatore, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_high], subj/agent)/10
- ref_ex(sn19, lo, [+ ref, -def, nil, + pro, + ana, + me], 3, mas, sing, [human], obj/theme_bound)/15
- ref_ex(sn20, segretario, [+ ref, 0def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_low], ncomp/prop)/50

is_a(aviveduti, segretario)/is_a(suocero, senatore); sn17 = suocero; sn19 = avveduti

MAIN TOPIC : ref_ex(_, suocero, _, nil, mas, sing, [human, soc_high], subj/agent)
SECONDARY TOPIC : ref_ex(_, avveduti, _, 3, mas, sing, [human], obj/theme_bound)
state(2, continue)

30. [era, un, uomo, che, chiunque, avrebbe, sfruttato, ma, che, lui, preferiva, lasciare, perdere] / (He) pro was past/sing/3rd_pers a man that anyone had past/sing/3rd_pers exploited but that him(self) preferred to let go.
- ref_ex(sn5, pro, [+ ref, -def, nil, + pro, - ana, - me], 3, mas, sing, [], subj/theme_bound)/40
- ref_ex(sn13, lui, [+ ref, -def, nil, + pro, - ana, + me], 3, nil, sing, [human], subj/actor)/43
- ref_ex(sn10, pPro, [+ ref, nil, nil, - pro, - ana, - me], nil, _, _, [event, state, human], obj/theme_aff)/45
- ref_ex(sn6, uomo, [+ ref, -def, nil, - pro, - ana, + class], nil, mas, sing, [human], ncomp/prop)/50

sn5 = suocero; sn13 = avveduti

MAIN TOPIC : ref_ex(_, suocero, _, nil, mas, sing, [human, soc_high], subj/agent)
SECONDARY TOPIC : ref_ex(_, avveduti, _, 3, mas, sing, [human], obj/theme_bound)
POTENTIAL TOPICS : ref_ex(sn10, pPro, [+ ref, nil, nil, - pro, - ana, - me], nil, _, _, [event, state, human], obj/tema_aff); ref_ex(sn6, uomo, [+ ref, -def, nil, - pro, - ana, + class], nil, mas, sing, [human], ncomp/prop)
state(3, continue)

In the following we show how, by varying the second sentence of the text, different inferences are triggered. Texts are limited to the second sentence which determine strictly what should happen in the third sentence. In text 2, we see how the appearance of an indefinite NP causes the algorithm to establish a new expected topic and RETAINing is the new state rather than CONTINUE. In particular, then only one inference is triggered always by means of the grammatical representation. In text 3 a proper noun is used as SUBJECT and is turned into a Main Topic: inferences are required for the pronoun.

TEXT 2:
31. [un, senatore, lo, aveva, nominato, segretario, particolare] A senator him had_past/sing/3rd_pers appointed secretary general

ref_ex(sn17, senator, [+ ref, -def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_high], subj/agent)/10
ref_ex(sn19, lo, [+ ref, -def, nil, + pro, + ana, + me], 3, mas, sing, [human], obj/theme_bound)/15
ref_ex(sn20, segretario, [+ ref, 0def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_low], ncomp/prop)/50
is_a(argveduti, segretario); sn19 = argveduti

EXPECTED TOPIC: ref_ex(, senatore, _, nil, mas, sing, [human, soc_high], subj/agent)
SECONDARY TOPIC: ref_ex(, argveduti, _, 3, mas, sing, [human], obj/theme_bound)
state(2, retaining)

TEXT 3.

32. [trabucchi, lo, aveva, nominato, segretario, particolare] Trabucchi him had_past/sing/3rd_pers appointed secretary general

ref_ex(sn43, trabucchi, [+ ref, 0def, nil, - pro, - ana, - class], nil, mas, sing, [human], subj/agent)/10
ref_ex(sn45, lo, [+ ref, -def, nil, + pro, + ana, + me], 3, mas, sing, [human], obj/theme_bound)/15
ref_ex(sn46, segretario, [+ ref, 0def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_low], ncomp/prop)/50
is_a(argveduti, segretario); sn45 = argveduti

MAIN TOPIC: ref_ex(, trabucchi, _, nil, mas, sing, [human], _/agent)
SECONDARY TOPIC: ref_ex(, argveduti, [+ ref, 0def, nil, - pro, - ana, - class], nil, mas, sing, [human], _/)
state(2, continue)

The following texts are in a condition of complete ambiguity but the inferential mechanism may still work as long as the grammatical representation allows it to recover the nominal head which is an open complement where the object pronoun is interpreted.

TEXT 4.

33. [lo, aveva, nominato, segretario, particolare] / (he) pro him had_past/sing/3rd_pers appointed secretary general

ref_ex(sn37, pro, [+ ref, +def, nil, + pro, - ana, - me], 3, _11419, sing, [human], subj/agent)/10
ref_ex(sn39, lo, [+ ref, +def, nil, + pro, + ana, + me], 3, mas, sing, [human], obj/theme_bound)/15
ref_ex(sn40, segretario, [+ ref, 0def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_low], ncomp/prop)/50
is_a(argveduti, segretario); sn37 = suocero / sn39 = argveduti

MAIN TOPIC: ref_ex(, suocero, _, 3, mas, sing, [human], subj/agent)
SECONDARY TOPIC: ref_ex(, argveduti, _, 3, mas, sing, [human], obj/theme_bound)
state(2, continue)

TEXT 5.

34. [nel, 1950, lo, avevano, nominato, senatore] In 1950 they pro him had_past/plur/3rd_pers appointed senator

ref_ex(sn43, pro, [+ ref, +def, nil, + pro, - ana, - me], 3, mas, plur, [human], subj/agent)/10
ref_ex(sn45, lo, [+ ref, +def, nil, + pro, + ana, + me], 3, mas, sing, [human], obj/theme_bound)/15
ref_ex(sn46, senatore, [+ ref, 0def, nil, - pro, - ana, + class], nil, mas, sing, [human, soc_high], ncomp/prop)/50
ref_ex(sn40, 1950, [+ ref, +def, nil, - pro, - ana, - class], nil, mas, sing, [time], adj/temporale)/1050
is_a(argveduti, senatore) failed; is_a(suocero, senatore) / sn43 = indefinite / sn45 = suocero

MAIN TOPIC: ref_ex(, suocero, _, 3, mas, sing, [unano], _/theme_bound)
SECONDARY TOPIC: ref_ex(, argveduti, [+ ref, 0def, nil, - pro, - ana, - class], nil, mas, sing, _/

POTENTIAL TOPICS: ref_ex(sn40, 1950, [+ ref, +def, nil, - pro, - ana, - class], nil, mas, sing, [time], adj/temporal)
state(2, continue)

5. Quantifiers

It is a well known fact that quantifiers and quantified NPs do not refer in the text or discourse, in the sense that they are unable to pick up a specific individual as antecedent to which they may corefer. However, when computing reference quantifiers either lexically expressed or unexpressed may be used by speakers to continue the
topic of discourse. In our referential system, ROSIE (see Delmonte & Bianchi 1991a, 1991b), both unexpressed and expressed quantified expressions are computed in the same way: unexpressed quantifiers, like big PROs or little pros in sentences with arbitrary or generic interpretation lack in some or all Φ-features (number, gender, person).

5.1 Quantifiers and quantified NPs as antecedents

In her works, Webber (1977, 1983) extensively deals with the problem of the interpretation of quantified expressions. Quantifier scope requires a separate level of representation, logical form, which builds on grammatical representation. In other words, it would seem that once f-structures have been built, this level of representation must be turned into two separate levels: one required for anaphoric binding sentence internally, the other relevant to quantifier scoping (see Delmonte, 1989). However, at discourse level and from a procedural point of view, problems may arise only when a singular indefinite quantified expressions is referred to by a plural pronoun in a following sentence. In other words, whenever a plural pronoun looks for antecedents and the previous sentence contains quantified expressions, the availability of the logical form for the sentence is crucial. In our paper (Bianchi, Delmonte, 1989a) we discuss a modified version of Hobbs and Shieber’s algorithm for scope assignment to quantified expressions. In order to decide whether a singular indefinite expression can be treated as a plural one scope must be computed: in particular, its scope must be included in that of a universal quantifier with a distributive reading. We shall quote one of her examples (1983, 363-d25):

35.1 Last week Wendy bought each boy a green T-shirt at Macy’s.

2 She prefers them in more subdued colors, but these were on sale.

This is a typical example which requires at first scope assignment to be computed for the two quantified NPs, "each boy" and a "green shirt". As Webber remarks, a definite plural anaphor may also specify a generic set entity and this is possible even with a singular definite noun phrase as antecedent. The only condition seems to be procedural and based on "recency": quoting from Webber, ". . . the listener can generate new generic-set entities whose IDs are based on generalizations of a recent description the listener has either heard or derived." the only restriction constituted by the fact that these generalizations must somehow be shared by the speaker. We might add that such a generalization is reached through the interpretation process: a generic reference to a definite description is interpreted as such if it is not referential. In other words, there are strong restrictions to interpret an assertion as a generic statement, and they mainly concern the interpretation of tense and its modifiers. In order to compute the reference of a definite NP as generic, tense cannot be definite and referential, and adverbial modifiers cannot be deictic. This is clear if we look at some of Webber's examples (hers d22 through d28), as for instance (the underlining is mine):

36.1 Last week Wendy again bought each boy a green T-shirt at Macy’s.

2 She’s always buying them.

37.1 I see seven Japanese cars in the parking lot.

2 They’re really selling like hot cakes.

38.1 Last week Wendy bought each boy a green T-shirt at Macy’s.

2 She gave them to everyone.

39.1 Wendy bought some T-shirt yesterday.

2 Usually she charged them, but yesterday she paid cash.

40.1 Wendy wouldn’t buy a green T-shirt, because they always run in the wash.

All the items underlined are either the main verb or the adverbial modifier: tense is simple present, or progressive, and adverbials are "always, usually, really". The intended meaning conveyed by the sentences in 2. is iteration on events, the events are those of buying, selling, charging, giving. In other words it is quantification on events introduced by tense and adverbial modifiers which acts on propositions, just like quantifiers and determiners act on NPs (see Hinrichs, 1988; Bianchi & Delmonte, 1989a).

Expressed quantifiers and quantified expressions also require Neutralization of one feature: Number. Here below we show how the system computes a big Pro and indefinite by allowing the topic of discourse to continue. In particular, in case an indefinite is in the scope of a universal quantifier or a quantifier like "each" which must assume scope in order for the sentence to be grammatical, in lack of other sources of knowledge - which however in a given context might be available - the system rejects the quantified expression and takes the indefinite as
antecedent because it has the Number feature set to nil. In order to produce this result, LFG grammatical representations must be passed on the algorithm for scope assignment, which builds up a Logical Form, as explained below. The system calls up Logical Form automatically each time it spots an indefinite expressions in the list of possible arguments of discourse. The LF algorithm assigns scope to the quantified expressions contained in the sentence under analysis and decides whether the indefinite is in the scope of a universal quantifier. In that case it proceeds to modify the feature Number. Then it continues computing topics of discourse as usual.

41. [ogni, porcellino, ha, visto, un, lupo, nella, campagna] Every little pig saw a wolf in the countryside
    definito(s25, campagna(s25),
    state(7),
    and(situation(7),
    [l],
    ogn(s12, porcellino(s12),
    indefinito(s13, and(lupo(s13),
    in(s13, sn13, sn25)),
    vedere(s12, sn13, f7))))

    ref_ex(s12, porcellino, [+ ref, 0def, - part, nil, - pro, - ana, + class], nil, mas, sing, [edible, animato],
    sogg/esperien) 11
    ref_ex(s13, lupo, [+ ref, - def, nil, nil, - pro, - ana, + class], nil, mas, nil, [ferocious, animato],
    ogg/tema_nonaff)/15
    ref_ex(s25, campagna, [+ ref, + def, nil, nil, - pro, - ana, + class], nil, fem, sing, [luogo], ogg(nil))/1022

EXPECTED TOPIC: ref_ex(s12, porcellino, [+ ref, 0def, - part, nil, - pro, - ana, + class], nil, mas, sing, [edible, animato],
    sogg/esperien)

POTENTIAL TOPICS: ref_ex(s13, lupo, [+ ref, - def, nil, nil, - pro, - ana, + class], nil, mas, nil, [ferocious, animato],
    ogg/tema_nonaff)

ref_ex(s25, campagna, [+ ref, + def, nil, nil, - pro, - ana, + class], nil, fem, sing, [luogo], ogg(nil))

state(1, changing)

[erano, tutti, cattivi] they were all bad

ref_ex(s4, tutti, [+ ref, nil, nil, nil, - pro, - ana, - me], nil, mas, plur, [umano, animato], sogg/prop)/50

MAIN TOPIC: ref_ex(_, tutti, _, nil, mas, plur, [umano, animato], /prop)

SECONDARY TOPIC: ref_ex(_, porcellino, [+ ref, 0def, - part, nil, - pro, - ana, + class], nil, mas, sing, _, _)

state(2, resume)

5.2 Pronouns

Coming now to pronouns, they may be free, controlled or bound. A pronominal is bound only when its antecedent is a quantifier, a quantified NP or in case no controller is available at sentence level, a number of semantic conditions are met at the level of tense and mood specification and arbitrary reading is assigned to the whole sentence. In this latter case, the pronominal expression must either be an empty pronoun, big PRO or the nominative clitic si /one. Sentences with arbitrary or generic reading can in turn be assigned either universal quantification or existential quantification: only in the former case, when introduced in discourse, they may be iterated without producing incoherence. The matter is discussed at length in Cinque (1988) who comes to the conclusions that the difference in meaning is due simply to the particular tense and aspect of the sentence.

42a. Oggi a Beirut si è ucciso un innocente / Today in Beirut one killed an innocent

b. Oggi a Beirut si è sparato tutta la mattina / Today in Beirut one shot the whole morning
c. %Oggi a Beirut si è nati senza assistenza medica / Today in Beirut we were born with no medical assistance

Example c. is bad because of tense specification and of the nature of the syntactic class of the main verb, an Ergative, as contrasted by the two previous cases where we have a transitive and an unergative verb. The sentence becomes good if we change tense, by introducing present indicative which assigns to the sentence a non specific time reference,

d. Oggi a Beirut si nasce senza assistenza medica
As a matter of fact, present tense may be assigned both a generic interpretation and a specific interpretation in case the spatio-temporal location coincides with discourse time. In this case, the discourse might continue by introducing a single individual satisfying the description denoted by the sentence:
e. John's child, for example.
Cinque quotes examples from Carlson's approach to bare plurals in English, where the a. example receives a quasi-universal reading whereas the b. example receives a quasi-existential reading,
43a. Dogs run around in circles.
   b. Dogs are running around in circles.
Italian has a 3rd pers.plural construction which enters an arbitrary reading, as long as the pronoun is the empty little pro, as shown by his 47:545,
44a. Li, odiano gli stranieri / There they hate foreigners
   b. Qui lavorano anche di sabato / Here they work even on Saturday
As with other generic readings, also in this case the indetermined subject cannot be bound to a specific individual in discourse, and this is a function of tense interpretation. The interpretation of little pro however becomes specific in case the spatial adverbial is omitted, as in "Odiano gli stranieri" or "Lavorano anche di sabato". Arbitrary interpretation may also appear with 2nd person singular pronouns. Also pro-object discussed by Rizzi(1986) in detail follows the same pattern, receiving only a generic universal reading. However, it is clear from the examples quoted by Cinque, that little pro 3rd pers.plural may only be assigned generic reading when an adverbial specifying a spatial location is present, and tense is present indicative. As for impersonal "si" and middle constructions we have the same effect always due to tense and aspect specification.
5.3 Arbitrary or Generic Reading
In our system, all [+ana] marked pronouns do not possess intrinsic reference, being also marked [-ref] and two consequences ensue: they must be bound in their sentence and cannot look for antecedents in the discourse, unless there are additional conditions intervening, i.e. tense must be specific and not generic, mood must be real. Else, they can be assigned ARBITRARY interpretation, when a controller is lacking, and a series of semantic conditions are met as to tense and mood specification. Since ARBITRARY interpretation is a generic quantification on events this can be produced with untensed propositions or tensed ones, but with no deictic or definite import as shown by:
45a. I think that [prop+[arbitrary]]killing oneself is foolish
   b. I think that [prop+[definite]]killing oneself has been foolish
A further argument may be raised for Arbitrary PROs which in LFG are introduced each time the clause does not contain a controller because being a closed function it does not need one: we quote here Bresnan(1982,345) example in Italian,
46. E' difficile andarsene /It is difficult to leave
   where the infinitive "to leave" may be analysed as an extraposed COMP bound to the SUBJECT. The PRO
generated as SUBJECT of the predicate "LEAVE" receives [arbitrary] interpretation. In general, reflexive pronouns lacking the ability to refer independently receive their reference from their binders; in case no binder is available reflexive pronouns are assigned arbitrary or generic reference. This may be detected both from structural cues and from properties associated with the predicate of the matrix clause. In 46, the copulative sentence is a typical case in question: the adjective "difficult" may or may not select a binder for the infinitive which should appear with the preposition "for", thus turning the PRO from arbitrary to controlled,
46i. E' difficile per Gino andarsene /It is difficult for John to leave.
A similar case may be raised for anaphoric possessive pronouns, whenever they are contained in a subject NP.
Possessives pronouns are obviative according to whether or not they are contained in a predicative or open function, as shown by the following examples,
47. La propria_{arb} libertà è una cosa importante/One's freedom is an important thing
The sentence contains a generic statement absolutely parallel to the reading of 46; the same happens whenever the anaphoric pronoun is contained in the subject position of a closed function like a sentential complement,
48. Marta pensar che la propria_{arb} libertà sia una cosa importante/ Martha thinks that one's freedom be an important thing in a parallel way to the behaviour of PRO
48.i. Mary thinks that [ PRO to behave oneself is important. 

We may note at this point the fact that English possessive pronouns behave in a different way from Italian ones: in particular "his" may be bound by a quantifier through PRO, and it may be taken to corefer to a non c-commanding NP, differently from what happens in Italian,

49. *La sua salute preoccupa ognuno

50. PRO Knowing his father pleases every boy ≠ Conoscere proprio padre fa piacere a ogni ragazzo

51. His mother loves John ≠ Suo madre ama Gino

In particular, "his" seems to possess the ability to be bound by quantifiers like "proprio" does: in 50, the Italian version becomes analogous to the English one if we substitute "proprio" to "suo". In other words, Italian has two separate lexical pronouns for bound and unbound reference whereas English has only one and the conditions on binding are simply structural whereas in Italian they are both structural and lexical. The peculiarity of long-distance anaphors emerges from the dependency of binding on the presence of a feature at sentence level, the one related to the mood of the subordinate clause. In particular, as also detected in other languages (cf. Zaenen, 1983) the choice of Indicative vs. Subjunctive Mood is relevant for the binding possibilities of anaphors contained in the clause. The presence of the Indicative, in the most embedded clause, the one containing the long-distance anaphor seems to block binding from the matrix clause, as shown in:

52. Gino pensa che tu sia convinto che la propria famiglia sia la cosa più importante.

53. Gino pensa che tu sei convinto che la propria famiglia è la cosa più importante.

John thinks that you be/are convinced that self's family be/is the most important thing, where we changed subjunctive in 52, to indicative in 53: only 52 allows binding, hence bound reference, and disallows arbitrary reference; on the contrary 53 only allows arbitrary reference i.e. no reference at all. As discussed at length in Zaenen (1983) the choice of the mood is bound by the matrix verb which permits only certain kind of referential acts to be realized by the complement clause. Being lexical, this information can be easily transmitted in features to the c-structure and percolated according to the usual LFG conventions (see Giorgi, 1984, for a lexical typology of the governing verbs).

The same applies to derived nominals like "suspicion" which can be the head of a sentential complement, inducing long-distance binding or preventing according to the presence of [+BOUND] feature,

54. Gino ritiene che il sospetto di Carlo che la propria sorella sia un assassino abbia determinato la sua condanna.

55. Gino ritiene che l'affermazione di Carlo che la propria sorella è un assassino abbia determinato la sua condanna.

/ John believes that the Karl's suspicion that self's sister be/is a murdered had determined his/her trial.

The relations intervening between tense, mood and pronominal expressions which possess the feature [+ana] are computed by a special module for interpretation, within the binding algorithm, at sentence level. The output of the module is made visible by the attribute "interpretation".

5.4 Quantifiers and quantified NP's as antecedents

As a first approach to the problem of quantifiers, the algorithm takes care of precedence whenever a quantified NP is indicated as possible antecedent for a pronoun. Quantified antecedents are individuated by the presence of the feature ±part in SPEC, as follows,

56. quantified(ante) :- node(N):index:ante,
   node(N):spec:part:._

This predicate is used for quantified antecedents in a simple declarative with psychic verbs: thus, binding of a possessive long distance anaphor can take place from a quantified antecedent contained at clause level.

However, when we want to deal with quantifiers and quantified NPs as possible antecedents of little pros, clitics or independent pronouns a different procedure must be called in, and is the following one,

57. a. non_quantif(ante) :- node(N):index:ante,
   not node(N):spec:part:_, !.

b. non_quantif(ante) :- node(N):index:ante,
   node(N):spec:part:X,
\( (X = \cdot) \),
\( \text{node}(N) : \text{spec} = :^* \).

This procedure is integrated into the predicate for referring clitics, in particular as follows,

58. \text{refer}(\text{Net}, \text{Ind}, [+ref,+pro,+ana,+me], \text{Ante}/N) :-
    \text{node}(\text{node}):(\text{index}: \text{Ind},
    \text{node}(\text{node}):(\text{cat}: \text{features},
    \text{node}(\text{node}):(\text{num}: \text{number},
    \text{node}(\text{node}):(\text{gen}: \text{gender},
    \text{find}_\text{gender}(\text{node}, \text{Gen}),
    \text{f}_\text{command}(\text{N}\text{Ante}, \text{F}_\text{ante}, \text{Ind}, N) > 0,
    \text{f}_\text{structure}(\text{N}\text{Ante}, \text{F}_\text{ante}, N, \text{ante}),
    \text{not contains}(\text{N}\text{Ante}, \text{Ind}),
    \text{node}(\text{N}_\text{ante}): \text{F}_\text{sup}: \text{node}(\text{N}2),
    \text{node}(\text{N}2): \text{F}/\text{R}: \text{index}: \text{Ante},
    \text{non}_\text{quantiff}(\text{Ante}),
    \text{not node}(\text{N}2): \text{path}(\cdot): \text{Ind},
    \text{node}(\text{N}2): \text{F}/\text{R}: \text{cat}: \text{Cat},
    \text{features}(\text{Cat}, \text{features}),
    \text{node}(\text{N}2): \text{F}/\text{R}: \text{gen}: \text{Gen}_\text{ante},
    \text{node}(\text{N}2): \text{F}/\text{R}: \text{num}: \text{Num}_\text{Ante},
    \text{number} = \text{Num}_\text{Ante},
    \text{node}(\text{N}2): \text{F}/\text{R}: \text{ref}_\text{tab}: \text{List},
    \text{pos}_\text{ante}(\text{Ind}, \text{Ante}, \text{List}),
    \text{non}_\text{referred}_\text{in}(\text{Ind}, \text{Ante}).

In this way we can account for lack of coreference between a clitic pronoun contained in a fronted subordinate clause and a quantified NP contained in the main clause, as in the a.
example

59a. When I insulted him, every student went out of the room.

b. When I insulted him, John went out of the room.

as opposed to the b. example, where coreference is allowed as usual. This notion of binding which is relevant for long-distance anaphors is also important for quantifiers as discussed in another work (Delmonte, 1989), in particular the fact that pronouns embedded in an Indicative or [-BOUND] clause need referential antecedents and not arbitrary or generic ones, as shown by the pair

60a. A woman requires/demands that many/every men be in love with her, *and John knows her.

b. A woman believes that many men like her, and John knows her.

in 60a., both in English and Italian, the indefinite "a woman" is computed as generic in the main clause and the same happens to the pronoun "her" in the complement clause introduced by "that"; but the conjoined sentence is expressed in the indicative and requires a specific woman to be picked up for referring the pronoun "her", which in this case must be computed as referential and not as generic, so the sentence is ungrammatical. The opposite happens in 60b., where the indefinite is taken to refer to a specific woman in the discourse, and the two occurrence of "her" to be bound to this individual. As clearly shown, the referential capabilities of pronouns are tightly linked to the ones of their antecedent: but the opposite may happen, i.e. the referential abilities of the antecedents are bound by those of the pronouns, and these in turn are conditioned by the referential nature of the RD- referential domain - in which they are contained: an [-BOUND] domain is one containing indicative mood and reference is free, whereas a [+BOUND] domain is one containing subjunctive mood and reference not free but locally bound, for anaphors, or lacking in referential import for lexical pronouns. However, in order to produce an adequate result, the discourse module must also query tense and aspect interpretation module which alone may assess whether a generic or an existential reading is made available by the sentence under analysis. The two following examples show the output of the binding and the interpretation module,

611. [la, propria, salute, era, necessaria]/one's health was necessary
61ii. [la,propris, salute, e_necessaria]/one’s health is necessary
net(es17c)
index:f8
pred:essere
lex_form:[np/subj/theme_bound[/_ _]]
mood:ind
tense:imp
cat:esistenza
subj/theme_bound:index:sn49
cat:[stato]
pred:salute
gen:fem
num:sing
spec:inf:+
subj/poss:index:sn50
cat[]
pred:proprio
gen:fem
num:sing
spec:inf:+
tab_ref:[- ref, + pro, + ana, + me, - subj]
antecedent:external
interpretation:specific
tab_ref:[+ ref, - pro, - ana, + class]
acomp/prop:index:sa4

cat:[valutativo]
pred:necessario
gen:fem
num:sing
subj/prop:index:sn51
pred:vbl
binder:sn49

tab_ref:[+ ref, - pro, - ana, - me]
aspect:state
As appears from the value of the attribute "interpretation" only example ii. is assigned arbitrary reading. The discourse module activates then the Logical Form which computes scope for all quantified expressions. Logical Form translates f-structures into well-formed-formulas and then assigns scope. As a result the first sentence has a nominal expression "la salute" which is taken to be referential; in the second case, the arbitrary reading associated to the anaphoric pronoun "proprio" prevents the nominal expression to assume scope, as shown by the two following LF representations,

i. [la, propria, salute, era, necessaria]

input form

wff(essere, [term(definio, sn49), wff(and, [wff(salute, [sn49! _2185]), wff(necessario, [sn49])))], wff(necessario, [sn49], term(state, saa2), wff(and, [wff(situation, [saa2]), [], []]))), term(state, f8, wff(and, [wff(situation, [f8]), []])))

output form

definito(sn49,and(salute(sn49),
necessario(sn49)),
state(saa2),
and(situation(saa2),
[],
state(f8),
and(situation(f8),
[],
essere(sn49,
  necessario(sn49, saa2),
  f8)))))

ii. [la, propria, salute, e_, necessaria]

input form

wff(essere, [term(arbitraria, sn49), wff(and, [wff(salute, [sn49! _1741]), wff(necessario, [sn49])))], wff(necessario, [sn49], term(state, saa4), wff(and, [wff(situation, [saa4]), [], []]))), term(state, f8, wff(and, [wff(situation, [f8]), []])))

output form

state(saa4),
and(situation(saa4),
[],
state(f8),
and(situation(f8),
[],
arbitraria(sn49,and(salute(sn49),
necessario(sn49)),
  essere(sn49,
necessario(sn49, saa4),
  f8)))

A completely different result is produced when an empty pronominal, big PRO is introduce into discourse and no generic or arbitrary reading is triggered by the interpretation module. In this case, the pronominal is bound by the topic of discourse and a Continue is produced as a result of discourse state strategy. We include a brief excerpt from the story of the three little pigs:

[i, lupo, stava, gia_, arrivando, deciso, a, non, rinunciare, al, suo, pranzetto]/within seconds the wolf was arriving determined not to give up his meal

ref_ex(sn22, lupo, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [animato], subj/agent)/10
ref_ex(sn27, pranzetto, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [evento, oggetto],obl/theme_aff)/53

EXPECTED TOPIC : ref_ex(_, lupo, _, nil, mas, sing, [animato], _/_

SECONDARY TOPIC : ref_ex(_, porcellino, _, _, mas, plur, [uman0, animato], _/_
POTENTIAL TOPICS: ref_ex(sn27, pranzo), [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [evento, oggetto], obl/theme_aff
state(18, changing)

[sicuro, di, abbarbattere, anche, la, cassetta, di, mattoni, il, lupo, si, riempi, _i, polmoni, di, aria, e, comincio, _a, soffiare, con, forza, _i, alcune, volte] convinced that he could also blow the little brick house down, he filled his lungs with air and huffed and puffed a few times
ref_ex(sn699, lupo, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [animato], subj/agent)/30
ref_ex(sn709, aria, [+ ref, 0def, nil, - pro, - ana, + class], nil, fem, sing, [sostanza, luogo], obl/theme_nonaff)/43
ref_ex(sn701, polmone, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, plur, [oggetto], obj/theme_aff)/1035
ref_ex(sn718, forza, [+ ref, 0def, nil, - pro, - ana, + class], nil, fem, sing, [modo], obl/modal)/1080
ref_ex(sn719, volta, [+ ref, 0def, +, - pro, - ana, + class], nil, fem, plur, [tempo, ripetizione], adj/temporal)/1100
MAIN TOPIC: ref_ex(, lupo, _, nil, mas, sing, [animato],_/agent)
POTENTIAL TOPICS: ref_ex(sn709, aria, [+ ref, 0def, nil, - pro, - ana, + class], nil, fem, sing, [sostanza, luogo], obl/theme_nonaff)
ref_ex(sn701, polmone, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, plur, [oggetto], obj/theme_aff)
ref_ex(sn718, forza, [+ ref, 0def, nil, - pro, - ana, + class], nil, fem, sing, [modo], obl/modal)
ref_ex(sn719, volta, [+ ref, 0def, +, - pro, - ana, + class], nil, fem, plur, [tempo, ripetizione], adj/temporal)
state(19, continue)

[niente, da, fare]/no way
ref_ex([], pPro, [+ ref, nil, nil, + pro, + ana, - me], _, _ _, [umanco, animato], subj/agent)/30
pPro=lupo
MAIN TOPIC: ref_ex(_, lupo, _, nil, mas, sing, [animato],_/agent)
state(20, continue)

[la, casa, non, si, mosse, di, un, solo, palmo]/the house did not move an inch
ref_ex(sn218, casa, [+ ref, +, nil, - pro, - ana, + class], nil, fem, sing, [oggetto, luogo], sogg/esperienz)/11
ref_ex(sn241, palmo, [+ ref, -, nil, - pro, - ana, + class], nil, mas, sing, [misura, luogo, oggetto], obl/misura)/1020
EXPECTED TOPIC: ref_ex(_, casa, _, _, fem, sing, _ _, _ _)
SECONDARY TOPIC: ref_ex(_, lupo, _, nil, mas, sing, [animato], _) _
POTENTIAL TOPICS: ref_ex(sn241, palmo, [+ ref, -def, nil, - pro, - ana, + class], nil, mas, sing, [misura, luogo, oggetto], obl/misura)
state(21, changing)

[alla, fine, esausto, il, lupo, si, accascio, _a, terra] in the end the wolf fell to the ground exhausted
ref_ex(sn50, lupo, [+ ref, +def, nil, - pro, - ana, + class], nil, mas, sing, [animato], subj/experienzi)/11
ref_ex(sn49, fine, [+ ref, +def, nil, - pro, - ana, + class], nil, fem, sing, [evento, tempo], adj/temporal)/70
ref_ex(sn57, terra, [+ ref, 0def, nil, - pro, - ana, + class], nil, fem, sing, [luogo], subj/nil)/1032
MAIN TOPIC: ref_ex(_, lupo, _, nil, mas, sing, [animato], _/experienzi)
SECONDARY TOPIC: ref_ex(_, casa, _, _, fem, sing, _ _, _ _)
POTENTIAL TOPICS: ref_ex(sn49, fine, [+ ref, +def, nil, - pro, - ana, + class], nil, fem, sing, [evento, tempo], adj/temporal)
ref_ex(sn57, terra, [+ ref, 0def, nil, - pro, - ana, + class], nil, fem, sing, [luogo], subj/nil)
state(22, continue)

6. CONCLUSIONS
The number of residual problems is clearly very high, particularly in the role of inferencing mechanisms. We list here below some of these cases:
21. Sentence pronominalization
I don't think you should go to the party, your wife wouldn't like it

22. Deverbal nominalization and metonymic reference

I no longer want to ski in the Dolomites. People practise that sport because it is fashionable.

We quote from Moravia, 1954, 2:

A. 'The German army had stolen in the villa of the prince the box containing silver cutlery...'

further down in the same text the event reappears as,

B. 'The robbery had actually taken place and the prince had never found his silver cutlery again.'

'The robbery' is related to the action of 'stealing' referred to in a previous part of the text: however to do this, the system must be equipped with very powerful inferencing mechanisms, so that the verb phrase is at first decomposed in its primitives, denoting a certain event, and the same event is recalled by the noun phrase in the following discourse. Work is now in progress in this area using KL-Prolog. As clearly shown by Webber (1988), parts of an event may also be referred to in a subsequent discourse: in this sense tensed clauses, possessing both tense and aspect specification of states and events being spoken of, have a much more complicated structure than definite NPs, even though from a procedural point of view they must be computed in a similar fashion when binding anaphors.

We are also working at a module for summarizing texts which makes of the output of the discourse system ROSIE, and generates natural language from a conversion of LFG grammatical representation into Situational Semantics schema (Delmonte and Pianta 1991).

P.S. All modules have been implemented in Prolog and run both under MS-Dos and VMS in Quintus Prolog, and Macintos under MacProlog.

References


BOSCH P. (1980), The modes of pronominal reference and their constraints, in J. Kreiman, A. Ojeda (eds), Papers from the Parasession on Pronouns and Anaphora, CLS, Chicago, IL, 64-78.


BRESNAN J. (1990), Levels of Representation in Syntax, Invited Speech, GLOW, Cambridge


DELMONTE R. (1989a), Grammatica e quantificazione in LFG, Manuscript, University of Venice.


HALLIDAY M.K., R. HASAN (1975), Cohesion in English, Longman.

HINRICH'S E. (1988), Tense, Quantifiers, and Contexts, in Computational Linguistics - Special Issue on Tense and Aspect - 14, 2, 3-14.