Explanatory AI, Indexical Reference, and Perception

Lawrence D. Roberts
Philosophy And Computers & Cognitive Science (PACCS)
State University of New York
Binghamton, New York 13902-6000
larry@turing.pacss.binghamton.edu

Abstract

Researchers in AI often say that certain types of reference are based on perception. Their models, however, do not reflect perceptual functioning, but instead represent denotation, an intellectually modeled relation, by using exact feature matching in a serial device as the basic mechanism for reference. I point out four problems in this use of denotation: substitution of an intellectual model for a perceptual one; unclarity about the nature of referential identification; relative neglect of the role of contrast in reference; and inexact matches. I then suggest an alternative theoretical account for perceptually based indexical reference, the figure-ground model, and I explain how this model handles the four problems.

I. Introduction

Reference concerns the relation of symbols to things outside the system of symbols. Indexical noun phrases (e.g., 'this,' 'I,' 'you,' 'that dog,') are important vehicles of reference in natural language. In their most basic use in communication, they serve to connect discourse with things present in the physical surroundings. To model such reference, AI workers have used *exact feature matching in a serial device*. They combine this mechanism with various devices for restricting the search space containing possible referents. In regard to devices for restricting search spaces, much progress has been made since the early 1970's when Winograd developed his work on reference in *Shrdlu*. For example, Grosz (1977, 1981), Sidner (1979), Reichman (1978, 1985), and Bruce (1981)) have developed theories aimed at restricting lists of possible referents. The motivation for these theories was general theories of the structure of discourse, rather than theories of reference. Despite this progress in restricting the search space for reference, the basic mechanism for picking the referent out of such search spaces has remained unchanged since Winograd's work.

In this paper, I argue that four issues in the AI literature on reference indicate serious problems in the supposition that exact feature matching in a serial device provides an explanatory mechanism for reference. I discuss the AI literature as though it aimed at *explanatory* rather than *purely specificatory* models for reference. An explanatory device aims at replicating certain results (e.g., the understanding of reference in natural language) by using means that are similar or parallel to the natural ones. In contrast, a purely specificatory device aims at replicating the results, but not at using similar means. I (1993) have discussed the distinction between explanatory and specificatory theories elsewhere. *My assumption of explanatory goals for AI studies of reference is not always correct*, but I make it for two reasons. First, it would take up too much space in a short paper to sort out which type of AI is the goal for each piece of work. More importantly, I am interested in explanatory AI for reference, and I want to show that current AI theories of reference have some deficiencies if construed as explanatory theories.

The four problems in the AI literature that I discuss concern the distinction between perceptual and intellectual bases for reference, the nature of referential identification, the role of contrast in referential identification, and inexact matches. To deal with these issues, I suggest an alternative approach to perceptually based indexical reference, the figure-ground model.

II. The Four Problems

A) Perceptual vs. Intellectual Bases for Reference

The basis for connecting indexical phrases to their referents in the physical surroundings is *perception*: I use 'that dog' or 'you' to refer in such a way that your perceptual attention is directed to the referent. My claim that perception is at the foundation of a basic variety of indexical reference is generally accepted by many AI workers (e.g., Cohen (1984), Goodman (1986), Appelt and Kronfeld (1987), Grosz (1981), Lesperance and Levesque (1993), Dale (1988), and Claasen (1992)). Despite their agreement about the *datum* that perception is the basis for certain types of reference, none of these authors provides any AI device to model perception. Instead, the basic mechanism they use to model reference is *exact feature matching in a serial device*. They use this mechanism in conjunction with various devices for restricting the search space containing possible referents.

Exact feature matching in a serial device models *denotation*, rather than perception. AI scholars, in fact, often seem to take denotation as equivalent to reference. Philosophers, since Strawson's (1950) "On Referring," have usually agreed that reference is an action of a speaker, involving his use of words and actions in a context to determine a referent. In contrast, denotation is simply a relation between a description and an object, by which the description is true of the object and of nothing else. Thus denotation leaves out of consideration the speaker, his actions, and the context. Nevertheless, many philosophers have used denotation to explicate the basic functioning of words in reference. In this way, the basic mechanism of reference is taken to be a type of predication.

The tradition of using denotation to explicate reference goes back to Frege (1879, 1950) and Russell. In his "On Denoting," Russell (1905) analyzed "The present king of France is bald" as "There is one and only one present king of France, and he is bald." This analysis takes denotation to be the mechanism by which "the present king of France" functions. Interestingly enough, Russell held that indexical reference is determined in a perceptual way, rather than by denotation (cf. my (1984)). But his view on indexicals has been ignored in the literature (though Castaneda (1977) and Smith (1982) offer similar views) while his denotational account is commonly accepted, not only for definite descriptions, but also for indexicals (for a prominent instance of the latter, cf. Quine (1960, p. 163)). Frege and Russell

influenced later analyses of reference not only by their papers on the topic, but also by their development of the predicate calculus, a formal system in which descriptions have no function but that of being predicates. The common use of this system as a tool for theorizing about natural language has promoted the view that descriptions in referring expressions function in virtue of being true of referents.

Using denotation to explicate the basic mechanisms of reference treats these mechanisms as functioning at an intellectual level, in virtue of true description, rather than at a perceptual level. Since one can perceive things without describing them, and can describe things without perceiving them, it is clear that description and perception are different. Nevertheless, most (but not all) AI scholars omit commenting on the fact that their models for perceptually based reference treat something quite different from such reference, namely, true description of the referent.

B) The Nature of Referential Identification

Some AI scholars explicitly note that their denotational models for reference do not treat the perceptual identification involved in certain kinds of reference. Philip Cohen (1984a, 1984b), as well as Appelt and Kronfeld (1987), say that perceptual identification is an additional step in referring, after the one represented by a true description of the referent. They also say that there is no adequate theory of just how people identify referents. On their view, not only are our AI models deficient in modeling referential identification, but we also do not know what referential identification is, i.e., we don't have any underlying philosophical account of such identification. This is a serious problem for AI theories of reference, since these theories take referential identification to be the heart of reference.

C)The Role of Contrast in Referential Identification

Several AI scholars say that contrast is basic to the functioning of referential identification (e.g., Dale (1988), Grosz (1981), Goodman (1986), Appelt and Kronfeld (1987), and Reichman (1985)). This position was held earlier by the psychologist, Olson (1970), who is cited by Grosz. Contrast of the referent to its surroundings is compatible with denotation of the referent, in that one could pick a referent out of some set by a description that was true of it and of nothing else in the set. Despite this compatibility, it is not clear that true description is *necessary* for distinguishing an intended referent from a set of possible referents because descriptions which are false of the referent often suffice in natural language for communicating reference. For example, one could use 'the man with the martini' to refer to a man drinking water from a martini glass, or one could use 'that desk' to refer to a table. Therefore, if contrast is to play the major role in reference, it is not clear that denotation, or exact feature matching in a serial device, is also needed.

D) Inexact Matches

If contrast is the most basic device for reference, then a description that is false of a referent, but sufficient for contrasting it to other possible referents, may suffice for reference. Such a description would "inexactly match" its referent: calling such a match "inexact," however, is based on the assumption that *true description* is the basic device for matching, rather than contrast. If true description is not the relevant device, then the use of 'the man with the martini' in the circumstances described above may count as an exact match according to some other basis for matching.

Both Barbara Grosz (1977) and Bradley Goodman (1986) discuss the fact that inexact matches of descriptions to referents often suffice for communicating reference in natural language. Grosz (1981) says that the question of inexact matches needs further study, and Goodman (1986) provides such a study. He has developed interesting devices for replacing or removing descriptions that do not match any possible referents. His device has preset rankings of which kinds of descriptions are most likely to be wrong, and it tries replacing or removing them, and then matching the resultant description set to the list of possible referents. Goodman's methods and algorithms are ingenious, and produce useful results. However, his approach is justified practically rather than theoretically, and has been called "*ad hoc*" by Charniak (1988: 288). In my opinion, several parts of Goodman's treatment are theoretically motivated, but not the assumption that exact matches like those modeled by denotation are required for reference in natural language. If contrast to other things in the surroundings, rather than denotation, is the basic mechanism of a certain type of reference in natural language, then many cases of "inexact matches" are exact enough. It is an empirical question whether people use mechanisms other than denotation to make reference.

Summary of section II. I have argued that there are several theoretical problems in using denotation, and exact feature matching in a serial device, as bases for explanatory models of perceptually based reference: denotation captures an intellectual rather than perceptual relation; current models leave it unclear just what constitutes referential identification; contrast rather than true description may be the basic mechanism of reference; "inexact matches" are commonly used in referring in natural language. In what follows, I describe a theoretical model for perceptually based indexical reference which deals with these problems.

III. The Figure-Ground Model

My theoretical model for the determination of indexical reference, which I (1986, 1993) have developed at length elsewhere, is called the "figure-ground model," and has three steps:

A. The *use* of an indexical determines a *context* containing the referent. (To simplify matters, I take the context to be the physical surroundings, but there are additional kinds.)

B. Gestures may direct the hearer's attention to one part of the physical surroundings as containing the referent.

C. Descriptions in the indexical phrase (or connected with it in certain other ways) function as a *figure* to make the referent stand out by *contrast to the background.*

Consider an example. Suppose a person, at a bar known by the conversers to be frequented by transvestites, says while nodding in a certain direction,

"That strikingly beautiful woman is not a woman."

In this context, the nodding is a gesture indicating that a certain part of the surroundings contains the referent, and the descriptive terms 'strikingly beautiful woman' function as a figure to make the referent stand out (for the hearer) qua particular from that background. This perceptual

functioning of the description requires the hearer to scan the indicated part of the physical surroundings for something that fits the perceptual content conveyed by 'strikingly beautiful woman.' This initial presentation of the figure-ground model is too brief, but I will clarify it further by contrasting it to denotation, and by explaining how it deals with the four problems about perceptually based reference.

Denotation functions in virtue of a predication model, in that the *denotatum* or referent is whatever the description in the referring expression is true of. The predicational mechanism identifies a thing *qua member of a kind*, in that the referent is whatever satisfies the description. For example, most uses of "the winner of the next New York state lottery" and of "whoever is in the next room" pick an individual out not qua individual, but rather qua having certain features, or falling under a certain description. A contrasting manner of identification is *identification qua individual* (or *qua particular*). For example, most uses of "this dog" or "that car" based on perception would identify a thing qua individual, because the hearer's attention is directed to pick out a certain individual from the surroundings. The mechanism of the figure-ground model naturally produces identification qua particular by the way it functions: the figure functions as a template that makes a particular thing stand out for the hearer from its background. Such contrastive functioning is quite different from directing attention to whatever it may be that a certain description is true of. The two different manners of identifying a referent have effects on the logic of negation and on rigidity of designation (cf. my (1993)). Note that I am not proposing that either the figure-ground model or the predication model is more fundamental than the other. Instead I view descriptive terms as having a lexical meaning which may be used either to provide a figure for drawing a contrast, or to provide a predicate that is to be true of something (or in various other ways--I do not attempt a complete account of the varieties of reference).

Next, I discuss the four problems. First, in contrast to denotation, the figure-ground model provides a perceptual model for perceptually based reference, as shown by its three main features: *hearing* the use of the indexical is the basis on which the hearer picks out the ground containing the referent; gestures direct the hearer's *perceptual* attention to a part of the surroundings; and the figure functions in a *perceptual* way in drawing a contrast of an individual to its surroundings. Even though this paper focuses on perceptual applications of the figure-ground model, the contrast of figure to ground is also present in higher level cognition, so that the figure-ground model is applicable also to indexical reference that is not perceptually based (cf. my 1993, ch. 3).

The perceptual nature of the figure-ground model enables it to give *contrast*, rather than true description, the major role in reference. Both gestures and descriptions function in the figure-ground model in virtue of contrasts to backgrounds. Descriptions need not be true of the referent as long as they suffice for distinguishing it from other possible referents in the indicated part of the surroundings. Because the description need not be true of its referent, as long as it provides the needed contrast, certain types of *"inexact matches"* of descriptions to referents are explained: a description may function by creating a perceptual contrast, rather than function at an intellectual level by being true of something. An AI model of such functioning could use a connectionist device, which allows a less rigid type of feature matching, that better fits perceptual matching. John Moulton and I are now working on such a model (Moulton and Roberts, 1994).

The figure-ground model also helps in regard to the problem of the nature of referential identification, both by offering a model for the mechanism of such identification, and by providing a basis for explicating identification qua particular. I submit that identification qua particular on the figure-ground model explicates what Cohen, Appelt, and Kronfeld call "referential identification" in the case of indexical reference. But this is not an account of referential identification in general. I believe that there are at least three types of referential identification based on uses of descriptions. One type is based on the figure-ground model, and a second type is based on the predication model and approximates denotation. The third type, which I have not discussed here, directs attention to a thing qua kind, and produces generic reference, as in "An elephant never forgets." I view reference made by means of descriptive noun phrases as a complex type of attention directing, which puts the descriptive content of the phrases to work in various ways, and not just by means of true predication.

I conclude that the figure-ground model provides an alternative to denotation as a model for perceptually based reference, and that this alternative overcomes several problems that exist for a purely denotational model.

References

Appelt, Douglas, and Kronfeld, Amichai (1987). A Computational Model of Referring. *IJCAI*, 640-647.

Bruce, B.C. (1981). Natural Communication between Person and Computer. *Strategies for Natural Language Processing*, ed. by W. Lehnert and M. Ringle, Hillsdale: Erlbaum.

Castaneda, H. (1977). Perception, Belief, and the Structure of Physical Objects and Consciousness. *Synthese*, 35, 285-351.

Charniak, Eugene (1988). Motivation Analysis, Abductive Unification, and Nonmonotonic Equality. *Artificial Intelligence*, 34, 275-295.

Claasen, Wim (1992). Generating Referring Expressions in a Multimodal Environment. *Aspects of Automated Natural Language Generation*, Heidelberg: Springer Verlag, 247-262.

Cohen, Philip (1984). The Pragmatics of Referring and the Modality of Communication. *Computational Linguistics*, 10, 97-146.

Dale, Robert (1988). The Generation of Subsequent Referring Expressions in Structured Discourses. In M. Zock and G. Sabah, *Advances in Natural Language Generation*. vol 2., (pp. 58-75). London: Pinter.

Frege, Gottlob, (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Halle: Nebert, 1879, reprinted in Angelelli, I., (ed.) *Begriffsschrift und andere Aufsaetze*, Hildesheim: Olm, 1964, vii-88.

Frege, Gottlob, (1950). On Sense and Reference. trans. by M. Black. *Philosophical Writings of Gottlob Frege* (56-78). Oxford: Blackwell.

Goodman, Bradley (1986). Reference Identification and Reference Identification Failures. *Computational Linguistics*, 12, 273-305.

Grosz, Barbara (1977). The Representation and Use of Focus in Dialogue Understanding. Technical Report 151, Artificial Intelligence Center, SRI International.

Grosz, Barbara (1981). Focusing in Descriptions in Natural Language Dialogues. *Elements of Discourse Understanding*, Cambridge: Cambridge Univ. Press, 84-105.

Lesperance, Yves, and Levesque, Hector (1993). Indexical Knowledge and Robot Action. manuscript, Burnaby and Toronto.

Moulton, John, and Roberts, Lawrence D. (1994). An AI Module for Reference Based on Perception. In Mc Kevitt, Paul (ed.) *Proceedings of the AAAI-94 Workshop on the Integration of Natural Language and Vision Processing*. Seattle.

Olson, D.R.(1970). Language and Thought: Aspects of a Cognitive Theory of Semantics. *Psychological Review*. 77, 257-273.

Quine, W. V. (1960). *Word and Object*. Cambridge: MIT Press.

Reichman, Rachel (1985). *Getting Computers to Talk Like You and Me*. Cambridge: MIT Press.

Roberts, Lawrence D. (1984). Russell on the Semantics and Pragmatics of Indexicals. *Philosophia*. 14, 111-127.

Roberts, Lawrence D. (1986). The Figure-Ground Model for the Explanation of the Determination of Indexical Reference. *Synthese*, 68, 441-486.

Roberts, Lawrence D. (1993). *How Reference Works*. Albany: State University of New York Press.

Russell, Bertrand (1905). On Denoting. *Mind*. 14, 479-93.

Searle, John (1980). Minds, Brains, and Programs. *Behavioral and Brain Science*, 3, 417-457.

Sidner, C. L. (1979). *Towards a Computational Theory of Definite Anaphora Comprehension in English Discourse*, technical report 537, Artificial Intelligence Laboratory, MIT.

Smith, D. W. (1982). What is the Meaning of 'This'? *Nous*, 16, 181-208.

Winograd, Terry (1972). *Understanding Natural Language*, New York: Academic Press.