Anaphora Resolution as Equality by Default

Ariel Cohen

Ben-Gurion University, Beer Sheva 84105, Israel,
`arikc@bgu.ac.il`,
WWW home page: `http://www.bgu.ac.il/~arikc`

Abstract. The resolution of anaphora is dependent on a number of factors discussed in the literature: syntactic parallelism, topicality, etc. A system that attempts to resolve anaphora will have to represent many of these factors, and deal with their interaction. In addition, there must be a principle that simply says that the system needs to look for an antecedent. Without such a principle, if none of the factors recommend a clear winner, the system will be left without an antecedent. This principle should work in such a way that, if there is exactly one good candidate antecedent, the system will choose it; if there are more than one, the system will still attempt to identify one, or, at least, draw some inferences about the likely antecedent; and, in case there is no candidate, the system will produce an accommodated or deictic reading.

Many systems embody some version of this principle procedurally, as part of the workings of their algorithm. However, because it is not explicitly formalized, it is hard to draw firm conclusions about what the system would do in any given case. In this paper I define a general principle of Equality by Default, formalize it in Default Logic, and demonstrate that it produces the desired behavior. Since all other factors can also be formalized in Default Logic, the principle does not need to be left implicit in the algorithm, and can be integrated seamlessly into the rest of the explicit rules affecting anaphora resolution.

Key words: anaphora, Default Logic, equality, DOAP

1 The Search for an Antecedent

Identifying the antecedent of an anaphoric trigger (a pronoun, definite DP, etc.) depends on the interaction of many factors: syntactic (e.g. Binding Theory), semantic (e.g. selectional restrictions), and pragmatic (e.g. Centering Theory). Some of these factors, such as selectional restrictions and syntactic binding requirements rule out certain antecedents, while other factors, e.g. topicality, suggest that a certain antecedent should be chosen.

Most, perhaps all of these factors are defeasible. Consider, for example, the following discourse, from [1]:

(1) The Vice-President entered the President's office. *He* was nervous and clutching his briefcase. After all, he couldn't fire the Vice-President without making trouble for himself with the chairman of the board.

The pronoun in the second sentence has two potential antecedents: *the Vice-President* or *the President*. Clearly *the Vice-President* is preferred: it has the same syntactic position (subject) as the pronoun, and it is more salient. However, by the time the third sentence is processed, it is clear that this choice is wrong, and the intended antecedent is, in fact, *the President*.

Even what appear to be inviolable constraints, such as number agreement, can sometimes be overruled. For example, [2] note that

> numeric agreement in this corpus of Wall Street Journal articles is a defeasible constraint, because it includes so many mentions of organizations. An organization, such as "Wellington Industries" appears syntactically to be plural, but can be re-mentioned with the pronoun *it*.

Such examples abound; and they indicate that all anaphora resolution factors, or almost all of them, are best thought of as defaults, which may be overridden. It is therefore attractive to model anaphora resolution as a system of prioritized defaults (e.g. [2–5]).

Most such systems do not encode the constraints explicitly, but rather procedurally, as part of the algorithm. There are, however, strong arguments for having a declarative, explicit definition of the constraints, as argued by [2]. They implement a system of constraints for anaphora resolution proposed by [6], formulated in Optimality Theory [7]. They point out that a program that uses an explicit definition of constraints is easy to test, debug, and revise. It is also much easier to modify, say in order to apply it to another genre or another language. If constraints need to be added, removed, or the priorities between them changed, this can be done quickly, reliably, and transparently.

In this paper I am not going to consider the question of identifying these factors or their relative strengths. What I do wish to argue is that formalizing all these factors is not enough, and an additional rule is necessary; I will propose a formalization of this rule in Default Logic [8].

The rest of the paper is organized as follows. The next section contains a discussion of the additional rule: Don't Overlook Anaphoric Possibilities—DOAP [9]. The following two sections provide necessary background for the formalization of DOAP: section 3 discusses DRT as an underspecified representation for anaphora, and the significance of treating anaphoric relations as equality. Section 4 presents a brief overview of Default Logic, to be applied to equality in section 5, where DOAP is formalized as Equality by Default. Section 6 discusses the inferences that can be drawn using this relation, and contains examples demonstrating that they obey the desired patterns. The final section concludes the paper and points out potential additional applications of the theory.

2 Don't Overlook Anaphoric Possibilities

Consider the discourse in (1) again.

The antecedent that is eventually chosen, *the President*, is not suggested by any of the well known factors discussed in the literature: it is neither topical, nor

a subject, nor does it have the same syntactic position as the pronoun, etc. This antecedent is simply chosen as a last resort, since the other potential candidate is ruled out. This "last resort" rule must be defined somehow, for, without it, no antecedent would be chosen. Indeed, in the linguistics literature, such a rule has been proposed [9, p. 603]:

(2)   *Don't Overlook Anaphoric Possibilities* (DOAP)
        Opportunities to anaphorize text must be seized.

Essentially, this rule says simply that, when we encounter a trigger , we must try to find an antecedent. If we find an antecedent that is suggested by some rule, so much the better; but even a dispreferred antecedent is better than no antecedent at all. DOAP has been used by [10], who propose an Optimality Theoretic system of prioritized defaults for anaphora resolution.

However, while factors such as syntactic parallelism or selectional restrictions are, at least conceptually, easy to implement, it is not clear how to formalize DOAP in such a way that it could be implemented. This paper is an attempt to provide such a formalization, which, in combination with other factors, has the potential to bring about a fully explicit system of anaphora resolution.

Of course, in practice almost all anaphora resolution algorithm obey DOAP, in the sense that they always attempt to find (at least) one antecedent, even if the anaphora is ambiguous. However, if DOAP is not defined explicitly in the object level of the logic, but is left to a metalevel description, it is hard to be clear on, let alone prove, what the system will do when there is no clear choice of antecedent: which, if any, antecedent it will choose, and which inferences it will draw. Hence, formalization of DOAP on a par with all other factors is a desirable goal.

Take, for example, systems that use model building techniques. Such systems typically generate minimal models. Minimality could be seen as an implementation of DOAP: A model in which the antecedent of a referring expression is not identified is not minimal (since it has an additional entity, namely the reference of the trigger); it is therefore dispreferred, and the anaphoric reading is chosen, if possible.

However, for many of these systems, the model cannot always be relied upon to be minimal [11]. Even where it can, minimality of the model is not sufficient to ensure that an antecedent is chosen.

Consider, for example, the following discourse:

(3)   John met Mary. *He* didn't talk to *her*.

A model builder would generate a model whose universe consists of John and Mary, and where the denotation of the predicate *talk to* is the empty set. This model satisfies the discourse in (3), and is clearly minimal, yet it says nothing about which antecedents the pronouns refer to.

An explicit formalization of DOAP should be able to deal with cases where there is one clear antecedent, as well as with cases where there isn't. In general, when an anaphoric trigger is encountered, there are three possibilities.

One possibility is that there is exactly one appropriate antecedent:

(4)   John was eating ice cream. *He* was upset.

In this case, John is the only appropriate antecedent, and we would want to resolve the anaphora by equating the pronoun with John.

The second possibility is that there is no appropriate antecedent in the text:

(5)   John was eating ice cream. *The waitress* brought him the check.

The text provides no appropriate antecedent for the definite description, so one must be accommodated. If the anaphoric trigger is a pronoun, whose informational content is minimal, accommodation may be impossible [12]. In this case, the pronoun will be interpreted deictically:

(6)   John was eating ice cream. *She* brought him the check.

In (6) we interpret the pronoun as referring to some individual that is not introduced in the discourse, and is, perhaps, identified by pointing.

The third possibility is that there is more than one good candidate antecedent:

(7)   John and Bill met at the ice cream parlor. *He* was upset.

There are few reasons to prefer either John or Bill as the antecedent of the pronoun. In this case, we have two choices: we can decide on some antecedent, perhaps at random, perhaps using some criterion such as recency; alternatively, we can acknowledge that the anaphora is genuinely ambiguous. Even if we take the latter course of action, all is not lost: although we do not know who the pronoun refers to, we can still draw some conclusions about him. For example, we know that, whoever *he* is, he was at the ice cream parlor.

3   An Underspecified Representation for Anaphora

Before formalizing DOAP, we need to say something about how the relation between trigger and antecedent is represented. Consider a simple case of ambiguous anaphoric reference:

(8)   John shook hands with Bill and Mary. *He* hung out with *her* the whole evening.

What can we say about the resolution of the anaphora? The pronoun *her* probably refers to Mary, and the pronoun *he* is ambiguous between John and Bill, but probably refers to John. And, in the right context and/or intonation, either pronoun (or both) may be used deictically, referring to some other individual that is not denoted by a linguistic antecedent. What we would like is a system that allows us to represent all these options, pick those we consider plausible, and draw some inferences even in the absence of a clear resolution.

As the discourse in (8) exemplifies, anaphora is often ambiguous. Moreover, it is always possible, in principle, that what we had identified as the antecedent

of a trigger actually is not, and we need to get an accommodated or deictic reading. In the case of (8), since we have two pronouns, one with three possible interpretations (John, Bill, or the deictic use) and the other with two (Mary or deictic), we will have six potential interpretations. We need to be able to represent the ambiguity, but still draw inferences as best we can on the basis of what we know. This calls for some sort of underspecified representation, and some inference mechanism to derive conclusions from it.

Many special formalisms have been proposed, whose sole purpose is to allow efficient representation of and reasoning with underspecification. I will not, however, go down this road, for several reasons. A formalism that is not independently motivated on linguistic grounds, and whose sole justification is to represent underspecification, may work in a practical system, but its explanatory adequacy from a linguistic point of view would be dubious.

To give one example, recall that deictic readings of a pronoun are always (given the right intonation and/or context) possible, and this is the case across languages. Why is this? Why don't we have languages where pronouns are restricted to linguistic antecedents only, and deictic readings are indicated only by, say, demonstratives? A formalism that is only geared toward underspecification would be quite adequate even if pronouns could only refer to linguistic antecedents, and it is hard to see why it would necessitate the availability of deictic readings. It is, of course, preferable to have the possibility of deictic readings follow directly from the representation, thus explaining the puzzle.

Furthermore, a nonstandard representation will typically require nonstandard inference methods, especially tailored for the representation.[1] Again, these inference methods would not be independently justified, unlike rules of commonsense inference that must, in one way or another, be used in order to understand natural language.

An additional reason for keeping the representation as simple and as close to standard linguistic representations as possible is the fact that it is not likely to be replaced by a fully specified representation during the interpretation process. Normally, one uses an underspecified representation in the hope that, in the fullness of time, or as the need arises, it will be fully specified. In this sense, an underspecified representation is only a "temporary measure." Unlike a fully specified representation, it is not really a description of the world (which has a truth value), but rather a description of readings. However, as examples like (1) demonstrate, we may choose some antecedent, only to find later on that it is inappropriate. Even if there is only one candidate antecedent, it is possible that it will later be ruled out, leaving us with an accommodated or deictic reading. Hence, the representation of anaphora cannot be thought of as a temporary measure, to be discarded once the ambiguity is resolved. The underspecified representation cannot therefore be *ad hoc*, and must be fully motivated.

I suggest that we don't need to look far for a representation and its associated inference method. A standard, linguistically motivated representation, without

---

[1] Though see [5], who uses a nonstandard representation of anaphora, but applies Default Logic to generate its perceived readings.

special machinery for underspecification, will do: Discourse Representation Theory [13].[2]

Using this theory, the discourse in (8) will be represented by the following (simplified) DRS:

(9)

$x\ y\ z\ u\ v$
John(x)
Bill(y)
Mary(z)
shake-hands(x,y)
shake-hands(x,z)
male(u)
female(v)
hang-out(u,v)

Note that this DRS does not resolve the anaphora. In this representation, u and v are subject to existential closure, and all we know is that *some* antecedents exist. So, in effect, the DRS (9) is an underspecified representation, containing all the possible ways of resolving the anaphora.

The relation between anaphoric trigger and antecedent is represented in DRT as an equality relation. Thus, any specific resolution of the anaphora results in the addition of equalities identifying the referents of the pronouns. For example, if we identify *he* with John and *her* with Mary, we get the following DRS:

(10)

$x\ y\ z\ u\ v$
John(x)
Bill(y)
Mary(z)
shake-hands(x,y)
shake-hands(x,z)
male(u)
female(v)
hang-out(u,v)
u=x
v=z

While equalities such as the ones above are often treated as a mere notational convenience, it is clear from the formal definitions of [13] that they are *real* equalities, in the strictest logical sense. This means that we can apply the full power of the equality axioms, and get various desirable results for free. For example, if the antecedent has a certain property, then it immediately follows that the trigger has this property too.

---

[2] Of course, it may be the case that some sort of special underspecified representation is needed for other reasons, e.g., to represent scope ambiguities. All I claim is that such special representations are not necessitated by the need to represent anaphora.

In this paper I propose a simple formalization of DOAP using Default Logic [8]. The idea is that a trigger and a potential antecedent are equated by default, unless this is prevented by some rule. This default rule is assigned low priority, so that other factors affecting anaphora resolution can rule out inadmissible antecedents, or suggest an antecedent before DOAP applies. The result it that this principle would apply only if there is no strong preference for any antecedent; but when it does apply, the behavior of the resulting system complies with the desiderata described above.

4   Default Logic

The relation between trigger and antecedent is equality, so the problem of anaphora resolution becomes the problem of inferring the necessary equalities from the representation. As discussed above, this inference must be defeasible, so some form of nonmonotonic reasoning is necessary to formalize it.

One could, following [10], use Optimality Theory to state DOAP, but this would be problematic. While Optimality Theory is suitable for expressing defeasible, prioritized constraints, it does not employ a formal language; constraints in Optimality Theory are typically expressed in natural language, and may consequently be underspecified or vague—indeed, [10] use nothing more precise than the natural language definition of DOAP in (2). Since the goal of the current paper is a formal system, which could be implemented, and about which statements could actually be proved, this is not good enough. I will, instead, use a formal system with well defined syntax and semantics—Default Logic [8].[3]

Default Logic is one of the most widely used nonmonotonic formalisms. A substantial body of theoretical work has been devoted to it, and a number of theorem provers have been implemented.

A *default theory* is a pair (D, A), where D is a set of defaults and A is a set of first-order sentences. Defaults are expressions of the form

$$(11) \quad \frac{\alpha(x) : \beta_1(x), \ldots, \beta_m(x)}{\gamma(x)},$$

where $\alpha(x), \beta_1(x), \ldots, \beta_m(x),$ and $\gamma(x)$ are formulas of first-order logic whose free variables are among $x = x_1, \ldots, x_n$. Note that the presence of $\alpha(x)$ is optional.

The intuitive meaning of a default is as follows. For every n-tuple of objects $t = t_1, \ldots, t_n$, if $\alpha(t)$ is believed, and the $\beta_i(t)$s are consistent with one's beliefs, then one is permitted to deduce $\gamma(t)$.

For example, the following rule says that if something is a bird, and you don't know anything to the contrary, you may believe that it flies:

$$(12) \quad \frac{\mathbf{bird}(x) : \mathbf{fly}(x)}{\mathbf{fly}(x)}$$

---

[3] See [14] on implementing Optimality Theory in Default Logic.

Crucial to the interpretation of Default Logic is the notion of an *extension*. Roughly speaking, an extension of a default theory is a set of statements containing all the logical entailments of the theory, plus as many of the default inferences as can be consistently believed. A default theory may have more than one extension, as in the well known *Nixon diamond*. Suppose we have the following two defaults:

1. $\dfrac{\mathbf{Quaker}(x) : \mathbf{pacifist}(x)}{\mathbf{pacifist}(x)}$

2. $\dfrac{\mathbf{Republican}(x) : \neg\mathbf{pacifist}(x)}{\neg\mathbf{pacifist}(x)}$

The first rule says that Quakers are pacifist by default, and the second rule says that, by default, Republicans are not pacifist. If Nixon is both a Quaker and a Republican, in one extension he will be a pacifist, and in another he won't be. So, is Nixon a pacifist or isn't he?

When faced with multiple extensions, there are two general strategies we can use to decide which conclusions to accept: skeptical or credulous reasoning. Skeptical reasoning means accepting only what is true in all extensions. So, we will believe neither that Nixon is a pacifist, nor that he is not a pacifist. Credulous reasoning means picking one extension, based on whatever principles one deems appropriate, and accepting its conclusions. This means we will pick one extension, perhaps using our knowledge of Nixon's statements and actions, and based on this extension, conclude whether he is a pacifist or not.

A useful feature of some formalizations of Default Logic (e.g [15]) is the possibility of assigning priorities to defaults. Intuitively, this means that if default d_1 outranks default d_2, then it applies first, in the sense that there is no extension of the default theory that contains the conclusion of d_2 but not the conclusion of d_1, if both are applicable. While ranking is a very useful device, and we will use it too, it is important to emphasize that it doesn't add to the formal power of the system: for every ranked default theory, an equivalent unranked default theory can be constructed [16].

The semantics of Default Logic can be provided by *Herbrand models* [17, 18]. Suppose we have a first order language $\mathcal{L}$, and we augment it with a set of new constants, b, calling the resulting language $\mathcal{L}_b$. The set of all closed terms of the language $\mathcal{L}_b$ is called the *Herbrand universe* of $\mathcal{L}_b$ and is denoted $\boldsymbol{T}_{\mathcal{L}_b}$. A *Herbrand b-model* is a set of closed atomic formulas of $\mathcal{L}_b$.

5   Equality by Default

Resolving anaphora means generating an equality between two discourse referents. I suggest generating such an equality by default: we assume that two elements are equal if they cannot be proved to be different. The idea underlying this notion has been proposed, though not formalized, in [19]. Charniak's approach is further explored in [20], and formalized more fully in [21, 22], in which its potential for anaphora resolution is noted.

Equality by Default can be implemented in Default Logic very simply, with the following default:

(13)　　$\dfrac{:\ x = y}{x = y}$.

This rule means that whenever it is consistent to assume that two elements are equal, we conclude that they are. It would, of course, be *in*consistent to assume $x = y$ if we know that $x \neq y$. By the axioms of equality, then, (13) is equivalent to saying that we assume $x = y$ unless there is some property ϕ s.t. we know $\phi(x)$ but we also know $\neg\phi(y)$.

It might be objected that this is rather too liberal an assumption of equality, and that we allow two many elements to be equal by default. This, however, is not the case. Equality by Default does not apply in isolation; any reasonable system drawing inferences from natural language will require many more defaults, some of which deal specifically with anaphora, while others don't. If we assign low priority to Equality by Default, so that, if other defaults can apply, they will, inappropriate equalities will be ruled out, and rather few equalities will remain.

For example:

(14)　a.　John saw Bill. *He* greeted *him*.
　　　b.　John hates *him*.
　　　c.　John doesn't have a car. *It* is red.
　　　d.　A man came into the bar. *She* was upset.

The most likely interpretation of (14.a) is that the first pronoun refers to John, and the second one to Bill, hence they are not equal. This interpretation is brought about by a default rule that prefers antecedents that share the grammatical position of the pronoun (parallelism). In general, Equality by Default is a principle of *last resort*: it will not be invoked if other rules suggest some antecedent. Since in this case the parallelism rule applies, Equality by Default will not apply, and we are in no danger of concluding erroneously that the referents of the two pronouns are equal.

Sentence (14.b) does not have an interpretation where *him* is equated with John, for syntactic reasons. In (14.c), the pronoun *it* should not be equated with the discourse referent representing the indefinite *a car*, because, according to the rules of DRT, the indefinite is not accessible to the pronoun. The discourse in (14.d) is an example where the pronoun cannot be associated with the antecedent because of a gender mismatch. If all such constraints are formalized— as indeed they must be for any anaphora resolution system—and given a higher priority than Equality be Default, inadmissible antecedents will be ruled out.

We could restrict the definition of Equality by Default to apply only to anaphoric triggers and potential antecedents. However, this is not really necessary. Spurious equalities between arbitrary discourse referents will not be generated, because of independently motivated principles. Consider the following examples:

(15)　a.　John talked to Bill.

    b.   An officer talked to a gentleman.

    c.   John is meeting a woman tonight. His mother told me so.

    d.   John went to the clinic. The doctor had a busy day.

Sentence (15.a) involves two different names. Usually, it is assumed that two different names denote two different individuals; this is known as the Unique Names Assumption [23]. It might appear that our system cannot have the Unique Names Assumption, because different terms are assumed to be equal, rather than different, by default. However, this is not the case because, in DRT, names get their reference by *anchoring* them to individuals in the model, rather than by equality [13]. If the names *John* and *Bill* are anchored to different individuals, with different properties, then they must be different and cannot be equal by default.

Sentence (15.b) involves two indefinites. Standardly, indefinites are assumed to be novel [24]. This means that an indefinite must be different from any previously introduced discourse referent; hence, the referent of *a gentleman* must be different from the referent of *an officer*.

In sentence (15.c), *his mother* does not refer back to *a woman*. The reason is due to conversational implicature [25]: a speaker who knows that John is meeting his mother should say so, hence we conclude that the woman is someone else.

Sentence (15.d) is an example of bridging: *the doctor* is identified with the doctor associated with the clinic. Could it be equal to John by default? The answer is, in fact, yes, and the sentence does have this reading. But (15.d) also has another, perhaps more plausible reading, where John is a patient rather than a doctor. This reading is obtained because the notion of a clinic also introduces the notion of patients, together with the restriction that the patients are different from the doctor. According to one default conclusion, John is equated with the doctor, but according to another, he is equated with one of the patients, and is different from the doctor. Clearly, the two default conclusions are incompatible, hence we will have two extensions, one for each reading. We can then apply credulous reasoning to choose one of the readings, or skeptical reasoning, in which case the ambiguity remains unresolved.[4]

Thus, although the assumption of Equality by Default appears very permissive, in fact it allows rather few elements to be equal by default. These are intended to be anaphoric triggers and their potential antecedents, when no antecedent is suggested by an anaphora resolution factor.

Like other default theories, Herbrand models can provide a semantics for Equality by Default [22]. A clarification, however, is in order. Since the Herbrand universe of a language $\mathcal{L}_b$ is the set of all closed terms of $\mathcal{L}_b$, then, by definition, in a Herbrand model no two terms are identical. But in our default theory, two terms may be equal by default. Is this a contradiction? The answer is no. Equality is *any* relation that satisfies the equality axioms, and is not necessarily

---

[4] See section 6 for more on how the proposed system deals with ambiguity.

identity.[5] Hence, there is no problem about two terms being *equal*, even though they are not *identical*.

6 Unresolved Anaphora

As mentioned above, there are two cases where anaphora may remain unresolved: when there is no appropriate antecedent, or when there is more than one. In the first case, the trigger needs to be interpreted as referring to an entity not provided by the linguistic content (i.e. an accommodated or deictic interpretation). In the second case, the anaphora is truly ambiguous, and this ambiguity needs to be either resolved arbitrarily, or left unresolved, drawing as many inferences as possible.

6.1 No Potential Antecedent

It turns out that using Herbrand models has a consequence that is particularly important for our purposes. Note that the new elements introduced in b, by being new, are equal by default to any term. In particular, they are equal by default to any anaphoric trigger; this is how accommodated and deictic readings are possible.

This theory allows accommodated and deictic readings, but only as a last resort, when no other readings are possible. More precisely, it has been shown [22] that if E is an extension for Equality by Default, and w is a Herbrand b-model of E, then w is *minimal*. That is to say, there is no Herbrand b-model w' of E such that

(16) $\{\langle t_1, t_2 \rangle : w \models t_1 = t_2\} \subset \{\langle t_1, t_2 \rangle : w' \models t_1 = t_2\}$.

Now, consider a model w of extension E where trigger u is accommodated or interpreted deictically. This means that, in w, for every x_i, a potential antecedent of u, $u \neq x_i$; and for some new element $n \in b$, $u = n$. Since w is minimal, there is no Herbrand b-model w' of E that contains all the equalities in w and adds to them. Therefore, it is not only in w, but in all models of E, that u is different from all its potential antecedents.

What this means is that there is at least one extension, i.e. at least one plausible way of reasoning from the known facts, that is inconsistent with an anaphoric reading of u. Hence, accommodated or deictic readings are only available when the anaphoric reading is implausible (or impossible).

6.2 Multiple Potential Antecedents

Suppose we have two acceptable antecedents for some trigger. For example, in (7), repeated below, the pronoun may be equated with *John* or with *Bill*.

---

[5] Of course, we can a have a non-Herbrand model where equality *is* identity—such models are called *normal*, see [26, p. 100] for details.

(17)   John and Bill met at the ice cream parlor. *He* was upset.

If we make the standard assumption, as described above, that different names are anchored to different individuals, we know that *John* is different from *Bill*, so it is impossible to believe that the pronoun is equal to both. We will therefore have two extensions: in one of them, the pronoun is equated with *John*, and in the other—with *Bill*.

How do we deal with these extensions? We may decide to force a decision for one or the other; for example, we can decide that the most recent antecedent (Bill) is appropriate, or that the first one mentioned (John) is more prominent, hence preferred. So, in effect, we would apply credulous reasoning and pick one extension.

Note that, by the axioms of equality, once such a choice is made, any property of the antecedent becomes also a property of the trigger. For example, if we choose the extension where *he* is equated with Bill, and if Bill is bald, it will immediately follow that *he* is bald.

There are cases, however, where the anaphora is genuinely ambiguous, and we may have no reason to prefer one reading over the other. But even if we decide not to resolve the anaphora, there are still inferences we can make. Recall that, given (17), we want to conclude that whoever the pronoun refers to was at the ice cream parlor.

In this case, it makes sense to apply skeptical reasoning, and accept only what is true in all extensions. This will generate the desired inference, since in both extensions, the pronoun has the properties that its antecedent has.

Of course, there are extensions where *he* is equated with one or more of the new terms introduced in *b*. However, this makes no difference to the inference pattern described above, for the following reason. Even if *he* is equated with one or more new elements, it must also be equated with John or Bill. This is because so long as it is possible to find at least one antecedent for the pronoun, a model for the deictic reading, i.e. a model where the pronoun is equated with a new element but with no other element, will not be minimal, hence it will not be the model of any extension. In every extension, then, the pronoun will be equated with some old discourse referent x (which may be equated with any number of new elements in *b*). Since x is either John or Bill, and both are at the ice cream parlor, x is at the ice cream parlor. Since this is the case in every extension, skeptical reasoning will still conclude that the *he* was at the ice cream parlor.

Now let us consider cases where one possible antecedent has a property than the other one lacks, or is not known to have:

(18) a.   John walked along the sidewalk and saw that Bill was inside the ice cream parlor. *He* was upset.
      b.   John saw that Bill was eating ice cream. *He* was upset.

In (18.a), Bill is inside the ice cream parlor, but John is outside. Thus, in one extension, *he* will have the property of being inside the ice cream parlor, and in the other—its negation. If we apply skeptical reasoning, we will be able to conclude nothing—this appears intuitively correct.

In (18.b), we know that Bill was eating ice cream, but we do not know whether John was. Intuitively, we cannot conclude that *he* was eating ice cream, although this is consistent with the pronoun being equated with either John or Bill. Skeptical reasoning predicts this result: while in one extension the property of eating ice cream is predicated of the antecedent of the pronoun, in the other extension, neither this property nor its negation will be so predicated. Therefore, it is not true in all extensions that *he* was eating ice cream.

7   Conclusions and Further Applications

I have proposed a formalization of DOAP, a rule that tells us to exhaust all anaphoric possibilities before accommodating or interpreting a pronoun deictically. This rule is formalized using a standard linguistic representation (DRT) and a standard default reasoning system (Default Logic); no special mechanisms for representation or inference are required. Yet this conceptually simple theory appears to produce exactly the sort of inferences regarding anaphora that are intuitively desirable. It is ensured that if it is possible to find an antecedent, we do so; if more than one is a good candidate, we can use Default Logic techniques for dealing with multiple extensions; and if there is none, we accommodate or interpret the pronoun deictically. Thus, the resolutions such a system would make, and the inferences it would draw, can be proved explicitly, rather than be left implicit in the workings of the algorithm.

In this paper I have concentrated on anaphora. Yet I believe the theory can be extended to additional, related phenomena. The phenomena of presupposition and bridging immediately come to mind. Intuitively, these phenomena share with anaphora the notion of some trigger that is looking for an antecedent. In all three cases, a DOAP-like principle applies: it is preferred to choose an antecedent that is, in some sense, already given, than introduce a new one.

The theory presented here can be applied to these other phenomena as well, provided that the association of trigger with antecedent be represented as an equality relation, so that the rule of Equality by Default may be applied. Fortunately, this requirement can, indeed, be easily satisfied in standard theories of bridging and presupposition, in the following manner.

Regarding presupposition, we have already considered example (5) above, where the presuppositions of a definite description is accommodated if no antecedent can be found. This phenomenon is not restricted to definite descriptions, however, and appears to be a general fact about presupposition [12]: binding is preferred to accommodation.

For example, consider the following sentence:

(19)   If Jack comes in, then Mary will realize that a dangerous criminal came in.

Despite the factive verb in the consequent of the conditional, (19) does not presuppose that a dangerous criminal came in. Rather, the presupposition is bound by the antecedent of the conditional; in order for this to be possible, the

discourse referent corresponding to *a dangerous criminal* must be equated with Jack.

In general, resolving presuppositions involves adding, for each discourse referent x from the universe of the presuppositional DRS, a condition of the form $x = y$, where y is some discourse referent of the antecedent DRS [12].

The same holds for cases of bridging, even when the existence of the entity in question is not entailed by the antecedent, but is merely associated with it. The following example is from [27]:

(20)   John entered the room. He saw the chandelier sparkling brightly.

We tend to associate the chandelier with the room John entered, rather than with some other room, not mentioned in the discourse (which John can look into, say through a window).

What is the relation between the chandelier and the room mentioned in the discourse? Since normally rooms have some sort of light in them, we can assume that mentioning the room introduces a discourse referent representing this light source. Then, the relation between the chandelier and the light is that of equality, and Equality by Default can apply to produce the desired result.

It appears, then, that the proposed formalization of DOAP may be fruitfully applied to other phenomena besides pure anaphora. The precise details, however, must await another occasion.

References

1. Asher, N.: Linguistic understanding and non-monotonic reasoning. In: Proceedings of the 1st International Workshop on Nonmonotonic Reasoning, New Paltz (1984)
2. Byron, D., Gegg-Harrison, W.: Evaluating optimality theory for pronoun resolution algorithm specification. In: Proceedings of the Discourse Anaphora and Reference Resolution Conference (DAARC2004). (2004) 27–32
3. Lascarides, A., Asher, N.: Temporal interpretation, discourse relations and common sense entailments. Linguistics and Philosophy **16** (1993) 437–493
4. Mitkov, R.: An uncertainty reasoning approach for anaphora resolution. In: Proceedings of the Natural Language Processing Pacific Rim Symposium (NLPRS'95), Seoul, Korea (1995) 149–154
5. Poesio, M.: Semantic ambiguity and perceived ambiguity. In van Deemter, K., Peters, S., eds.: Semantic Ambiguity and Underspecification. CSLI, Stanford (1996) 159–201
6. Beaver, D.: The optimization of discourse anaphora. Linguistics and Philosophy **27**(1) (2004) 3–56
7. Prince, A., Smolensky, P.: Optimality theory: Constraint interaction in generative grammar. Technical report, Rutgers University, New Brunswick, NJ and University of Colorado at Boulder (1993)
8. Reiter, R.: A logic for default reasoning. Artificial Intelligence **13** (1980) 81–132
9. Williams, E.: Blocking and anaphora. Linguistic Inquiry **28** (1997) 577–628
10. Hendriks, P., de Hoop, H.: Optimality theoretic semantics. Linguistics and Philosophy **24** (2001) 1–32

11. Bry, F., Yahya, A.: Positive unit hyperresolution tableaux and their application to minimal model generation. Journal of Automated Reasoning **25** (2000) 35–82

12. van der Sandt, R.: Presupposition projection as anaphora resolution. Journal of Semantics **9** (1992) 333–377

13. Kamp, H., Reyle, U.: From Discourse to Logic. Kluwer Academic Publishers, Dordrecht (1993)

14. Besnard, P., Mercer, R., Schaub, T.: Optimality theory through default logic. In Günther, A., Kruse, R., Neumann, B., eds.: KI'03: Advances in Artificial Intelligence, Proceedings of the Twenty-sixth Annual German Conference on Artificial Intelligence. Volume 2821 of Lecture Notes in Artificial Intelligence., Springer-Verlag (2003) 93–104

15. Brewka, G.: Adding priorities and specificity to Default Logic. In Pereira, L., Pearce, D., eds.: Proceedings of the 4th European Workshop on Logics in Articial Intelligence (JELIA-94). (1994) 247–260

16. Delgrande, J.P., Schaub, T.: Expressing preferences in Default Logic. Artificial Intelligence **123** (2000) 41–87

17. Lifschitz, V.: On open defaults. In Lloyd, J., ed.: Computational Logic: Symposium Proceedings, Berlin, Springer–Verlag (1990) 80–95

18. Kaminski, M.: A comparative study of open default theories. Artificial Intelligence **77** (1995) 285–319

19. Charniak, E.: Motivation analysis, abductive unification and nonmonotonic equality. Artificial Intelligence **34** (1988) 275–295

20. Cohen, A., Makowsky, J.A.: Two approaches to nonmonotonic equality. Technical Report CIS-9317, Technion—Israel Institute of Technology (1993)

21. Cohen, A., Kaminski, M., Makowsky, J.A.: Indistinguishability by default. In Artemov, S., Barringer, H., d'Avila Garcez, A.S., Lamb, L.C., Woods, J., eds.: We Will Show Them: Essays in Honour of Dov Gabbay. College Publications (2005) 415–428

22. Cohen, A., Kaminski, M., Makowsky, J.A.: Notions of sameness by default and their application to anaphora, vagueness, and uncertain reasoning. ms., Ben-Gurion University and The Technion (2006)

23. Reiter, R.: Equality and domain closure in first order databases. Journal of the ACM **27** (1980) 235–249

24. Heim, I.: The Semantics of Definite and Indefinite NPs. PhD thesis, University of Massachusetts at Amherst (1982)

25. Grice, H.P.: Logic and conversation. In Cole, P., Morgan, J.L., eds.: Syntax and Semantics 3: Speech Acts. Academic Press (1975)

26. Mendelson, E.: Introduction to mathematical logic. Chapman and Hall, London (1997)

27. Clark, H.: Bridging. In Johnson-Laird, P., Wason, P., eds.: Thinking. Readings in Cognitive Science. Cambridge University Press, Cambridge (1977) 411–420