Pronominal Anaphora Processing

Anna Maria Di Sciullo

Abstract—Discourse pronominal anaphora resolution must be part of any efficient information processing systems, since the reference of a pronoun is dependent on an antecedent located in the discourse. Contrary to knowledge-poor approaches, this paper shows that syntax-semantic relations are basic in pronominal anaphora resolution. The identification of quantified expressions to which pronouns can be anaphorically related provides further evidence that pronominal anaphora is based on domains of interpretation where asymmetric agreement holds.

Keywords—asymmetric agreement, pronominal anaphora, quantifiers and indefinite expressions.

I. THE PROBLEM

A pronoun lacks independent reference. Proper names and definite descriptions are possible referents to pronouns. This is also the case for indefinite expressions and quantifiers under certain conditions. This paper focuses on cases where a pronoun is anaphorically related to a quantifier in a precedent sentence. In (1), the existential quantifier can be the antecedent of a pronoun. In (2), the universal quantifier cannot be anaphorically related to the individual pronoun *he*, whereas this is possible in (3), where an indefinite expression, a *corporate number*, can be reconstructed as a restrictor for the proposition including the pronoun. This phenomenon is referred to in terms of telescoping in [30].

(1) *Someone* came to the meeting. *He* was expected to vote on the motion.

(2) *Everyone* came to the meeting. *He* was expected to vote on the motion.

(3) *Everyone with a corporate number* came to the meeting. *He* was expected to vote on the motion.

Information processing systems, including information extraction and question answering, must be able to identify the possible antecedents of pronouns, since the information requested in a query or in a question can be traced through discourse pronominal anaphora. For example, the answer to the question in (4) can be accessed through the linking of the pronoun *he* in the second sentence in (3) to the quantifier phrase *everyone with a corporate number* in the preceding sentence.

(4) Who voted on the motion?

So-called “knowledge-poor” systems for discourse pronominal anaphora resolution systems [21], [25] use limited linguistic knowledge to identify the antecedents of pronouns. They rely on the string-linear position of pronouns and possible antecedents, without taking into account the fine-grained syntax-semantic properties of the expressions they are part. Consequently, their success rate is poor. For example, Mitkov Anaphora Resolution System (MARS) [27] incorrectly identifies the constituent *the meeting*, instead of the constituent *everybody with a corporate number*, as being the antecedent of the pronoun *he* in the examples (1)-(3) above:

(5) MARS result for (3): *He* appears in paragraph 2, sentence 2, from position 1 to position 1. It is singular. The antecedent is indicated to be *the meeting* in paragraph 2, sentence 1, from position 7 to position 8.

(6) MARS result for (2): *He* appears in paragraph 1, sentence 2, from position 1 to position 1. It is singular. The antecedent is indicated to be *the meeting* in paragraph 2, sentence 1, from position 3 to position 4.

(7) MARS result for (1): *He* appears in paragraph 3, sentence 2, from position 1 to position 1. It is singular. The antecedent is indicated to be *meeting* in paragraph 3, sentence 1, from position 1 to position 1.

The efficiency of information processing systems depends on their ability to process fine-grained syntax-semantic properties of linguistic expressions. What are the syntax-semantic properties of quantifiers and indefinites that make them possible antecedents for pronouns? Contrary to proper names and definite descriptions, quantifiers and indefinites are usually considered to be non-referential expressions, i.e., they do not refer to individuals in the universe of interpretation.1 However, a quantified

1 A quantifier requires generalizing over the individual entities of the universe of interpretation. For example, the truth of a quantificational statement such as *everyone wrote a program* requires that, for all the individuals in the universe that can be substitutes for *x* in “*x* wrote a program”, the outcome is true. The truth of a quantificational statement such as *someone wrote a program* requires finding some individual or other in the universe that can be
expression can be the antecedent of a pronoun, as illustrated above.

The properties of discourse pronominal anaphora have been widely discussed in semantic theory \cite{15}, \cite{8}, \cite{20}, \cite{23}, and different approaches to this phenomenon are available in the literature. For example, in the variable binding approach \cite{17}, \cite{18}, \cite{2}, \cite{3}, indefinite expressions and quantificational expressions are essentially of the same type. The binding effects observed in examples such as (1) and (3) fall out of an extension of the scope domain of the quantifier. However, the universal quantifier is usually static (but it also has dynamic definitions), which accounts for the lack of binding effect in (2). In the restrictor reconstruction approach \cite{30}, telescoping is viewed as the reconstruction of the restrictor of a sentence. In this approach, it is the discourse, including the syntactic properties of the linguistic expressions, that makes it clear that a given sentence is interpreted relative to a restrictor. In fact, the variable binding and the restrictor reconstruction approaches have in common that quantifiers and anaphoric pronouns are asymmetrically related, and that must also be semantically related.

I have shown in \cite{11} that a definite description can be a possible antecedent of a pronoun if it asymmetrically agrees with that pronoun. I focus here on cases where a pronoun is anaphorically related to a quantifier or an indefinite, in order to show that asymmetric agreement is also at play.

The organization of this paper is the following. First, I define the notion of asymmetric relation. Second, I illustrate that pronouns can be bound in different domains. Third, I show how discourse pronominal anaphora resolution based on asymmetric agreement makes correct predictions for the processing of pronouns bound by quantifiers and indefinites.

II. ASYMMETRY THEORY

In Set Theory \cite{32}, asymmetry is a property of a relation \(R\) such that there are no ordered pairs in \(R\) whose members are inverted. Symmetric relations do not have this property. 2 In linguistic theory \cite{5}, \cite{22} the structure of linguistic expressions is represented in terms of oriented graphs, where asymmetric relations are defined in terms of precede, dominate, and asymmetric c-command. Asymmetric c-command (8) is relevant across the board in grammar, including in binding and movement \cite{6}, \cite{7}, \cite{10}, \cite{12}, \cite{19}, \cite{22}, \cite{27}, \cite{19}, \cite{28}. Thus in (9), \(X\) asymmetrically c-commands \(Y\).

\textbf{(8) a. C-command:} \(X\) c-commands \(Y\) iff \(X\) and \(Y\) are categories and \(X\) excludes \(Y\), and every category that dominates \(X\) dominates \(Y\). \cite{22}

\textbf{(8) b. Asymmetric c-command:} \(X\) asymmetrically c-commands \(Y\), if \(X\) c-commands \(Y\) and \(Y\) does not c-command \(X\). \cite{22}

\textbf{(9)} \[ \begin{array}{c} Z \\
\vee \\
X \\
\vee \\
Z \\
\vee \\
Y \\
\end{array} \]

In Asymmetry Theory \cite{10}, \cite{12}, asymmetric relations are part of morphological and syntactic expressions. The theory correctly predicts that a change in morphological relations gives rise to either gibberish or a difference in semantic interpretation. It also correctly predicts that a change in the syntactic relations brings about a change in information structure. In this theory, the operations of the grammar apply under asymmetric Agree (10), (11). Thus, the features of two related elements are in a proper subset relation. 3

\textbf{(10) a. Shift \((\alpha,\beta)\)}

Given two objects \(\alpha, \beta\), Shift \((\alpha,\beta)\) derives a new object \(\delta\) projected from \(\alpha\).

\textbf{b. Link \((\alpha, \beta)\)}

Given two objects \(\alpha\) and \(\beta\), Link \((\alpha, \beta)\) creates a new object where \(\alpha\) and \(\beta\) are featurally related.

\textbf{(11) Agree \((\varphi_1, \varphi_2)\)}

Given two sets of features \(\varphi_1\) and \(\varphi_2\), Agree \((\varphi_1, \varphi_2)\) applies if and only if \(\varphi_1\) properly includes \(\varphi_2\).

In this theory, only elements in asymmetric relation are optimally interpretable at the interfaces with the external systems, conceptual-intentional and sensorimotor.

Asymmetry Theory has implications for natural language technologies, including information extraction and question answering, as shown in \cite{13}, \cite{14}. The processing of the asymmetric properties of linguistic expressions is expected to improve any area where human users can benefit by communicating with their computers in a natural way.

III. LOCAL DOMAINS OF INTERPRETATION

Locality is another salient property of natural languages. It has been shown that the syntactic operations apply to local domains and that semantic interpretation is domain-dependent. More recently, the notion of local domain has been thought of in terms of the notion of phase \cite{5}, \cite{6}, \cite{31}. A syntactic phase is a unit of the computation and interpretation: it has an internal structure, it is subject to impenetrability, and it is isolatable at the interfaces.

\footnote{2} If \(R \subseteq A \times A\), then \(R\) is symmetric if \(f(\forall x y)(\langle x, y \rangle \in R \rightarrow \langle y, x \rangle \in R)\).

If \(R \subseteq A \times A\), then \(R\) is asymmetric if \(f(\forall x y)(\langle x, y \rangle \in R \rightarrow \langle y, x \rangle \notin R)\).

\footnote{3} Given two sets \(A\) and \(B\), if all the members of \(A\) are also the members of \(B\), \(A\) is a subset of \(B\).

\(A\) is a proper subset of \(B\), or is properly included in \(B\), whenever \(A\) is a subset of \(B\) but \(A\) is not equal to \(B\).
A. Morphological Domain

In Asymmetry Theory, local domains of interpretation are also part of the derivation of morphological expressions, such as reflexive pronouns and quantifiers [9], [12]. A morphological domain typically includes two layers: an affix occupies the higher layer, and a root heads the lower layer, as in (12) where \( a, \beta, \) and \( \delta \) are placeholders for morphological features. A morphological domain is derived by the operations of the grammar applying under asymmetric Agree.

\[
(12) \quad [x \ a \ X \ [v \ \beta \ Y \ \delta]]
\]

The configuration in (16) is the minimal domain form for interpretable morphological expressions. Considering quantifiers, operator-variable-restrictor linking applies in their morphological domain. For example, in (13a) every is the operator that locally links, under asymmetric Agree, a variable \( X \), and one is the local restrictor of \( X \). In syntax, quantifier raising [26], [16] recovers the syntactic operator-variable relation, as illustrated in (14). Considering reflexive pronouns, linking also applies in their morphological domains. The pronoun him is a restrictor of the reflexive operator, and it is anaphorically linked to the reflexive head self which projects its features to the whole reflexive construct (13b)). Consequently, himself is a reflexive pronoun that must be bound by an antecedent in the syntactic domain.

\[
(13) \quad \text{a. } X \quad \text{b. } X
\]

(14) a. Everybody trusts somebody.

b. [Everybody [somebody [everybody trusts somebody]]

\( \forall X \exists Y (\text{QR: wide scope universal quantifier}) \]

c. [Somebody [everybody [everybody trusts somebody]]]

\( \exists X \forall Y (\text{QR: wide scope existential quantifier}) \]

B. Syntactic Domain

The examples in (15) and (16) illustrate that in the syntactic domain, e.g., in a proposition, a reflexive pronoun, such as himself, must be linked to an antecedent, whereas a pronoun, such as him, must be free.

\[
(15) \quad [\text{Everyone [trusts himself]}].
\]

\[
(16) \quad [\text{Everyone [trusts him]}].
\]

The Binding Theory (17), [4], [5] expresses this generalization in terms of two conditions that hold locally in a Binding Domain (BD). If two arguments are bound, they have the same reference. They do not have the same reference if they are free. The notions of ‘bound’ and ‘free’ are defined in terms of the asymmetric c-command relation, (18).

(17) Binding Theory

A. An anaphor is bound in its BD.

B. A pronominal is free in its BD.

(18) \( a \) is bound by \( \beta \) iff \( a \) and \( \beta \) are co-indexed and \( \beta \) asymmetrically c-commands \( a \) \( a \) is free iff \( a \) is not bound.

Asymmetric c-command is a necessary condition for binding. A reflexive anaphor, such as himself, must be asymmetrically c-commanded by its local antecedent. A pronoun, such as him, must be free in its BD under asymmetric c-command, however, it may be bound by an antecedent that does not asymmetrically c-command it, as illustrated in (19), and (20), where TP stands for Tense Phrase, QP stands for Quantifier Phrase, DP stands for Determiner Phrase, and VP stands for Verb Phrase).

(19) \( [\text{Everyone who owns a corporate number [uses it]].} \]

Moreover the syntax-semantic properties of the antecedent of a pronoun are also at play, as the following example illustrates.

(21) \( \text{Everybody with a corporate number thinks that he will vote.} \]

In (21), the antecedent of the pronoun he is the full QP constituent everybody with a corporate number, see (22) where PP stands for Prepositional Phrase, and CP for Complementizer Phrase).

(22) \( \text{Everybody with a corporate number thinks that he will vote.} \]

International Scholarly and Scientific Research & Innovation 1(11) 2007 3648

ISNI:0000000091950263
It is not surprising that MARS fails to identify the antecedent of the pronouns in these cases as well. No antecedent is found for the pronoun it in (19), see (23). Likewise for the pronoun he in (21), see (24).

(23) MARS result for (26): it appears in paragraph 1, sentence 1, from position 6 to position 6. It is singular. The antecedent is indicated to be ![NOTHING!! in paragraph , sentence , from position to position .

(24) MARS result for (28): he appears in paragraph 1, sentence 2, from position 7 to position 7. It is singular. The antecedent is indicated to be ![NOTHING!! in paragraph , sentence , from position to position .

These facts further show that knowledge-poor pronominal anaphora resolution systems are not optimal. They also point to the correctness of the view that the fine-grained syntax-semantic properties of the linguistic expressions are crucial for pronominal anaphora resolution.

C. Discourse Domain

While pronouns must be free (under asymmetric c-command) in their BD, they must be bound by an antecedent outside of their BD. The example in (3) above shows that the presence of an indefinite expression in the domain of the universal quantifier makes the whole quantified expression a possible antecedent for a pronoun. These data show that discourse pronominal anaphora resolution must be sensitive to the syntax-semantic properties of the constituents of the discourse. In particular, the presence of an operator in a previous sentence may bring about the reconstruction of the restrictor of a sentence including a pronoun. Moreover, the formal and semantic features of the quantifiers and the pronouns are also crucial. For example, without an embedded indefinite, the quantifier everyone cannot be a possible antecedent for the individual pronoun he, whereas it can be for plural pronoun they (25). The quantifier someone differs from everyone in this respect (26).

(25) [(Everyone came to the meeting). [\#He/They was/were z
…………..……………m
expected to vote on the motion]].

(26) [(Someone came to the meeting). [ He/They was/were z
…………..……………m
expected to vote on the motion]].

MARS finds no antecedent for the pronoun they in (25), it also finds no antecedent for the pronoun he in (26), see (27).

However, the quantifiers everyone and someone in the preceding sentences are possible antecedent for these pronouns.

(27) MARS result for (35): They appears in paragraph 2, sentence 2, from position 1 to position 1. It is plural. The antecedent is indicated to be ![NOTHING!! in paragraph , sentence , from position to position .

MARS result for (26): He appears in paragraph 2, sentence 2, from position 1 to position 1. It is plural. The antecedent is indicated to be ![NOTHING!! in paragraph , sentence , from position to position .

I develop the view that discourse pronominal anaphora resolution is basically determined by linguistic knowledge, and in particular by the fine-grained syntax-semantic properties (see also [1], [11], [14] for richer syntax-semantic knowledge-based approaches to discourse pronominal anaphora resolution than knowledge-poor approaches).

IV. PRONOMINAL ANAPHORA RESOLUTION USING ASYMMETRIC AGREEMENT

I assume the interface condition on pronominal anaphora in (28), proposed in [11] in terms of the Link operation of Asymmetry Theory.

(28) DD-Linking (Discourse Domain-Linking)

A pronominal must be linked in its DD.

Like the other operations of this theory, Link (12) applies under asymmetric Agree, see (13). Pronominal anaphora resolution is essentially the identification of the closest DP/QP with respect to which the features of a pronoun DPro stand in a proper inclusion relation.

I take the elements that enter into bound pronominal anaphora to have the formal features determiner (D) and quantifier (Q), and the phi-features person (pers), number (num), gender (gen). Both pronouns and definite determiners are D, but differ in their phi-features, definite determiners not being specified for person and gender in English. DPs differ from DPros, Ns are inherently 3rd pers. DPs, QPs, and DPros have semantic features that participate in anaphoric relations. DPs and QPs have independent reference [+Ir], and DPros are [-Ir]. An anaphoric relation has only one [+Ir] feature, and the [-Ir] feature of DPros is linked by the [+Ir] feature of the antecedent DPs or QPs. Given the Binding Theory, an anaphoric pronoun, such as himself, must be bound under asymmetric c-command by an antecedent in its BD, whereas pronouns must be free. Given DD-Linking, a pronoun, such as him, must be linked in its discourse domain.

The formal and semantic features that are necessary for pronominal anaphora resolution based on asymmetric agreement are specified in (29). The feature specifications are provided for DPros, QPs, and DPs. They differ with respect to the formal phi-feature specifications, including person (pers.), number (num.), and gender (gen.). QPs like DPs are 3rd pers., whereas this is not necessarily the case for DPros. QPs are specified for number features, but not for gender features in languages such as English (this is not the case in some other languages, including the Romance languages). The semantic features include the independent reference feature ( [+Ir]),
along with the animic ([±ani]) feature, the part-whole ([±w]) feature, and the group ([±gr]) feature. The [±ani] feature differentiates he from it, and the [±w] feature differentiates anaphoric pronouns, such as himself, from non-anaphoric pronouns, such as he and him, and from DPs. Non-anaphoric pronouns and DPs are [±w], anaphoric pronouns are [±w].

The [±gr] feature differentiates QPs with a group reading, such as everyone, from those that do not, such as someone.

(30) DPros, QP, and DPs formal and semantic features

<table>
<thead>
<tr>
<th>Form: pers, num, gen</th>
<th>Sem: Ir, ani, w, gr</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPro</td>
<td>+       +       -/+   +/-   +/-</td>
</tr>
<tr>
<td>QP</td>
<td>3rdpers + u</td>
</tr>
<tr>
<td>DP</td>
<td>3rdpers + +</td>
</tr>
</tbody>
</table>

V. Predictions

Pronominal anaphora resolution based on (28) and (29) makes correct predictions. A reflexive pronoun is bound and a pronominal is free. In (30), everyone is the antecedent of himself, illustrating this.

(31) [Everyone trusts him] who owns a corporate number]

The BD for pronouns and anaphors is limited to the embedded propositions in the examples in (32) and (33), and within these propositional domains, the reflexive pronoun is bound and the pronominal is free. In (32), everyone is the antecedent of himself, the features of the antecedent is the superset of the features of the anaphor himself. In (33), everyone, which is located outside of the binding domain of the pronoun him, is a possible antecedent, as the dotted line indicates, for the pronoun him. The linking relation is also obtained under asymmetric Agree.

(32) [Someone thinks [that everyone trusts himself]].

(33) [Everyone thinks [that the president trusts him]].

In the example in (34) repeated here in (35), he makes correct predictions. A reflexive pronoun is bound and a pronominal is free under asymmetric c-command in their BD.

(34) [Everyone who owns a corporate number]

Asymmetric Agree is also at play in the domain of the discourse, given (28). The examples in (35) and (36), where the antecedent of the pronoun is a quantifier, illustrate this point.

(35) [[Someone came to the meeting] He was expected to vote on the motion].

(36) [[Everyone came to the meeting] They were expected to vote on the motion].

The proposed system, based on asymmetric Agree, is flexible enough to account for the fact that everyone can be the antecedent of a plural pronoun, such as they in (36), even though its set of formal features includes +sing. The set of features of everyone is the superset of the features of the plural pronoun they. Moreover, the system can also handle cases where everyone is a possible antecedent for the singular pronoun he (37). The difference between someone and everyone takes the form of a difference in the value of the semantic feature ±gr. Someone is -gr, and can only be a possible antecedent for a +sing pronoun, whereas everyone is +gr and, in the context of an indefinite expression, such as with a corporate number, can be an antecedent for a +sing pronoun (38). In both cases, the quantifier is in asymmetric agreement relation with the pronoun.

(37) [[Someone with a corporate number] came to the meeting].

International Scholarly and Scientific Research & Innovation 1(11) 2007 3650

ISNI:0000000091950263
(38) [[Everyone with a corporate number] came to the meeting]. [[He was expected to vote on the motion]].

\[
\{ +Ir, +ani, +w, -gr \} \\
\{ +3rdpers, + sing, +masc \}
\]

Asymmetric Agree holds in all the domains of interpretation.

VI. SUMMARY

Knowledge-poor systems for pronominal anaphora resolution cannot handle cases where a quantifier is anaphorically related to a pronoun. The syntax-semantic properties of quantified expressions, indefinites, and pronouns cannot be dealt with by systems that mainly process string-linear properties of linguistic expressions. Knowledge-rich systems are necessary for efficient (bound) pronominal anaphora resolution. DD-Linking is a syntax-semantic discourse interface condition requiring that a pronoun, i.e., an element that lacks independent reference, be linked to an antecedent with which it asymmetrically agrees. Linking applies to domains of interpretation, which may in some cases reconstruct for the interpretation of pronouns related to quantifiers. Pronominal anaphora resolution crucially relies on the dynamic syntax-semantic processing of these domains.

REFERENCES