# 17 DISCOURSE*

> Gracie: Oh yeah...and then Mr. and Mrs. Jones were having matrimonial trouble, and my brother was hired to watch Mrs. Jones.
> George: Well, I imagine she was a very attractive woman.
> Gracie: She was, and my brother watched her day and night for six months.
> George: Well, what happened?
> Gracie: She finally got a divorce.
> George: Mrs. Jones?
> Gracie: No, my brother's wife.
>
> George Burns and Gracie Allen in *The Salesgirl*

Up to this point of the book, we have focused primarily on language phenomena that operate at the word or sentence level. Of course, language does not normally consist of isolated, unrelated sentences, but instead of collocated, related groups of sentences. We refer to such a group of sentences as a **discourse**. DISCOURSE

The chapter you are now reading is an example of a discourse. It is in fact a discourse of a particular sort: a **monologue**. Monologues are characterized by a *speaker* (a term which will be used to include writers, as it is here), and a *hearer* (which, analogously, includes readers). The communication flows in only one direction in a monologue, that is, from the speaker to the hearer. MONOLOGUE

After reading this chapter, you may have a conversation with a friend about it, which would consist of a much freer interchange. Such a discourse is called a **dialogue**. In this case, each participant periodically takes turns DIALOGUE

*This chapter was written by Andrew Kehler.

being a speaker and hearer. Unlike a typical monologue, dialogues generally consist of many different types of communicative acts: asking questions, giving answers, making corrections, and so forth.

HCI

Finally, computer systems exist and continue to be developed that allow for *human-computer interaction*, or **HCI**. HCI has properties that distinguish it from normal human-human dialogue, in part due to the present-day limitations on the ability of computer systems to participate in free, unconstrained conversation. A system capable of HCI will often employ a strategy to constrain the conversation in ways that allow it to understand the user's utterances within a limited context of interpretation.

While many discourse processing problems are common to these three forms of discourse, they differ in enough respects that different techniques have often been used to process them. This chapter focuses on techniques commonly applied to the interpretation of monologues; techniques for dialogue interpretation and HCI will be described in Chapter 19.

Language is rife with phenomena that operate at the discourse level. Consider the discourse shown in example (17.1).

(17.1) John went to Bill's car dealership to check out an Acura Integra. He looked at it for about an hour.

What do pronouns such as *he* and *it* denote? No doubt that the reader had little trouble figuring out that *he* denotes John and not Bill, and that *it* denotes the Integra and not Bill's car dealership. On the other hand, toward the end of the exchange presented at the beginning of this chapter, it appears that George had some trouble figuring out who Gracie meant when saying *she*.

What differentiates these two examples? How do hearers interpret discourse (17.1) with such ease? Can we build a computational model of this process? These are the types of questions we address in this chapter. In Section 17.1, we describe methods for interpreting *referring expressions* such as pronouns. We then address the problem of establishing the *coherence* of a discourse in Section 17.2. Finally, in Section 17.3 we explain methods for determining the *structure* of a discourse.

Because discourse-level phenomena are ubiquitous in language, algorithms for resolving them are essential for a wide range of language applications. For instance, interactions with query interfaces and dialogue interpretation systems like ATIS (see Chapter 9) frequently contain pronouns and similar types of expressions. So when a user spoke passage (17.2) to an ATIS system,

(17.2) I'd like to get from Boston to San Francisco, on either December 5th or December 6th. It's okay if it stops in another city along the way.

the system had to figure out that *it* denotes the flight that the user wants to book in order to perform the appropriate action.

Similarly, information extraction systems (see Chapter 15) must frequently extract information from utterances that contain pronouns. For instance, if an information extraction system is confronted with passage (17.3),

(17.3) First Union Corp is continuing to wrestle with severe problems unleashed by a botched merger and a troubled business strategy. According to industry insiders at Paine Webber, their president, John R. Georgius, is planning to retire by the end of the year.

it must correctly identify *First Union Corp* as the denotation of *their* (as opposed to *Paine Webber*, for instance) in order to extract the correct event.

Likewise, many text summarization systems employ a procedure for selecting the important sentences from a source document and using them to form a summary. Consider, for example, a news article that contains passage (17.3). Such a system might determine that the second sentence is important enough to be included in the summary, but not the first. However, the second sentence contains a pronoun that is dependent on the first sentence, so it cannot place the second sentence in the summary without first determining the pronoun's denotation, as the pronoun would otherwise likely receive a different interpretation within the summary. Similarly, natural language generation systems (see Chapter 20) must have adequate models for pronominalization to produce coherent and interpretable discourse. In short, just about any conceivable language processing application requires methods for determining the denotations of pronouns and related expressions.

## 17.1 Reference Resolution

In this section we study the problem of **reference**, the process by which speakers use expressions like *John* and *he* in passage (17.1) to denote a person named John. Our discussion requires that we first define some terminology. A natural language expression used to perform reference is called a **referring expression**, and the entity that is referred to is called the **referent**. Thus, *John* and *he* in passage (17.1) are referring expressions, and John is their referent. (To distinguish between referring expressions and their referents, we italicize the former.) As a convenient shorthand, we will sometimes

REFERENCE

REFERRING EXPRESSION

REFERENT

speak of a referring expression referring to a referent, e.g., we might say that *he* refers to John. However, the reader should keep in mind that what we really mean is that the speaker is performing the act of referring to John by uttering *he*. Two referring expressions that are used to refer to the same entity are said to **corefer**, thus *John* and *he* corefer in passage (17.1). There is also a term for a referring expression that licenses the use of another, in the way that the mention of *John* allows John to be subsequently referred to using *he*. We call *John* the **antecedent** of *he*. Reference to an entity that has been previously introduced into the discourse is called **anaphora**, and the referring expression used is said to be **anaphoric**. In passage (17.1), the pronouns *he* and *it* are therefore anaphoric.

COREFER

ANTECEDENT

ANAPHORA

ANAPHORIC

Natural languages provide speakers with a variety of ways to refer to entities. Say that your friend has an Acura Integra automobile and you want to refer to it. Depending on the operative **discourse context**, you might say *it*, *this*, *that*, *this car*, *that car*, *the car*, *the Acura*, *the Integra*, or *my friend's car*, among many other possibilities. However, you are not free to choose between any of these alternatives in any context. For instance, you cannot simply say *it* or *the Acura* if the hearer has no prior knowledge of your friend's car, it has not been mentioned before, and it is not in the immediate surroundings of the discourse participants (i.e., the **situational context** of the discourse).

DISCOURSE CONTEXT

SITUATIONAL CONTEXT

The reason for this is that each type of referring expression encodes different signals about the place that the speaker believes the referent occupies within the hearer's set of beliefs. A subset of these beliefs that has a special status form the hearer's mental model of the ongoing discourse, which we call a **discourse model** (Webber, 1978). The discourse model contains representations of the entities that have been referred to in the discourse and the relationships in which they participate. Thus, there are two components required by a system to successfully produce and interpret referring expressions: a method for constructing a discourse model that evolves with the dynamically-changing discourse it represents, and a method for mapping between the signals that various referring expressions encode and the hearer's set of beliefs, the latter of which includes this discourse model.

DISCOURSE MODEL

We will speak in terms of two fundamental operations to the discourse model. When a referent is first mentioned in a discourse, we say that a representation for it is **evoked** into the model. Upon subsequent mention, this representation is **accessed** from the model. The operations and relationships are illustrated in Figure 17.1.

EVOKED

ACCESSED

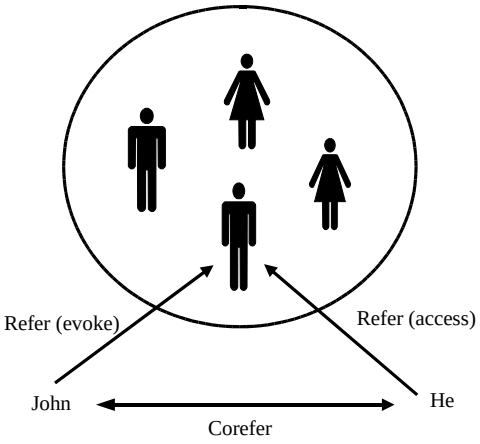


**Figure 17.1** Reference operations and relationships.

We will restrict our discussion to reference to entities, although discourses include reference to many other types of referents. Consider the possibilities in example (17.4), adapted from Webber (1991).

(17.4) According to John, Bob bought Sue an Integra, and Sue bought Fred a Legend.

a. But *that* turned out to be a lie.
b. But *that* was false.
c. *That* struck me as a funny way to describe the situation.
d. *That* caused Sue to become rather poor.
e. *That* caused them both to become rather poor.

The referent of *that* is a speech act (see Chapter 19) in (17.4a), a proposition in (17.4b), a manner of description in (17.4c), an event in (17.4d), and a combination of several events in (17.4e). The field awaits the development of robust methods for interpreting these types of reference.

## Reference Phenomena

The set of referential phenomena that natural languages provide is quite rich indeed. In this section, we provide a brief description of several basic reference phenomena. We first survey five types of referring expression: *indefinite noun phrases*, *definite noun phrases*, *pronouns*, *demonstratives*, and

*one-anaphora*. We then describe three types of referents that complicate the reference resolution problem: *inferrables*, *discontinuous sets*, and *generics*.

**Indefinite Noun Phrases** Indefinite reference introduces entities that are new to the hearer into the discourse context. The most common form of indefinite reference is marked with the determiner *a* (or *an*), as in (17.5), but it can also be marked by a quantifier such as *some* (17.6) or even the determiner *this* (17.7).

(17.5) I saw *an Acura Integra* today.

(17.6) *Some Acura Integras* were being unloaded at the local dealership today.

(17.7) I saw *this awesome Acura Integra* today.

Such noun phrases evoke a representation for a new entity that satisfies the given description into the discourse model.

The indefinite determiner *a* does not indicate whether the entity is identifiable to the speaker, which in some cases leads to a *specific*/*non-specific* ambiguity. Example (17.5) only has the specific reading, since the speaker has a particular Integra in mind, particularly the one she saw. In sentence (17.8), on the other hand, both readings are possible.

(17.8) I am going to the dealership to buy an Acura Integra today.

That is, the speaker may already have the Integra picked out (specific), or may just be planning to pick one out that is to her liking (nonspecific). The readings may be disambiguated by a subsequent referring expression in some contexts; if this expression is definite then the reading is specific (*I hope they still have it*), and if it is indefinite then the reading is nonspecific (*I hope they have a car I like*). This rule has exceptions, however; for instance definite expressions in certain modal contexts (*I will park it in my garage*) are compatible with the nonspecific reading.

**Definite Noun Phrases** Definite reference is used to refer to an entity that is identifiable to the hearer, either because it has already been mentioned in the discourse context (and thus is represented in the discourse model), it is contained in the hearer's set of beliefs about the world, or the uniqueness of the object is implied by the description itself.

The case in which the referent is identifiable from discourse context is shown in (17.9).

(17.9) I saw an Acura Integra today. *The Integra* was white and needed to be washed.

Examples in which the referent is either identifiable from the hearer's set of beliefs or is inherently unique are shown in (17.10) and (17.11) respectively.

(17.10) *The Indianapolis 500* is the most popular car race in the US.

(17.11) *The fastest car in the Indianapolis 500* was an Integra.

Definite noun phrase reference requires that an entity be accessed from either the discourse model or the hearer's set of beliefs about the world. In the latter case, it also evokes a representation of the referent into the discourse model.

**Pronouns** Another form of definite reference is pronominalization, illustrated in example (17.12).

(17.12) I saw an Acura Integra today. *It* was white and needed to be washed.

The constraints on using pronominal reference are stronger than for full definite noun phrases, requiring that the referent have a high degree of activation or **salience** in the discourse model. Pronouns usually (but not always) refer to entities that were introduced no further than one or two sentences back in the ongoing discourse, whereas definite noun phrases can often refer further back. This is illustrated by the difference between sentences (17.13d) and (17.13d').

SALIENCE

(17.13) a. John went to Bob's party, and parked next to a beautiful Acura Integra.
b. He went inside and talked to Bob for more than an hour.
c. Bob told him that he recently got engaged.
d. ?? He also said that he bought *it* yesterday.
d.' He also said that he bought *the Acura* yesterday.

By the time the last sentence is reached, the Integra no longer has the degree of salience required to allow for pronominal reference to it.

Pronouns can also participate in **cataphora**, in which they are mentioned before their referents are, as in example (17.14).

CATAPHORA

(17.14) Before *he* bought *it*, John checked over the Integra very carefully.

Here, the pronouns *he* and *it* both occur *before* their referents are introduced.

Pronouns also appear in quantified contexts in which they are considered to be **bound**, as in example (17.15).

BOUND

(17.15) Every woman bought *her* Acura at the local dealership.

Under the relevant reading, *her* does not refer to some woman in context, but instead behaves like a variable bound to the quantified expression *every woman*. We will not be concerned with the bound interpretation of pronouns in this chapter.

**Demonstratives** Demonstrative pronouns, like *this* and *that*, behave somewhat differently than simple definite pronouns like *it*. They can appear either alone or as determiners, for instance, *this Acura*, *that Acura*. The choice between two demonstratives is generally associated with some notion of spatial proximity: *this* indicating closeness and *that* signaling distance. Spatial distance might be measured with respect to the discourse participants' situational context, as in (17.16).

(17.16) [John shows Bob an Acura Integra and a Mazda Miata]
Bob (pointing): I like *this* better than *that*.

Alternatively, distance can be metaphorically interpreted in terms of conceptual relations in the discourse model. For instance, consider example (17.17).

(17.17) I bought an Integra yesterday. It's similar to the one I bought five years ago. *That one* was really nice, but I like *this one* even better.

Here, *that one* refers to the Acura bought five years ago (greater temporal distance), whereas *this one* refers to the one bought yesterday (closer temporal distance).

**One Anaphora** *One*-anaphora, exemplified in (17.18), blends properties of definite and indefinite reference.

(17.18) I saw no less than 6 Acura Integras today. Now I want *one*.

This use of *one* can be roughly paraphrased by *one of them*, in which *them* refers to a plural referent (or generic one, as in the case of (17.18), see below), and *one* selects a member from this set (Webber, 1983). Thus, *one* may evoke a new entity into the discourse model, but it is necessarily dependent on an existing referent for the description of this new entity.

This use of *one* should be distinguished from the formal, non-specific pronoun usage in (17.19), and its meaning as the number one in (17.20).

(17.19) One shouldn't pay more than twenty thousand dollars for an Acura.

(17.20) John has two Acuras, but I only have one.

**Inferrables** Now that we have described several types of referring expressions, we now turn our attention to a few interesting types of referents that

complicate the reference resolution problem. First, we consider cases in which a referring expression does not refer to an entity that has been explicitly evoked in the text, but instead one that is inferentially related to an evoked entity. Such referents are called *inferrables* (Haviland and Clark, 1974; Prince, 1981). Consider the expressions *a door* and *the engine* in sentence (17.21).

(17.21) I almost bought an Acura Integra today, but *a door* had a dent and *the engine* seemed noisy.

The indefinite noun phrase *a door* would normally introduce a new door into the discourse context, but in this case the hearer is to infer something more: that it is not just any door, but one of the doors of the Integra. Similarly, the use of the definite noun phrase *the engine* normally presumes that an engine has been previously evoked or is otherwise uniquely identifiable. Here, no engine has been explicitly mentioned, but the hearer infers that the referent is the engine of the previously mentioned Integra.

Inferrables can also specify the results of processes described by utterances in a discourse. Consider the possible follow-ons (a-c) to sentence (17.22) in the following recipe (from Webber and Baldwin (1992)):

(17.22) Mix the flour, butter, and water.
a. Kneed *the dough* until smooth and shiny.
b. Spread *the paste* over the blueberries.
c. Stir *the batter* until all lumps are gone.

Any of the expressions *the dough* (a solid), *the batter* (a liquid), and *the paste* (somewhere in between) can be used to refer to the result of the actions described in the first sentence, but all imply different properties of this result.

**Discontinuous Sets** In some cases, references using plural referring expressions like *they* and *them* (see page 10) refer to sets of entities that are evoked together, for instance, using another plural expression (*their Acuras*) or a conjoined noun phrase (*John and Mary*):

(17.23) John and Mary love their Acuras. *They* drive *them* all the time.

However, plural references may also refer to sets of entities that have been evoked by discontinuous phrases in the text:

(17.24) John has an Acura, and Mary has a Mazda. *They* drive *them* all the time.

Here, *they* refers to John and Mary, and likewise *them* refers to the Acura and the Mazda. Note also that the second sentence in this case will generally receive what is called a *pairwise* or *respectively* reading, in which John

drives the Acura and Mary drives the Mazda, as opposed to the reading in which they both drive both cars.

**Generics** Making the reference problem even more complicated is the existence of *generic* reference. Consider example (17.25).

(17.25) I saw no less than 6 Acura Integras today. *They* are the coolest cars.

Here, the most natural reading is not the one in which *they* refers to the particular 6 Integras mentioned in the first sentence, but instead to the class of Integras in general.

### Syntactic and Semantic Constraints on Coreference

Having described a variety of reference phenomena that are found in natural language, we can now consider how one might develop algorithms for identifying the referents of referential expressions. One step that needs to be taken in any successful reference resolution algorithm is to filter the set of possible referents on the basis of certain relatively hard-and-fast constraints. We describe some of these constraints here.

**Number Agreement** Referring expressions and their referents must agree in number; for English, this means distinguishing between *singular* and *plural* references. A categorization of pronouns with respect to number is shown in Figure 17.2.

| Singular | Plural | Unspecified |
|---|---|---|
| she, her, he, him, his, it | we, us, they, them | you |

**Figure 17.2** Number agreement in the English pronominal system.

The following examples illustrate constraints on number agreement.

(17.26) John has a new Acura. It is red.

(17.27) John has three new Acuras. They are red.

(17.28) * John has a new Acura. They are red.

(17.29) * John has three new Acuras. It is red.

**Person and Case Agreement** English distinguishes between three forms of person: first, second, and third. A categorization of pronouns with respect to person is shown in Figure 17.3.

The following examples illustrate constraints on person agreement.

(17.30) You and I have Acuras. We love them.

| | First | Second | Third |
|---|---|---|---|
| Nominative | I, we | you | he, she, they |
| Accusative | me, us | you | him, her, them |
| Genitive | my, our | your | his, her, their |

**Figure 17.3** Person and case agreement in the English pronominal system.

(17.31) John and Mary have Acuras. They love them.

(17.32) * John and Mary have Acuras. We love them. (where *We*=John and Mary)

(17.33) * You and I have Acuras. They love them. (where *They*=You and I)

In addition, English pronouns are constrained by case agreement; different forms of the pronoun may be required when placed in subject position (nominative case, e.g., *he, she, they*), object position (accusative case, e.g., *him, her, them*), and genitive position (genitive case, e.g., *his Acura, her Acura, their Acura*). This categorization is also shown in Figure 17.3.

**Gender Agreement** Referents also must agree with the gender specified by the referring expression. English third person pronouns distinguish between *male*, *female*, and *nonpersonal* genders, and unlike some languages, the first two only apply to animate entities. Some examples are shown in Figure 17.4.

| masculine | feminine | nonpersonal |
|---|---|---|
| he, him, his | she, her | it |

**Figure 17.4** Gender agreement in the English pronominal system.

The following examples illustrate constraints on gender agreement.

(17.34) John has an Acura. He is attractive. (he=John, not the Acura)

(17.35) John has an Acura. It is attractive. (it=the Acura, not John)

**Syntactic Constraints** Reference relations may also be constrained by the syntactic relationships between a referential expression and a possible antecedent noun phrase when both occur in the same sentence. For instance, the pronouns in all of the following sentences are subject to the constraints indicated in brackets.

(17.36) John bought himself a new Acura. [himself=John]

(17.37) John bought him a new Acura. [him$\neq$John]

(17.38) John said that Bill bought him a new Acura. [him≠Bill]

(17.39) John said that Bill bought himself a new Acura. [himself=Bill]

(17.40) He said that he bought John a new Acura. [He≠John; he≠John]

REFLEXIVES

English pronouns such as *himself*, *herself*, and *themselves* are called **reflexives**. Oversimplifying the situation considerably, a reflexive corefers with the subject of the most immediate clause that contains it (ex. 17.36), whereas a nonreflexive cannot corefer with this subject (ex. 17.37). That this rule applies only for the subject of the most immediate clause is shown by examples (17.38) and (17.39), in which the opposite reference pattern is manifest between the pronoun and the subject of the higher sentence. On the other hand, a full noun phrase like *John* cannot corefer with the subject of the most immediate clause nor with a higher-level subject (ex. 17.40).

Whereas these syntactic constraints apply to a referring expression and a particular potential antecedent noun phrase, these constraints actually prohibit coreference between the two regardless of any other available antecedents that denote the same entity. For instance, normally a nonreflexive pronoun like *him* can corefer with the subject of the previous sentence as it does in example (17.41), but it cannot in example (17.42) because of its syntactic relationship with the coreferential pronoun *he* in the second clause.

(17.41) John wanted a new car. Bill bought him a new Acura. [him=John]

(17.42) John wanted a new car. He bought him a new Acura. [he=John; him≠John]

These rules oversimplify the situation in a number of ways, and there are many cases that they do not cover. Indeed, upon further inspection the facts actually get quite complicated. In fact, it is unlikely that all of the data can be explained using only syntactic relations (Kuno, 1987). For instance, the reflexive *himself* and the nonreflexive *him* in sentences (17.43) and (17.44) respectively can both refer to the subject *John*, even though they occur in identical syntactic configurations.

(17.43) John set the pamphlets about Acuras next to himself. [himself=John]

(17.44) John set the pamphlets about Acuras next to him. [him=John]

For the algorithms discussed later in this chapter, however, we will assume a syntactic account of restrictions on intrasentential coreference.

**Selectional Restrictions** The selectional restrictions that a verb places on its arguments (see Chapter 16) may be responsible for eliminating referents, as in example (17.45).

(17.45) John parked his Acura in the garage. He had driven it around for hours.

There are two possible referents for *it*, the Acura and the garage. The verb *drive*, however, requires that its direct object denote something that can be driven, such as a car, truck, or bus, but not a garage. Thus, the fact that the pronoun appears as the object of *drive* restricts the set of possible referents to the Acura. It is conceivable that a practical NLP system would include a reasonably comprehensive set of selectional constraints for the verbs in its lexicon.

Selectional restrictions can be violated in the case of metaphor (see Chapter 16); for example, consider example (17.46).

(17.46) John bought a new Acura. It drinks gasoline like you would not believe.

While the verb *drink* does not usually take an inanimate subject, its metaphorical use here allows *it* to refer to *a new Acura*.

Of course, there are more general semantic constraints that may come into play, but these are much more difficult to encode in a comprehensive manner. Consider passage (17.47).

(17.47) John parked his Acura in the garage. It is incredibly messy, with old bike and car parts lying around everywhere.

Here the referent of *it* is almost certainly the garage, due in part to the fact that a car is probably too small to have bike and car parts laying around "everywhere". Resolving this reference requires that a system have knowledge about how large cars typically are, how large garages typically are, and the typical types of objects one might find in each. On the other hand, one's knowledge about Beverly Hills might lead one to assume that the Acura is indeed the referent of *it* in passage (17.48).

(17.48) John parked his Acura in downtown Beverly Hills. It is incredibly messy, with old bike and car parts lying around everywhere.

In the end, just about any knowledge shared by the discourse participants might be necessary to resolve a pronoun reference. However, due in part to the vastness of such knowledge, practical algorithms typically do not rely on it heavily.

### Preferences in Pronoun Interpretation

In the previous section, we discussed relatively strict constraints that algorithms should apply when determining possible referents for referring ex-

pressions. We now discuss some more readily violated *preferences* that algorithms can be made to account for. These preferences have been posited to apply to pronoun interpretation in particular. Since the majority of work on reference resolution algorithms has focused on pronoun interpretation, we will similarly focus on this problem in the remainder of this section.

**Recency** Most theories of reference incorporate the notion that entities introduced in recent utterances are more salient than those introduced from utterances further back. Thus, in example (17.49), the pronoun *it* is more likely to refer to the Legend than the Integra.

(17.49) John has an Integra. Bill has a Legend. Mary likes to drive it.

**Grammatical Role** Many theories specify a salience hierarchy of entities that is ordered by the grammatical position of the expressions which denote them. These typically treat entities mentioned in subject position as more salient than those in object position, which are in turn more salient than those mentioned in subsequent positions.

Passages such as (17.50) and (17.51) lend support for such a hierarchy. Although the first sentence in each case expresses roughly the same propositional content, the preferred referent for the pronoun *him* varies with the subject in each case – John in (17.50) and Bill in (17.51). In example (17.52), the references to John and Bill are conjoined within the subject position. Since both seemingly have the same degree of salience, it is unclear to which the pronoun refers.

(17.50) John went to the Acura dealership with Bill. He bought an Integra. [ he = John ]

(17.51) Bill went to the Acura dealership with John. He bought an Integra. [ he = Bill ]

(17.52) John and Bill went to the Acura dealership. He bought an Integra. [ he = ?? ].

**Repeated Mention** Some theories incorporate the idea that entities that have been focused on in the prior discourse are more likely to continue to be focused on in subsequent discourse, and hence references to them are more likely to be pronominalized. For instance, whereas the pronoun in example (17.51) has Bill as its preferred interpretation, the pronoun in the final sentence of example (17.53) is more likely to refer to John.

(17.53) John needed a car to get to his new job. He decided that he wanted something sporty. Bill went to the Acura dealership with him. He bought an Integra. [ he = John ]

**Parallelism** There are also strong preferences that appear to be induced by parallelism effects, as in example (17.54).

(17.54) Mary went with Sue to the Acura dealership. Sally went with her to the Mazda dealership. [ her = Sue ]

The grammatical role hierarchy described above ranks Mary as more salient than Sue, and thus should be the preferred referent of *her*. Furthermore, there is no semantic reason that Mary cannot be the referent. Nonetheless, *her* is instead understood to refer to Sue.

This suggests that we might want a heuristic which says that non-subject pronouns prefer non-subject referents. However, such a heuristic may not work for cases that lack the structural parallelism of example (17.54), such as example (17.55), in which Mary is the preferred referent of the pronoun instead of Sue.

(17.55) Mary went with Sue to the Acura dealership. Sally told her not to buy anything. [ her = Mary ]

**Verb Semantics** Certain verbs appear to place a semantically-oriented emphasis on one of their argument positions, which can have the effect of biasing the manner in which subsequent pronouns are interpreted. Compare sentences (17.56) and (17.57).

(17.56) John telephoned Bill. He lost the pamphlet on Acuras.

(17.57) John criticized Bill. He lost the pamphlet on Acuras.

These examples differ only in the verb used in the first sentence, yet the subject pronoun in passage (17.56) is typically resolved to John, whereas the pronoun in passage (17.57) is resolved to Bill. Some researchers have claimed that this effect results from what has been called the "implicit causality" of a verb: the implicit cause of a "criticizing" event is considered to be its object, whereas the implicit cause of a "telephoning" event is considered to be its subject. This emphasis results in a higher degree of salience for the entity in this argument position, which leads to the different preferences for examples (17.56) and (17.57).

Similar preferences have been articulated in terms of the thematic roles (see Chapter 16) that the potential antecedents occupy. For example, most hearers resolve *He* to John in example (17.58) and to Bill in example (17.59). Although these referents are evoked from different grammatical role positions, they both fill the Goal thematic role of their corresponding verbs, whereas the other potential referent fills the Source role. Likewise, hearers

generally resolve *He* to John and Bill in examples (17.60) and (17.61) respectively, providing evidence that fillers of the Stimulus role are preferred over fillers of the Experiencer role.

(17.58) John seized the Acura pamphlet from Bill. He loves reading about cars. (Goal=John, Source=Bill)

(17.59) John passed the Acura pamphlet to Bill. He loves reading about cars. (Goal=Bill, Source=John)

(17.60) The car dealer admired John. He knows Acuras inside and out. (Stimulus=John, Experiencer=the car dealer)

(17.61) The car dealer impressed John. He knows Acuras inside and out. (Stimulus=the car dealer, Experiencer=John)

### An Algorithm for Pronoun Resolution

None of the algorithms for pronoun resolution that have been proposed to date successfully account for all of these preferences, let alone succeed in resolving the contradictions that will arise between them. However, Lappin and Leass (1994) describe a straightforward algorithm for pronoun interpretation that takes many of these into consideration. The algorithm employs a simple weighting scheme that integrates the effects of the recency and syntactically-based preferences; no semantic preferences are employed beyond those enforced by agreement. We describe a slightly simplified portion of the algorithm that applies to non-reflexive, third person pronouns.

Broadly speaking, there are two types of operations performed by the algorithm: discourse model update and pronoun resolution. First, when a noun phrase that evokes a new entity is encountered, a representation for it must be added to the discourse model and a degree of salience (which we call a **salience value**) computed for it. The salience value is calculated as the sum of the weights assigned by a set of **salience factors**. The salience factors used and their corresponding weights are shown in Figure 17.5.

SALIENCE VALUE

SALIENCE FACTORS

The weights that each factor assigns to an entity in the discourse model are cut in half each time a new sentence is processed. This, along with the added effect of the sentence recency weight (which initially assigns a weight of 100, to be cut in half with each succeeding sentence), captures the Recency preference described on page 14, since referents mentioned in the current sentence will tend to have higher weights than those in the previous sentence, which will in turn be higher than those in the sentence before that, and so forth.

| | |
|---|---|
| Sentence recency | 100 |
| Subject emphasis | 80 |
| Existential emphasis | 70 |
| Accusative (direct object) emphasis | 50 |
| Indirect object and oblique complement emphasis | 40 |
| Non-adverbial emphasis | 50 |
| Head noun emphasis | 80 |

**Figure 17.5** Salience factors in Lappin and Leass's system.

Similarly, the next five factors in Figure 17.5 can be viewed as a way of encoding a grammatical role preference scheme using the following hierarchy:

> subject > existential predicate nominal > object > indirect object or oblique > demarcated adverbial PP

These five positions are exemplified by the position of the italicized phrases in examples (17.62)–(17.66) respectively.

(17.62) *An Acura Integra* is parked in the lot. (subject)

(17.63) There is *an Acura Integra* parked in the lot. (existential predicate nominal)

(17.64) John parked *an Acura Integra* in the lot. (object)

(17.65) John gave *his Acura Integra* a bath. (indirect object)

(17.66) Inside *his Acura Integra*, John showed Susan his new CD player. (demarcated adverbial PP)

The preference against referents in demarcated adverbial PPs (i.e., those separated by punctuation, as with the comma in example (17.66)) is encoded as a positive weight of 50 for every other position, listed as the non-adverbial emphasis weight in Figure 17.5. This ensures that the weight for any referent is always positive, which is necessary so that the effect of halving the weights is always to reduce them.

The head noun emphasis factor penalizes referents which are embedded in larger noun phrases, again by promoting the weights of referents that are not. Thus, the Acura Integra in each of examples (17.62)–(17.66) will receive 80 points for being denoted by a head noun, whereas the Acura Integra in example (17.67) will not, since it is embedded within the subject noun phrase.

(17.67) The owner's manual for *an Acura Integra* is on John's desk.

Each of these factors contributes to the salience of a referent based on the properties of the noun phrase that denotes it. Of course, it could be that several noun phrases in the preceding discourse refer to the same referent, each being assigned a different level of salience, and thus we need a way in which to combine the contributions of each. To address this, Lappin and Leass associate with each referent an equivalence class that contains all of the noun phrases that have been determined to refer to it. The weight that a salience factor assigns to a referent is the highest of the weights it assigns to the members of its equivalence class. The salience weight for a referent is then calculated by summing the weights for each factor. The scope of a salience factor is a sentence, so, for instance, if a potential referent is mentioned in the current sentence as well as the previous one, the sentence recency weight will be factored in for each. (On the other hand, if the same referent is mentioned more than once in the same sentence, this weight will be counted only once.) Thus, multiple mentions of a referent in the prior discourse can potentially increase its salience, which has the effect of encoding the preference for repeated mentions discussed on page 14.

Once we have updated the discourse model with new potential referents and recalculated the salience values associated with them, we are ready to consider the process of resolving any pronouns that exist within a new sentence. In doing this, we factor in two more salience weights, one for grammatical role parallelism between the pronoun and the potential referent, and one to disprefer cataphoric reference. The weights are shown in Figure 17.6. Unlike the other preferences, these two cannot be calculated independently of the pronoun, and thus cannot be calculated during the discourse model update step. We will use the term *initial salience value* for the weight of a given referent before these factors are applied, and the term *final salience value* for after they have applied.

| | |
|---|---|
| Role Parallelism | 35 |
| Cataphora | -175 |

**Figure 17.6** Per pronoun salience weights in Lappin and Leass's system.

We are now ready to specify the pronoun resolution algorithm. Assuming that the discourse model has been updated to reflect the initial salience values of referents as described above, the steps taken to resolve a pronoun are as follows:

1. Collect the potential referents (up to four sentences back).

2. Remove potential referents that do not agree in number or gender with the pronoun.
3. Remove potential referents that do not pass intrasentential syntactic coreference constraints (as described on page 11).
4. Compute the total salience value of the referent by adding any applicable values from Figure 17.6 to the existing salience value previously computed during the discourse model update step (i.e., the sum of the applicable values in Figure 17.5).
5. Select the referent with the highest salience value. In the case of ties, select the closest referent in terms of string position (computed without bias to direction).

We illustrate the operation of the algorithm by stepping through example (17.68).

(17.68) John saw a beautiful Acura Integra at the dealership. He showed it to Bob. He bought it.

We first process the first sentence to collect potential referents and compute their initial salience values. The following table shows the contribution to salience of each of the salience factors.

| | Rec | Subj | Exist | Obj | Ind-Obj | Non-Adv | Head N | Total |
|---|---|---|---|---|---|---|---|---|
| John | 100 | 80 | | | | 50 | 80 | 310 |
| Integra | 100 | | | 50 | | 50 | 80 | 280 |
| dealership | 100 | | | | | 50 | 80 | 230 |

There are no pronouns to be resolved in this sentence, so we move on to the next, degrading the above values by a factor of two as shown in the table below. The *phrases* column shows the equivalence class of referring expressions for each referent.

| Referent | Phrases | Value |
|---|---|---|
| John | { *John* } | 155 |
| Integra | { *a beautiful Acura Integra* } | 140 |
| dealership | { *the dealership* } | 115 |

The first noun phrase in the second sentence is the pronoun *he*. Because *he* specifies male gender, Step 2 of the resolution algorithm reduces the set of possible referents to include only John, so we can stop there and take this to be the referent.

The discourse model must now be updated. First, the pronoun *he* is added in the equivalence class for John (denoted as $he_1$, to differentiate it from possible other mentions of *he*). Since *he* occurs in the current sentence and *John* in the previous one, the salience factors do not overlap between the two. The pronoun is in the current sentence (recency=100), subject position (=80), not in an adverbial (=50), and not embedded (=80), and so a total of 310 is added to the current weight for John:

| Referent | Phrases | Value |
|---|---|---|
| John | { *John*, $he_1$ } | 465 |
| Integra | { *a beautiful Acura Integra* } | 140 |
| dealership | { *the dealership* } | 115 |

The next noun phrase in the second sentence is the pronoun *it*, which is compatible with the Integra or the dealership. We first need to compute the final salience values by adding the applicable weights from Figure 17.6 to the initial salience values above. Neither referent assignment would result in cataphora, so that factor does not apply. For the parallelism preference, both *it* and *a beautiful Acura Integra* are in object position within their respective sentences (whereas *the dealership* is not), so a weight of 35 is added to this option. With the Integra having a weight of 175 and the dealership a weight of 115, the Integra is taken to be the referent.

Again, the discourse model must now be updated. Since *it* is in a nonembedded object position, it receives a weight of 100+50+50+80=280, and is added to the current weight for the Integra.

| Referent | Phrases | Value |
|---|---|---|
| John | { *John*, $he_1$ } | 465 |
| Integra | { *a beautiful Acura Integra*, $it_1$ } | 420 |
| dealership | { *the dealership* } | 115 |

The final noun phrase in the second sentence is *Bob*, which introduces a new discourse referent. Since it occupies an oblique argument position, it receives a weight of 100+40+50+80=270.

| Referent | Phrases | Value |
|---|---|---|
| John | { *John*, $he_1$ } | 465 |
| Integra | { *a beautiful Acura Integra*, $it_1$ } | 420 |
| Bob | { *Bob* } | 270 |
| dealership | { *the dealership* } | 115 |

Now we are ready to move on to the final sentence. We again degrade the current weights by one half.

| Referent | Phrases | Value |
|---|---|---|
| John | { *John*, $he_1$ } | 232.5 |
| Integra | { *a beautiful Acura Integra*, $it_1$ } | 210 |
| Bob | { *Bob* } | 135 |
| dealership | { *the dealership* } | 57.5 |

The reader can confirm that the referent of *he* will be resolved to John, and the referent of *it* to the Integra.

The weights used by Lappin and Leass were arrived at by experimentation on a development corpus of computer training manuals. This algorithm, when combined with several filters not described here, achieved 86% accuracy when applied to unseen test data within the same genre. It is possible that these exact weights may not be optimal for other genres (and even more so for other languages), so the reader may want to experiment with these on training data for a new application or language.

In Exercise 17.7, we consider a version of the algorithm that relies only on a noun phrase identifier (see also Kennedy and Boguraev (1996)). In the next paragraphs, we briefly summarize two other approaches to pronoun resolution.

**A Tree Search Algorithm** Hobbs (1978) describes an algorithm for pronoun resolution which takes the syntactic representations of the sentences up to and including the current sentence as input, and performs a search for an antecedent noun phrase on these trees. There is no explicit representation of a discourse model or preferences as in the Lappin and Leass algorithm. However, certain of these preferences are approximated by the order in which the search on syntactic trees is performed.

An algorithm that searches parse trees must also specify a grammar, since the assumptions regarding the structure of syntactic trees will affect the results. A fragment for English that the algorithm uses is given in Figure 17.7. The steps of the algorithm are as follows:

1. Begin at the noun phrase (NP) node immediately dominating the pronoun.
2. Go up the tree to the first NP or sentence (S) node encountered. Call this node X, and call the path used to reach it *p*.
3. Traverse all branches below node X to the left of path *p* in a left-to-right, breadth-first fashion. Propose as the antecedent any NP node that is encountered which has an NP or S node between it and X.
4. If node X is the highest S node in the sentence, traverse the surface parse trees of previous sentences in the text in order of recency, the

$$S \rightarrow NP\ VP$$
$$NP \rightarrow \left\{ \begin{array}{l} (Det)\quad Nominal\quad \left( \left\{ \begin{array}{l} PP \\ Rel \end{array} \right\} \right)^* \\ pronoun \end{array} \right\}$$
$$Det \rightarrow \left\{ \begin{array}{l} determiner \\ NP\ 's \end{array} \right\}$$
$$PP \rightarrow preposition\ NP$$
$$Nominal \rightarrow noun\ (PP)^*$$
$$Rel \rightarrow wh\text{-}word\ S$$
$$VP \rightarrow verb\ NP\ (PP)^*$$

**Figure 17.7** A grammar fragment for the Tree Search algorithm.

most recent first; each tree is traversed in a left-to-right, breadth-first manner, and when an NP node is encountered, it is proposed as antecedent. If X is not the highest S node in the sentence, continue to step 5.

5. From node X, go up the tree to the first NP or S node encountered. Call this new node X, and call the path traversed to reach it *p*.
6. If X is an NP node and if the path *p* to X did not pass through the Nominal node that X immediately dominates, propose X as the antecedent.
7. Traverse all branches below node X to the *left* of path p in a left-to-right, breadth-first manner. Propose any NP node encountered as the antecedent.
8. If X is an S node, traverse all branches of node X to the *right* of path *p* in a left-to-right, breadth-first manner, but do not go below any NP or S node encountered. Propose any NP node encountered as the antecedent.
9. Go to Step 4.

Demonstrating that this algorithm yields the correct coreference assignments for example (17.68) is left as Exercise 17.3.

As stated, the algorithm depends on complete and correct syntactic structures as input. Hobbs evaluated his approach manually (with respect to both parse construction and algorithm implementation) on one hundred examples from each of three different texts, reporting an accuracy of 88.3%. (The accuracy increases to 91.7% if certain selectional restriction constraints are assumed.) Lappin and Leass encoded a version of this algorithm within their system, and reported an accuracy of 82% on their test corpus. Although

this is less than the 86% accuracy achieved by their own algorithm, it should be borne in mind that the test data Lappin and Leass used was from the same genre as their development set, but different than the genres that Hobbs used in developing his algorithm.

**A Centering Algorithm** As we described above, the Hobbs algorithm does not use an explicit representation of a discourse model. The Lappin and Leass algorithm does, but encodes salience as a weighted combination of preferences. Centering theory (Grosz et al., 1995, henceforth GJW), also has an explicit representation of a discourse model, and incorporates an additional claim: that there is a single entity being "centered" on at any given point in the discourse which is to be distinguished from all other entities that have been evoked.

There are two main representations tracked in the discourse model. In what follows, take $U_n$ and $U_{n+1}$ to be two adjacent utterances. The *backward looking center* of $U_n$, denoted as $C_b(U_n)$, represents the entity currently being focused on in the discourse after $U_n$ is interpreted. The *forward looking centers* of $U_n$, denoted as $C_f(U_n)$, form an ordered list containing the entities mentioned in $U_n$, all of which could serve as the $C_b$ of the following utterance. In fact, $C_b(U_{n+1})$ is by definition the most highly ranked element of $C_f(U_n)$ mentioned in $U_{n+1}$. (The $C_b$ of the first utterance in a discourse is undefined.) As for how the entities in the $C_f(U_n)$ are ordered, for simplicity's sake we can use the grammatical role hierarchy encoded by (a subset of) the weights in the Lappin and Leass algorithm, repeated below.[1]

> subject $>$ existential predicate nominal $>$ object $>$ indirect object or oblique $>$ demarcated adverbial PP

Unlike the Lappin and Leass algorithm, however, there are no numerical weights attached to the entities on the list, they are simply ordered relative to each other. As a shorthand, we will call the highest-ranked forward-looking center $C_p$ (for "preferred center").

We describe a centering-based algorithm for pronoun interpretation due to Brennan et al. (1987, henceforth BFP). (See also Walker et al. (1994); for other centering algorithms, see Kameyama (1986) and Strube and Hahn (1996), inter alia.) In this algorithm, preferred referents of pronouns are computed from relations that hold between the forward and backward looking centers in adjacent sentences. Four intersentential relationships between a pair of utterances $U_n$ and $U_{n+1}$ are defined which depend on the relationship between $C_b(U_{n+1})$, $C_b(U_n)$, and $C_p(U_{n+1})$; these are shown in Figure 17.8.

[1] This is an extended form of the hierarchy used in Brennan et al. (1987), described below.

| | $C_b(U_{n+1}) = C_b(U_n)$ or undefined $C_b(U_n)$ | $C_b(U_{n+1}) \neq C_b(U_n)$ |
|---|---|---|
| $C_b(U_{n+1}) = C_p(U_{n+1})$ | Continue | Smooth-Shift |
| $C_b(U_{n+1}) \neq C_p(U_{n+1})$ | Retain | Rough-Shift |

**Figure 17.8** Transitions in the BFP algorithm.

The following rules are used by the algorithm:

- Rule 1: If any element of $C_f(U_n)$ is realized by a pronoun in utterance $U_{n+1}$, then $C_b(U_{n+1})$ must be realized as a pronoun also.
- Rule 2: Transition states are ordered. Continue is preferred to Retain is preferred to Smooth-Shift is preferred to Rough-Shift.

Having defined these concepts and rules, the algorithm is defined as follows.

1. Generate possible $C_b$-$C_f$ combinations for each possible set of reference assignments .
2. Filter by constraints, e.g., syntactic coreference constraints, selectional restrictions, centering rules and constraints.
3. Rank by transition orderings.

The pronominal referents that get assigned are those which yield the most preferred relation in Rule 2, assuming that Rule 1 and other coreference constraints (gender, number, syntactic, selectional restrictions) are not violated.

Let us step through passage (17.68), repeated below as (17.69), to illustrate the algorithm.

(17.69) John saw a beautiful Acura Integra at the dealership. ($U_1$)
He showed it to Bob. ($U_2$)
He bought it. ($U_3$)

Using the grammatical role hierarchy to order the $C_f$, for sentence $U_1$ we get:

$C_f(U_1)$: {John, Integra, dealership}
$C_p(U_1)$: John
$C_b(U_1)$: undefined

Sentence $U_2$ contains two pronouns: *he*, which is compatible with John, and *it*, which is compatible with the Acura or the dealership. John is by definition $C_b(U_2)$, because he is the highest ranked member of $C_f(U_1)$ mentioned in $U_2$

(since he is the only possible referent for *he*). We compare the resulting transitions for each possible referent of *it*. If we assume *it* refers to the Integra, the assignments would be:

$C_f(U_2)$: {John, Integra, Bob}
$C_p(U_2)$: John
$C_b(U_2)$: John
Result: Continue ($C_p(U_2)$=$C_b(U_2)$; $C_b(U_1)$ undefined)

If we assume *it* refers to the dealership, the assignments would be:

$C_f(U_2)$: {John, dealership, Bob}
$C_p(U_2)$: John
$C_b(U_2)$: John
Result: Continue ($C_p(U_2)$=$C_b(U_2)$; $C_b(U_1)$ undefined)

Since both possibilities result in a Continue transition, the algorithm does not say which to accept. For the sake of illustration, we will assume that ties are broken in terms of the ordering on the previous $C_f$ list. Thus, we will take *it* to refer to the Integra instead of the dealership, leaving the current discourse model as represented in the first possibility above.

In sentence $U_3$, *he* is compatible with either John or Bob, whereas *it* is compatible with the Integra. If we assume *he* refers to John, then John is $C_b(U_3)$ and the assignments would be:

$C_f(U_3)$: {John, Acura}
$C_p(U_3)$: John
$C_b(U_3)$: John
Result: Continue ($C_p(U_3)$=$C_b(U_3)$=$C_b(U_2)$)

If we assume *he* refers to Bob, then Bob is $C_b(U_3)$ and the assignments would be:

$C_f(U_3)$: {Bob, Acura}
$C_p(U_3)$: Bob
$C_b(U_3)$: Bob
Result: Smooth-Shift ($C_p(U_3)$=$C_b(U_3)$; $C_b(U_3)\neq C_b(U_2)$)

Since a Continue is preferred to a Smooth-Shift per Rule 2, John is correctly taken to be the referent.

The main salience factors that the centering algorithm implicitly incorporates include the grammatical role, recency, and repeated mention preferences. Unlike the Lappin and Leass algorithm, however, the manner in

which the grammatical role hierarchy affects salience is indirect, since it is the resulting transition type that determines the final reference assignments. In particular, a referent in a low-ranked grammatical role will be preferred to one in a more highly ranked role if the former leads to a more highly ranked transition. Thus, the centering algorithm may (often, but not always, incorrectly) resolve a pronoun to a referent that other algorithms would consider to be of relatively low salience (Lappin and Leass, 1994; Kehler, 1997a). For instance, in example (17.70),

(17.70) Bob opened up a new dealership last week. John took a look at the Acuras in his lot. He ended up buying one.

the centering algorithm will assign Bob as the referent of the subject pronoun *he* in the third sentence – since Bob is $C_b(U_2)$, this assignment results in a Continue relation whereas assigning John results in a Smooth-Shift relation. On the other hand, the Hobbs and Lappin/Leass algorithms will assign John as the referent.

Like the Hobbs algorithm, the centering algorithm was developed on the assumption that correct syntactic structures are available as input. In order to perform an automatic evaluation on naturally occurring data, the centering algorithm would have to be specified in greater detail, both in terms of how all noun phrases in a sentence are ordered with respect to each other on the $C_f$ list (the current approach only includes nonembedded fillers of certain grammatical roles, generating only a partial ordering), as well as how all pronouns in a sentence can be resolved (e.g., recall the indeterminacy in resolving *it* in the second sentence of example (17.68)).

Walker (1989), however, performed a manual evaluation of the centering algorithm on a corpus of 281 examples distributed over texts from three genres, and compared its performance to the Hobbs algorithm. The evaluation assumed adequate syntactic representations, grammatical role labeling, and selectional restriction information as input. Furthermore, in cases in which the centering algorithm did not uniquely specify a referent, only those cases in which the Hobbs algorithm identified the *correct* one were counted as errors. With this proviso, Walker reports an accuracy of 77.6% for centering and 81.8% for Hobbs. See also Tetreault (1999) for a comparison between several centering-based algorithms and the Hobbs algorithm.

## 17.2 TEXT COHERENCE

Much of the previous section focussed on the nature of anaphoric reference and methods for resolving pronouns in discourse. Anaphoric expressions