Direct Reference in Thought and Speech

Kirk A. Ludwig
Department of Philosophy
University of Florida
Gainesville, FL 32611-8545

1. Introduction

I want to begin by distinguishing between what I will call a pure Fregean theory of reference and a theory of direct reference. A pure Fregean theory of reference holds that all reference to objects is determined by a sense or content. The kind of theory I have in mind is obviously inspired by Frege, but I will not be concerned with whether it is the theory that Frege himself held.[1] A theory of direct reference, as I will understand it, denies that all reference to objects is determined by sense or content. We will also distinguish between a theory of reference for thought, and for language. This gives us a fourfold classification of theories.

What is puzzling about direct reference theories is not that the semantics of an expression in a public language should assign as its semantic value just a referent, but how such facts could be understood to reflect an underlying feature of thought. There are two interconnected aspects to this

---

[1] There is considerable controversy over the proper interpretation of Frege, and it is not clear that all that roles that Frege wanted to be filled by his notion of sense could be filled by a single thing. I will be taking as a starting point one aspect of the Fregean notion of sense to draw a distinction between sense based theories of reference and direct reference theories. I will not insist on the identification of what is true and what is the referent for Frege of a sentence in indirect discourse with the sense of a sentence. See John Perry, "Frege on Demonstratives," (*Philosophical Review,* 86 (1977): pp. 474-97) for more on these elements in Frege's thought, and Gareth Evans, "Understanding Demonstratives," (*Collected Papers* (Oxford: Oxford University Press, 1985): pp. 291-321) for a rejoinder. See also for further discussion Tyler Burge, "Belief *De Re,*" (*Journal of Philosophy,* 75 (1978): pp. 338-62) and "Sinning Against Frege," (*Philosophical Review,* 88 (1979): pp. 398-432).

puzzle, one metaphysical, and one epistemological. The metaphysical puzzle is how a thought could be essentially object involving, that is, have as a constituent of it an object as a leg is a constituent of a table. The metaphysical puzzle, though, is underlain by the epistemological puzzle. If some thoughts are directly referring, then their contents cannot be characterized independently of reference to an object. But there is no object in the world around us to which we stand in a privileged epistemic relation of the sort we suppose that we stand in to our own thoughts. A direct reference theory then seems to open up an epistemological gap between ourselves and the contents of our thoughts, and this seems scarcely imaginable.

The problem is generated (at least in part, I think) by the desire to have the contents of our thoughts satisfy the following three conditions:

(I)   A person has first person knowledge of the contents of his thoughts.

(II)   A person's behavior is sensitive to the contents of his thoughts.

(III)   The contents of a person's thoughts determine their truth conditions.

First person knowledge is that knowledge we have of our own mental lives which is different in kind from the knowledge that anyone else could have. The difference lies not in *what* is known by in *how* it is known. To say that a person's behavior is sensitive to the contents of his thoughts is to require that where there are no differences in content there is no difference in behavior and that where there is a difference there is at least a potential difference in behavior. This can be thought of as one feature of having first person knowledge of one's own thoughts. To require that the contents of a person's thoughts determine their truth conditions is to require that the conditions under which any thought of a person is true are determined entirely by reference to his thought contents. The difficulty is that if a thought is directly referring, it looks as if there is nothing that will play the role that we want content to play in these three conditions. I do not have first person knowledge of objects in my environment; but if a thought is directly referring, the facts that determine its truth conditions must mention the object that is its referent. So if my thoughts involve directly objects in my environment, it seems as if I will not be

able to have first person knowledge of their truth conditions, or even have first person knowledge that I have a thought with truth conditions at all, for I do not have first person knowledge of the existence of any external objects.

One solution to this problem is to deny that a direct reference theory is correct for thought. I will argue, however, that we cannot avoid a direct reference theory for thought if we are to accommodate a referential capacity which we know we have. But I will offer a conservative theory of direct reference for thought which preserves as much as possible of the intuitions which incline us to hold (I)-(III).

The plan of the paper is as follows. In section 2 of the paper, I will characterize more precisely the two theories of reference that I am contrasting. In section 3, I will give an epistemological argument for the claim that a direct reference theory for thought is correct. In section 4, I will argue that the requirements of the epistemological argument can be met if it is possible for us to refer directly to the self-at-a-time, and that we are independently committed to our being able to refer directly to the self-at-a-time. (Throughout my discussion I will use the expression 'the self' to mean 'oneself'; so if we are committed to referring directly to the self, each of us is committed to being able to refer directly to him or herself.) In section 5, I will consider how far an explanation of direct reference is possible, what this thesis commits us to, and I will show how this avoids the worst of the puzzles that afflict direct reference theories of thought content. In section 6, I will consider the relation between direct reference in thought and in language, and draw some conclusions about externalist claims about thought content. The final section will summarize the results of the paper.

2. Pure Fregean theories of reference and direct theories of reference

We will identify the sense of an expression in a language L as its cognitive significance to a fully competent speaker of L in virtue of his competence in that language. We will restrict our attention to indicative sentences and to those semantic properties of a sentence which have a bearing on its truth value. (A broader conception of sense would include, e.g., our competence in the use of grammatical

forms and their connections with the force of an utterance.) This Fregean conception of sense, as I have characterized it, does not provide an analysis of linguistic meaning (even as bearing on truth alone) because it must appeal to the unanalyzed conception of a speaker's competence in the use of a language, which is to appeal to his knowledge of the meanings of the terms in that language. In effect, we have said: suppose that X is a competent speaker of a language, that is, knows how to speak that language. If X would use two expressions which have senses interchangeably in indicative sentences when we restrict his knowledge to his knowledge of the proper use of those expressions, and all that is at issue is the truth of the sentences, then, and only then, those two expressions have the same sense.[2] Thus, the Fregean notion of sense is linked to the notion of an ideally competent speaker's knowledge of how to use expressions in a language, and in this sense is an epistemic conception.

Note that I do not say that a difference in the cognitive significance of expressions is tied to the possibility of taking different epistemic attitudes toward sentences otherwise alike containing them. The reason for this will be given below.

As we have characterized 'cognitive significance', the syntactical features of an expression are irrelevant to its cognitive significance, for its syntactical features, except in cases in which an expression

---

[2] This makes the fine-grainedness of senses relative to the language. A purely extensional language, e.g., would on this criterion individuate senses more coarsely than a language which includes propositional attitude idioms and modal operators. For our language, this generates the following criterion for the synonymy, or sameness of sense, of two expressions:

> Two expressions are synonymous iff they are inter-substitutable *salva veritate* everywhere in a language which contains modal operators and propositional attitude idioms.

If two expressions, A and B, meet this criterion, then the following is true

> Necessarily, X believes that S iff X believes that R

where 'X' names a person, 'S' and 'R' stand in for sentences, 'S' contains A, and 'R' is obtained from 'S' by substituting B for A. The suggestion in this form is due to Benson Mates, "Synonymity," in L. Linksy, *Semantics and the Philosophy of Language* (Urbana: University of Illinois Press, 1952): pp. 111-38.

is simultaneously used and mentioned, do not contribute to the truth conditions of a sentence in which the expression occurs. One may of course learn something when told that

(3)   all and only bushes which are furze are gorse,

even though 'furze' and 'gorse' are synonymous in English and their extensions are determined by their senses. But in an appropriate context what one learns here is that these two expressions have the same sense in the language, that is, what one learns in this case is metalinguistic, even though the sentence that one learns it from is not metalinguistic in form. If one is already fully competent in the use of these expressions, then one learns no more from (3) than from

(4)   all and only bushes which are furze are furze.

Therefore, (3) and (4) do not differ in cognitive significance in the sense defined above.

As we have characterized it, the sense of an expression is something a speaker learns when he learns a language. The cognitive significance of an expression is its significance in virtue of the semantic type under which it falls. Senses, then, are types under which expression types and tokens fall. The same expression can be used on different occasions (or even the same occasion) with different senses; different expressions can have the same sense; an expression may have no sense at all.

We will call a referring term which denotes a unique individual a singular term, and a referring term which has a set of individuals (possibly empty) as its extension a predicate. A pure Fregean theory of reference for language (PFTL), as I will understand it, holds the following:

(PFTL)   If T is a referring term, then T has a sense S, and S determines the referent of T if it is a singular term, and the extension of T if it is a predicate.

To put this another way, for any referring term, there is a function from the sense of that term to its referent or extension. It is also a standard feature of a Fregean theory of sense that different senses can determine the same referent or extension.

A direct reference theory for language (DRTL) denies PFTL, and therefore holds that

(DRTL)   There is at least one referring term T such that either T does not have a sense and has a referent or extension, or T has a sense but its sense does not determine its referent or extension.

This can be understood to entail that if a term T has a sense, there is no function from its sense alone to its referent. Thus, we can represent the difference between PFTL and DRTL in terms of a difference in the reference functions for languages which contain no directly referring terms and languages which contain directly referring terms. The reference function for a language which satisfies a PFT will have the form, $F(T,S) = R$, where T is a token utterance of a term, S is a sense, and R is a referent or extension. For a language which contains directly referring terms, the reference function must be of the form, $F(T,S,X) = R$, where 'X' represents some feature of the token in addition to its sense, e.g., its being in a certain context C when uttered.

If a PFT for a language is true, then there is a function from the sense of any sentence in the language to its truth conditions, that is, its sense together with the way the world is determines whether or not the sentence is true. (In the case of analytically true or false sentences, this is trivially true.) For a language for which a DRT is correct, there will be some sentences the sense of which underdetermines their truth conditions, so that the sense together with the way the world is does not determine whether the sentence is true.[3]

---

[3] It will be appropriate here to connect my terminology with some others which have become prominent. What I mean by the sense of an expression corresponds to what David Kaplan calls its character. These are not identical, because Kaplan thinks of the character as a function from a context to a content, and I make no such identification. But they correspond because what one knows in virtue of knowing a language is what the character of an expression is. Kaplan's notion of a content is

Let us suppose that we are concerned with a pure Fregean language (PFL), that is, a language for which a PFT is true. In this case, we can introduce an additional criterion for two expressions differing in cognitive significance. In a pure Fregean language two sentences differ in cognitive significance for a speaker competent in their use just in case that speaker could believe either one to be true without believing that the other was true. Thus, in a PFL, we will say that two sentences, 'S' and 'R', have the same sense just in case the biconditional

(1)   S iff R

has no more cognitive significance for a fully competent speaker than does

(2)   S iff S.

That is, the speaker learns nothing more from (1) than he does from (2). Two expressions, A and B, in a pure Fregean language L have the same sense just in case for any sentence *s* of L in which A occurs, if *r* is the sentence obtained by replacing A with B, *r* and *s* have the same sense for a fully competent speaker.

This criterion is a standard one for introducing cognitive significance, but it tracks our original characterization of sense only in a PFL. For in a language in which there are directly referring terms which do not have any sense at all, but only a referent, or which have a sense insufficient to determine a referent or extension, a competent speaker will not know from knowing the senses of the expressions in a sentence containing such terms the conditions under which it is true, and, therefore, will not know from knowing the senses that two sentences which are the same in *sense* have the same truth value. Doing so would be possible only if the sense of a sentence determined its truth conditions, but this is not

---

equivalent to what I have been calling the truth conditions of an utterance, where we understand this to be the conditions which have to be met for the sentence to be true as opposed to the states of affairs in the world that make it true.

so for sentences containing directly referring terms. A fully competent speaker will know when the sense of a sentence fails to determine its truth conditions, and therefore will know that sentences which are the same in sense but contain different directly referring terms may differ in truth value, and so will not believe one simply because he believes the other. This is why I have not used this standard device to give a general characterization of sense. For otherwise we would be forced to deny the possibility of directly referring terms without any sense, or sentences whose senses were insufficient to determine the conditions under which they were true.

We turn now to thought. To extend the distinction between a PFT and DRT of language to thought we need to substitute for linguistic expressions a different bearer of the sense. I will call this a thought. In contrast to Frege, in my usage, a thought is a psychological state, not a proposition. I will speak of the content of a thought where I would speak of the sense of a sentence. (My use of 'content' here is not to be confused with Kaplan's quite different use. If I wish to speak of Kaplan's notion, I will add a subscript 'k': content$_k$.) We can restrict our attention to beliefs and other cognitive attitudes as we have restricted our attention above to indicative sentences. To keep the discussion as general as possible, we can think of a thought simply as a token state of a person. A difficulty arises in the case of thought, however, which does not arise in the case of language. That is that there is no syntax which underlies contentful thoughts in the way that a syntax underlies meaningful sentences. Whether a physical state type falls under a syntactical type is a matter of convention. We impose syntactical types on physical types. Since the possibility of this presupposes our having thoughts, we cannot think of our thoughts as being syntactically individuated, and as then falling under content types. The states that are our thoughts cannot be individuated independently of their contents in the way that the linguistic bearers of meaning can be.[4] In this respect the states that are thoughts are like the bodily movements that are actions: the bodily movements that are actions cannot be individuated independently of their being identified as the movements which appropriately produced by an intention.

---

[4] See John Searle, "Is the Brain a Digital Computer?" *Proceedings and Addresses of the American Philosophical Association,* 64 (1990): pp. 21-37.

There may nonetheless be a physical structure underlying our thoughts which has a language-like structure, a language of thought, as it is often called, with some confusion.[5] Since I cannot see any reason to think that this is necessary, to keep the discussion as general as possible, I will not assume it.

We will instead, following a long tradition, call the elements of a thought concepts. Concepts are usually thought to be the elements in thought that correspond to predicates. Understanding a predicate can be thought of as associating with it a concept. Concepts are elements of thoughts in the sense that the same concept can be involved in different thoughts. We can extend this terminology to elements in thought which correspond to expressions which have logical functions; we can, e.g., talk of the concept of disjunction as being involved in the thought that Gainesville is either north or south of Miami. A difficulty arises, however, in the case of singular terms, if we wish to keep open the question whether there can be direct reference in thought, or object involving thoughts. One conception of an object involving thought is of a thought which has as a constituent the object which it is about. On this conception there need be no element in thought that corresponds to the object separate from the object itself. Therefore, when framing the question about whether a DR theory is correct for thought, I will ask whether the content of the thought as a whole, which is determined by the concepts which the thought involves, determines the thought's referent, what particular object it is about, if it is about a particular object. More generally, we can talk about whether the content of the thought determines the truth conditions of the thought. There could be two ways in which it could fail to do so. First, by failing to determine the thought's referent; second, by failing to determine whether or not the thought's referent falls under the concepts which the thought involves, that is, by failing to determine the extensions of its concepts.

This suggests the following formulation of the PFT for thought (PFTT):

(PFTT)   If T is a token thought, then T's content determines its truth conditions; thus, in particular, if T has a referent, T's content determines its referent.

---

[5] See Jerry Fodor, *The Language of Thought* (New York: Thomas Y. Crowell, 1975).

In contrast, then, a direct reference theory for thought (DRTT) would hold that

(DRTT)   There is at least one type of thought T such that T's content fails to determine T's referent or truth conditions.

We can represent this difference as above in terms of a difference in the functions each theory would determine from thought contents to the truth conditions for the thoughts. In the case of PFTT we have, $F(T,C) = p$, and in the case of DRTT we have, $F(T,C,X) = p$, where C is the content of the token thought, p is a proposition that gives the truth conditions, and X is some additional feature of the token thought, perhaps its being embedded in a particular context.

It is important that we think of the content of a thought as a type under which it falls; thus the content of a thought is never itself specified in an object involving way. Hence, it will always be possible for someone else to have a thought with the same content as any thought I have.

We have not characterized content in terms of cognitive significance in the way we characterized the sense of a linguistic expression. Cognitive significance was explained in terms of what a speaker knows in being competent in a language. In the case of thought, there is nothing equivalent to learning the meaning of an expression. But there is an analog for the cognitive significance of an expression. That is our knowledge of the truth conditions for the thought. But now there is another way in which thought differs from language: for as we will see in the sequel, it may be possible to know the truth conditions for a thought even though the thought is object involving. So in the case of thought, the content of a thought should not be identified with the extent of one's knowledge of its truth conditions, although in the case of language we are able to identify the sense of a sentence with the extent of the fully competent speaker's knowledge of its truth conditions.

A pure Fregean theory of reference is false for natural languages, and quite apart from this certainly would not be correct for every possible language. In natural languages, indexicals, demonstratives, and proper names are all expressions which are directly referring in the above sense. For example, a token utterance of

I am in Madrid

does not have its truth conditions determined by its sense alone, in the sense explained above, because the referent of 'I' is not determined by its sense alone, that is, there is no function that takes us from the sense of a token use of 'I' to its referent. For it is clear that the sense of 'I' is the same when I use it as when you do, though we do not refer to the same person. 'I' is therefore a directly referring term. Other familiar examples are 'you', 'he', 'now', 'then', 'here', 'there', 'today', 'yesterday', 'this' and 'that'. Even if there were no such expressions in any actual language, however, it is clear that it would be possible to introduce them. For example, we could introduce an expression in a language to function as a device for denoting a given individual, without our being concerned with how any particular speaker who uses the expression picks out that individual. The semantic value of the expression would then simply be the individual denoted. Arguably, this is how proper names function.[6]

These remarks about ordinary languages, however, do not show that a PFT is correct for thought. For it might be argued that every thought does have its truth conditions determined by its content, but that when we express thoughts in a public language, because of the social function of language and our interests in communication, we often do not express the full content of a thought. So, e.g., when I use a proper name, 'Napoleon', to designate an individual in order to say of him that he was great general, I do not care whether your way of picking out Napoleon is the same as my way, as long as we are talking about the same individual; but it does not follow that I do not have a way of picking out Napoleon in virtue of the content of a thought. The same is plausibly said of the use of the demonstratives, "this" and "that" as well. For the purposes of communication, it is often irrelevant how an individual picks out an object, as long as the parties to a conversation know that they pick out the same object. It is a mistake to suppose, as is sometimes done, that the semantics of an expression that

---

[6] Proper names are directly referring expressions, but not for the same reason that indexicals and demonstratives are. Proper names have no senses at all. Indexical and demonstratives have a sense, but their sense alone does not determine the referents of token uses of them, but only in a context.

a speaker utters gives the complete content of any of his thoughts. We will return to this point in the penultimate section of the paper.

In the next section of the paper, I will argue that being able to refer directly to something in thought is a necessary condition for a referential ability which we possess. The argument will be given without assuming that we can read off the content of a thought from the semantics of sentences used to express it.

3. An argument for a direct reference theory for thought

The argument which shows that a direct reference theory for thought is correct exploits four facts:

(i)   it is logically possible for a universe to be spatio-temporally symmetrical;

(ii)   we do not know that the universe is not spatio-temporally symmetrical.

(iii)   we can refer to unique spatio-temporal particulars;

(iv)   we know we can refer to unique spatio-temporal particulars;

A universe is spatio-temporally symmetrical provided that it has two or more spatio-temporal regions which are qualitatively indistinguishable.   It seems clear that it is logically possible for a universe to be spatio-temporally symmetrical. A simple example of a spatio-temporally symmetrical universe is provided by Nietzsche's hypothesis that a universe which exists for an infinite length of time will repeat each qualitatively described historical period exactly an infinite number of times. A simpler example of a spatio-temporally symmetrical universe is one which alternately contracts to a space-time singularity, and then explodes and expands until the gravitational attraction of its parts begins another cycle of contraction, such that each period between big bangs is qualitatively identical to each preceding period. An infinite universe could also be conceived of as spatially symmetrical; perhaps, like David

Lewis's possible worlds,[7] each spatially distinct duplicate of a given spatially infinite region of the universe is causally isolated from every other one.

We do not know that the actual universe is not a spatio-temporally symmetrical universe. We do not know, for instance, whether or not our universe instantiates Nietzsche's hypothesis. Nor do we know whether or not the universe is cyclical in the sense that it expands and collapses endlessly, reproducing after each new big bang a universe qualitatively identical to the previous period, or whether it is spatially symmetrical.

Despite our not knowing this, however, it is clear that we are able to refer, and to know that we refer, to unique spatio-temporal particulars. I form a thought now, e.g., about the hand which I hold up in front of me. Philosophical skepticism aside, I do not think that anyone would deny that I do know that I am referring to a unique hand, that I am able to have a thought about this hand in front of my face. Even if this were doubted, it would not yet be doubted that each of us refers to himself.

Suppose that the universe is spatio-temporally symmetrical. Suppose also, now, for a *reductio ad absurdum,* that I refer to a unique spatio-temporal particular in virtue of the content of my thought about it. Let us call a spatio-temporal region of the universe which has a duplicate a d-region. The content of my thought is a type. Therefore, the content of my thought is also the content of a thought of a duplicate of me in another d-region. The PFTT holds that the content of my thought, that is, the content type that the token thought falls under, determines its referents and the extensions of the concepts which it involves. If the content of my thought determines a unique spatio-temporal particular at its referent, then the content of any duplicate of me in a d-region will determine the same unique spatio-temporal particular as *its* referent. But which particular am I thinking about? Clearly, I am thinking about a spatio-temporal particular in the d-region in which I exist (if any). So if the content of my thought determines a unique spatio-temporal referent, it is a referent in the d-region in which I exist. The same argument, however, can be given for each individual with a thought with the same content. Thus, each individual with a thought with the same content refers to a different unique spatio-temporal individual. But this is now in conflict with our conclusion that each refers to the same unique spatio-

---

[7] See David Lewis, *On the Plurality of Worlds* (Oxford: Basil Blackwell, 1986).

temporal particular. Thus, if the universe is spatio-temporally symmetrical, we must give up one of the assumptions we started with: if we hold that we refer to unique spatio-temporal particulars in the d-region in which we exist, we must give up the assumption that the contents of our thoughts determine which unique spatio-temporal particulars we refer to.

This is a conditional conclusion. It is that if the universe is spatio-temporally symmetrical, we cannot refer to unique spatio-temporal particulars around us by the contents of our thoughts alone. Therefore, if the universe is spatio-temporally symmetrical, and we refer to unique spatio-temporal particulars, then a PFT for thought cannot be correct. So far, granting that we do refer to unique spatio-temporal particulars, this does not yet show that a PFT is correct for thought. To move to this conclusion, however, we do not have to assume that the universe is spatio-temporally symmetrical. We can move to this conclusion by invoking (iv), the fact that we know that (iii) is true, and (ii), which says that we do not know the universe is not spatio-temporally symmetrical. If (ii) is true, (iv) could not be true if we could refer to unique spatio-temporal particulars in virtue of the contents of our thoughts alone. For then we would refer to unique spatio-temporal particulars only if the universe were not spatio-temporally symmetrical. If we assume that knowledge is closed under known entailment, then since we know this conditional, and we know that we refer to unique spatio-temporal particulars, but we do not know that the universe is spatio-temporally symmetrical, it follows that we do not refer to unique spatio-temporal particulars solely in virtue of the contents of our thoughts.

We can therefore conclude that there is direct reference in thought, that is, that some reference in thought to spatio-temporal particulars is direct reference in the sense defined in the previous section. This is shown by the fact that it is epistemically possible that this is a spatio-temporally symmetric universe, despite our knowing that each duplicate of me would have same contents as I do but refer to different unique spatio-temporal particulars. Thus, the referent of each of duplicate's thoughts cannot be a function solely of the content of those thoughts.

The argument I have given is an epistemological argument, in the sense that one of its crucial premises is an epistemological assumption: that we do know that we refer to unique spatio-temporal particulars. I think this will generally be granted, but I do not intend to defend it in this paper. For

someone who disputes this assumption in the light of the consequence, the paper can be read as an exploration of what must be true if we accept that we do know that we refer to unique spatio-temporal particulars.

To establish that a DR theory is correct for thought it is not strictly necessary to appeal to symmetries on the scale of the universe. It is sufficient if we know that we are referring to a unique spatio-temporal particular even though we do not have any unique identifying content. This is possible as long as we know we refer to a particular even when we know we have no way to qualitatively distinguish two possible referents.

I have, in this section, discussed *spatio-temporal* particulars because it seems that it is only in this case that we find the kinds of symmetries that give rise to the necessity for direct reference. The argument depends upon, on the one hand, the characterization of content as being a type, and so not object involving, and, on the other, the epistemic possibility that the universe is spatio-temporally symmetrical. Where it is impossible to produce such symmetries, it is *prima facie* possible to refer to a unique particular in virtue of the content of a thought. This will be possible only if there are objects that are not to be met with in space and time, that is, abstract objects. Thus, if we countenance abstract objects, such as numbers, it is plausible that it is the content of my thoughts that determine which of them I refer to. For example, it is plausible that the content of my thought that the positive square root of 4 is greater than 1 determines the referents of my thought. It is the logical possibility of spatio-temporal symmetry which gives rise, in the light of our assumptions, to the need for direct reference in thought.

4. Direct reference to the self-at-a-time

Let us say that unique-reference is reference to a unique spatio-temporal particular which meets the requirement that it succeeds even if it is reference in a spatio-temporally symmetrical universe. Unique-reference is not the same as direct reference, for I may uniquely-refer to one object in virtue of specifying a unique relation to an object to which I directly refer. In this section, I will argue for two

conclusions. (1) It is in principle sufficient to uniquely-refer to any spatio-temporal particular that one is able to directly refer to the self-at-a-time. (2) Certain things we accept commit us to our being able to directly refer to the self-at-a-time. Take together, these two conclusions show that we are able to uniquely-refer to spatio-temporal particulars in virtue of being able to directly refer to the self-at-a-time.[8]

(1) There are two sorts of symmetries which raise difficulties for a PFT for thought: spatial symmetries and temporal symmetries. The problem of determining a unique individual can be thought of as the problem of specifying which d-region the individual is in and providing a uniquely identifying content which picks out the individual in that region. This can be accomplished if we can fix a spatio-temporal point relative to which we can describe the spatio-temporal positions of objects in a d-region. A complete qualitative description of a d-region plus a coordinate point in that region is in principle sufficient to refer to any particular in that region. A uniquely identifying description is a specification of a spatio-temporal coordinate or region together with a unique description of an object at that point or in that region. (In practice causal relations will play an important role in specifying unique relations to objects around us.) Direct reference to the self-at-a-time secures reference to a unique spatio-temporal coordinate point: the spatial coordinates are give by the spatial location of the self, and the temporal coordinate is give by the temporal location of the self. Thus, direct reference to the self-at-a-time is in principle sufficient to uniquely-refer to any object in a d-region.

(2) My starting assumption here is that unless an individual can uniquely-refer to himself, he cannot refer to anything, and that any individual who can uniquely-refer to himself can know that he does so when he does. The reason for this is Kantian in spirit: it is that nothing could unify the various thoughts an individual has as that individual's thoughts except an internal connection to that self, and that there could be such an internal connection only if it were possible for that individual to be conscious of

---

[8] This claim is inspired by Russell's discussion of ego-centric particulars in *Inquires into Meaning and Truth* (London: George Allen & Unwin, 1950). I do not follow Russell, however, in holding that indexical expressions are all translatable into expressions containing only 'I' and 'now'. I give reasons to reject this in section 6. My thinking on this topic has also been influenced by John Searle. See his *Intentionality* (Cambridge: Cambridge University Press, 1984).

that thought as *his* thought. While as it stands this reason is fairly obscure, I will not in this paper give further argument for this starting point. This starting assumption does not yet beg the question at issue because it does not by itself entail that anyone can *directly* refer to himself.

Suppose that one could refer directly to some object at a time which was not identical to oneself, but could not refer to oneself directly. To refer to oneself, then, one would have to specify a relation from that object to one's self in which only one's self could stand to it. If every person is identical to some physical object, then it is in principle possible to refer to oneself if one can refer to any object in a d-region. What we require is that in addition to this one should be able to know that one has referred to oneself. What this comes to is knowing that the person who is referring is identical to the person who is referred to. But it does not follow from the qualitative description of any object, including a description of its mental properties, and its spatial location, that it is identical to the person referring to that object. Thus, to know that that object was identical to oneself, one would have to be able to identify it as oneself independently of one's ability to indirectly refer to it. If one could do that, then one could also refer to the oneself directly. Thus, give our initial assumption, everyone can refer to him or herself directly.

The epistemological gap that this argument relies on is reflected in familiar examples which show that how we conceive of any object we refer to by means of a description is different from how we conceive of ourselves independently of any such description. There is, e.g., always a gap between the knowledge that (i) the man sitting in front of the only computer in 252B Dauer Hall at the University of Florida is F and the knowledge that (ii) I am F. This is shown by the possibility of F being something which is of immediate concern to me, and yet my actions not being appropriate even if I know that (i).[9]

Another argument to the same conclusion relies on the assumption, made plausible by Cartesian thought experiments, that one could refer to oneself even if one were being fooled about the existence of a external world by an evil demon or were otherwise in massive error about the world. In this case,

---

[9] See David Kaplan, *Demonstratives* in *Themes from Kaplan,* ed. by Almog, Perry, and Wettstein (Oxford: Oxford University Press, 1989) and John Perry, "The Essential Indexical," *Nous* 13 (1979): 3-21.

we do not suppose that there would be any difficulty about knowing that we were referring to the self, although presumably we would not know that we were referring to any external object. For unique-reference this would require direct reference to the self. (The only plausible option would be that one could directly refer to a thought and refer to the self as the person whose thought this thought is. However, this would not yet settle that the referrer was identical to the person whose thought had been referred to.)

5. Knowledge by Acquaintance

In this section, I consider some further consequences of accepting that we can directly refer to the self-at-a-time.

The argument in section 3 for direct reference in thought was based on the assumption that we know that we can make unique-reference to spatio-temporal particulars. As we have said, this would be possible only if we made direct reference to some spatio-temporal particular, and in the previous section I have argued that we can and do directly refer to the self-at-a-time (for convenience I will sometimes shorten this to 'the self').   But we also know, very often, that we are on a given occasion uniquely-referring to a spatio-temporal particular, e.g., I am currently representing to myself a pen which I hold in my hand. The difference between this claim and the former one is that the former requires only that we know that we sometimes directly refer to an object, not that on any particular occasion we know that we are. If I know that I am referring to a particular as I form this thought, if I can in fact identify, in some way, the unique particular to which I am referring, then I must know also that I am referring to the particular which secures an origin to my referential coordinate system. If I did not know this, then I could not know that I was uniquely-referring to a particular pen. My knowledge here cannot be thought of as mediated by knowledge of the truth of any general statement, or of any

description of the self which relates it to another spatio-temporal particular. For neither of these would be sufficient for me to know that I refer to myself. For this reason it follows that my knowledge cannot be knowledge by inference or be based on any evidence. Thus, knowledge that I am uniquely-referring to a spatio-temporal particular requires a maximally direct knowledge that I am referring to myself.[10]

This is what I think is correct in Russell's doctrine of knowledge by acquaintance,[11] and it is required by our commitment to our knowing on particular occasions that we are uniquely-referring to spatio-temporal particulars. What I have in mind, however, is not exactly what Russell did. Russell's discussion is not very clear, and it would take us too far afield to explore it here. Briefly, though, there are at least two important ways in which the epistemic relation I have in mind is different from that which Russell had in mind. Russell treated knowledge by acquaintance as a kind of awareness of an object, and treated sense data and universals indifferently as the objects to which we are related by the relation of acquaintance. It is difficult to understand how our relation to universals, if we admit them, could be the same as our relation to sense data. It is also difficult to see how to apply the notion of direct knowledge to awareness, since this is naturally spelled out in terms of non-inferential knowledge that something is the case, and this contrast does not apply to awareness of something. In contrast, I treat acquaintance with the self as a maximally direct knowledge *that* I am referring to myself.

By our characterization above, the content of a thought is something others could have as well. But a thought about myself that is directly referring is not something that anyone else could have. The content of my thought, as we have characterized it, does not determine its referent. But my thought, if it is to have the appropriate role in my behavior, must be sensitive to its referent. Thus, the thought has a cognitive role in my thoughts and behavior that outstrips its content. Thus, we need a characterization

___

[10] If we think of this special epistemic relation to the self as a way of thinking of a referent, and call this way of thinking a sense, we will have a sense that determines the referent of this thought. This would be to adopt Gareth Evans's suggestion about how to understand Fregean senses (see "Understanding Demonstratives"). However, this would not be a sense as we have characterized it in section 2, and would not show that reference to the self is not direct, in the sense that we have characterized. I do not myself find calling this direct epistemic access a 'sense' or a 'way of thinking of an object' very helpful.

[11] Bertrand Russell, "Knowledge by Description and Knowledge by Acquaintance," in *Mysticism and Logic* (London: George Allen & Unwin, 1917): pp. 152-167.

of my thought that goes beyond its content. Let us the introduce the notion of truth conditional content: a thought's truth conditional content is that set of intrinsic properties which determines its truth conditions. (This corresponds to content$_k$.) In the case of a thought about oneself, a specification of the truth conditional content must make reference to an actual particular, the self whose thought it is. If in reference to spatio-temporal particulars, reference to the self is what secures unique-reference, the truth conditional content of every thought must include a reference to the self whose thought it is.

I now want to give two arguments to show that all we can refer to directly is the self-at-a-time. First, an application of Occam's razor argues against our referring directly to any other object because it is not necessary to account for our ability to uniquely-refer to objects. Second, grounding unique-reference in direct reference to the self offers a kind of solution to the puzzle we began with. That puzzle was how any notion could fulfill the role of content in (I)-(III) if thoughts were directly referring. As we have characterized it, content does not fulfill that role because it is conceived of as a type and consequently in principle something that different people can share; but then it could not determine the truth conditions of a thought that was directly referring. However, what we have above called truth-conditional content does satisfy that role provided that we can have first person knowledge of the truth conditions involved in such content; and this condition is satisfied in the special case in which the referent of the thought is the self. It would not be satisfied if it were any external object. For any other object is an object whose existence and character we can call into doubt. This was the puzzle that we started with about direct reference theories of thought: how could an object in the environment around me be referred to directly in thought, for it seems as if that would make it impossible for me to have the sort of access to it which we think we must have to our own thoughts. This problem is solved if the only object to which we can refer directly is the self. Thus, we should conclude that not only can we refer directly to the self, but that we can refer directly only to the self. Thus, the conditional conclusion of the previous paragraph becomes unconditional: every unique-reference to a spatio-temporal particular involves reference to the self. In this sense, the present picture represents all thought about the world around us as essentially ego-centric.[12]

---

[12] Cf. David Lewis, "Attitudes *De Dicto* and *De Se," Philosophical Review* 88 (1979): 513-43.

A corollary of this is that the knowledge I have of the content of any of my thoughts if that knowledge extends to its truth conditions includes knowledge of the existence of the self; thus knowledge of every thought includes this acquaintance with the self.

I now want to address two requests for explanation that might be pressed on a account of the sort that I have advanced. (1) How are we able to refer directly to the self? (2) How are we able to know that we do so?

(1) In a certain sense, there is no answer to this question. We have inferred that we do refer directly to the self from a number of other assumptions we are inclined to hold. There is a answer to the question how we refer to an object if content can determine the referent by itself, or if it is an object that we refer to indirectly, for then we can say what relation that object bears to something to which we can directly refer. But neither of these answers can be given for the case of direct reference itself, and to demand them is to misunderstand what direct reference is. Direct reference functions as a foundation for reference to spatio-temporal particulars, and is not secured by satisfaction of any content. If there were an explanation of how we are able to refer directly to the self, we would not be referring to the self directly. Direct reference to the self must be taken as primitive. It is clear that if every thought requires a bearer, it will not be possible to have a thought without the possibility of referring to the self, that is, there could be no vacuous thought about the self, since the existence of the thought itself guarantees the existence of the self. But this observation does not seem to amount to an explanation of how we refer to the self, any more than a similar observation would help explain how one could refer to the very sentence that one is uttering on a give occasion. The only explanation we can give of it consists in showing what role it plays in our referring to other things, and our acting coherently, thereby showing its necessity.

(2) The same thing must be said about our second question. There are two possible answers to this question: we know that we refer to the self because we know that we have a thought whose content uniquely picks out the self, or we know that we refer to the self because we know we are referring to something else and can specify a unique relation that thing bears to the self. Both of these answers are ruled out by the hypothesis that we are referring to the self directly. It follows that the

question cannot be answered; for it has a false presupposition. Just as we must take direct reference to the self as a semantically primitive relation, so we must take acquaintance with the self as an epistemologically primitive relation. To the extent that we can give an explanation of it, it is in terms of its role in explaining how we can know other things that we suppose ourselves to know.

6. Thought and talk

In this section, I want to explore briefly and incompletely some questions about the relations between these conclusions about reference in thought and the semantics of directly referring expressions, in particular, indexicals and demonstratives, in ordinary language. (For convenience I will use the expression 'indexicals' to refer to expressions in both of these classes; where a distinction is necessary I will talk about true indexicals and demonstratives.) (i) Are directly referring terms required in language? (ii) Are indexicals for referring to times and the self required in language? (iii) Are indexicals required in language? (iv) Are all indexical expressions reducible to a canonical form or translatable into expressions containing other indexicals? Turning to more general questions: (v) How should we conceive of the relation between direct reference in thought and in language in general? (vi) Do directly referring expressions in language show that some thought contents are relational in character?

(i) The first conclusion we can draw is that any language in which spatio-temporal particulars are uniquely-referred to in the absence of knowledge of the large scale structure of the universe must include directly referring elements. Expressions in natural languages do not have their semantic properties independently of the capacities of their speakers. Therefore, if it is possible for us to uniquely-refer to spatio-temporal particulars using expressions in our language, those expressions cannot refer solely in virtue of their senses; for if they could, this would be because we were able in thought to uniquely-refer to a spatio-temporal particular without referring directly to any spatio-temporal particular. By the argument in section 3, we cannot do this. In particular, any singular term

which determines a unique spatio-temporal particular cannot do so in virtue of its sense alone. Thus, any uniquely-referring singular term is a directly referring term.[13]

It is clear that for cognitively limited agents like ourselves having the ability to directly refer to some object is very important, for it enables us to identify, think, and talk about objects even in less than ideal epistemic conditions.

(ii) If we are able to uniquely-refer to spatio-temporal particulars because we can refer to the self, does it follow that our language must contain indexical expressions designed specifically for referring to the self and times? I do not think it does, for it seems imaginable that we could purge pronouns and other true indexicals in favor of using just demonstratives such as 'that' in contexts with auxiliary descriptions or extralinguistic hints which would allow one's interlocutor to pick out what one is referring to. (David Kaplan's 'dthat' operator is such a device.[14]) For example, by uttering 'that person' and pointing to oneself one could achieve reference to oneself independently of using the first person pronoun; the pointing need not be considered a part of the semantics of any expression of the language but, instead, an extra-linguistic aid to the auditor to figure out what is being referred to. In appropriate circumstances, even the pointing could be eliminated. However, how the referent of a directly referring term is secured would have to be ultimately understood in terms of the capacity of individuals to refer directly to themselves at a time.

(iii) Are indexical expressions required in a language in which spatio-temporal particulars are uniquely-referred to? Since we require directly referring terms in any language in which unique-reference is made by speakers of the language, indexicals could be dispensed with only if we could refer to spatio-temporal particulars using expressions which functioned as proper names do, that is, functioned as expressions which directly refer but do not change their reference on different occasions.

---

[13] Definite descriptions that denote uniquely spatio-temporal particulars are not counterexamples to this claim. For remember that we defined 'uniquely referring' as referring in a way that succeeds even in a spatio-temporally symmetrical universe. A definite description may happen to denote a unique individual, but it would do so by uniquely referring only if it were conjoined with an indexical.

[14] See David Kaplan, "Dthat" in *Syntax and Semantics,* ed. P. Cole (New York: Academic Press, 1978): pp. 221-43.

It might be thought that this is not be possible because to introduce a name with a certain referent one would need to already to have available linguistic devices which played the role of indexicals. However, this does not seem to be necessary. One can learn the referent of a name without having it demonstrated for one, by, e.g., watching people who use the name, seeing where they look, and what they say. We learn many names in this way, and it seems imaginable that we could acquire all of them in this way.

The strongest argument for the necessity of indexicals in a language is the need to refer to a potentially indefinitely large number of new particulars. If we tried to introduce proper names for each new particular to which we wished to make a passing reference, we would in a very short time exhaust our capacity to keep track of which names refer to which entities. Having terms which can have different referents on different occasions is for practical purposes indispensable. Still, this argument does not establish that it is impossible for there to be a language in which there are directly referring terms but no indexicals or demonstratives. Such languages are conceivable in cases in which we are not interested in talking about very many particulars, or in cases in which one has a capacity to name and keep track of a very large number of individuals. God, e.g., would be capable of speaking a language in which every directly referring term was a proper name.

(iv) Are all indexical expressions reducible to a canonical form or translatable into other expressions which contain indexicals? It is tempting when analyzing indexical expressions to provide a translation of every indexical expression into a canonical form. For example, it might be suggested that every indexical expression could be translated into an expression which includes 'I' and 'now'. Thus, for instance, one might suggest that 'you' means 'the person I am now talking to', or that 'that' means 'what I am now designating'. I think this is a mistake, even if the translation seems to provide an expression that picks out the same individual as the original expression in every context of use. I will give two reasons for this. The first is that indexicals can fulfill their role without our having to suppose that there are any synonyms of them or that they are expressible in terms of expressions containing other indexicals. The second is that every obvious 'translation' fails to be synonymous with the original

expression because it entails something that the original does not. The general point can be made by considering an example, so I will concentrate on demonstratives.

In investigating this question, it is instructive to begin with non-indexical directly referring terms such as proper names. Proper names do not have a semantic value other than their referent. (I put aside the fact that to understand the use of a proper name one must understand its role in the language, which is of course different from knowing its referent; the referent is all that it contributes to the truth conditions of the sentences in which it appears.) Clearly, no proper name is equivalent to any expression in which 'I' and 'now' appear, for unlike indexicals, the referents of proper names do not vary across contexts of use. The referent of a proper name must be understood to be fixed for individuals by reference to each self who uses the proper name. (Though of course this doesn't mean that one can't misuse the name.) How different individuals pick out the referent of a proper name, however, makes no difference to its semantic role in the language. A case of particular interest is the original introduction of a proper name. Take a case in which I simply introduce a name in the course of a conversation — perhaps I introduce a nickname for someone, 'The Rat', I call him. Understanding the role of proper names in language, my interlocutors will recognize my intention to introduce a name for someone, and look for an appropriate referent, using whatever clues they can gather from the context. Clearly, however, these clues are not a part of a sense associated with the name itself.

Now consider the demonstratives 'this' and 'that'. These function as variable names, in the sense that they designate particulars without having any accompanying sense that determines which individual is designated, and that they can designate different individuals on different occasions. The semantic facts about demonstratives are that they are singular terms and that on different occasions of use they have different referents. Each new use of a demonstrative requires a new determination of a referent on the part of the speaker; the referent is not given by the demonstrative itself. It is the speaker who refers and not the term; the object referred to is therefore determined by the individual's referential intentions. The similarities to names suggest that no translation should be required to explicate the function of these words. We can think of 'this' and 'that' as bare referring terms, whose referents are determined on occasions of their use in the same way the referent of 'The Rat' was determined in the

example in the previous paragraph. The difference in the case of demonstratives is that there is no presumption that the next time I use the expression I will be referring to the same entity. Since it is the individual's referential intentions that determine what on a particular occasion a demonstrative refers to, we can then give the following account of how to determine the referent of a demonstrative on a particular occasion of use: the referent of a demonstrative on a particular occasion of use is whatever the individual using the demonstrative intends primarily to be referring to. How we figure out what an individual is referring to, however, just as in the case of introducing a new proper name, is not a part of the semantics of the demonstrative expression. We do this in any way we can.

In support of this, notice that it seems possible to introduce into a language a device which we think of as operating in exactly this way. That is, we could introduce into a language a device which we stipulate (1) functions to pick out an object of which something can then be predicated, and (2) picks out whatever the speaker using the device intends primarily as the object of predication. Such a device would have exactly the properties I have claimed that demonstrative expressions have.

Thus, the function of demonstratives clearly does not require that we think of them as being equivalent to any other expression. We understand their function when we recognize them as devices for denoting different objects on different occasions and know that the object denoted is determined by the speaker's referential intentions.

Let us turn to the second objection. I have argued so far only that there is no reason to suppose that demonstratives are equivalent to any other expressions containing indexicals. The temptation is, however, to treat the rule we give for determining the referent of a demonstrative as a synonymous expression. This is clearly a mistake, however, because even if the rule gives the referent of the demonstrative on various occasions of use, it has different entailment relations. Suppose that we attempt to translate 'that' as 'what I am now designating'. Then the following two sentences should be semantically equivalent:

(1)   That is an apple

and

(2)   What I am now designating is an apple.

Thus, if each is uttered on the same occasion, the proposition each expresses should entail the same propositions. But they do not. For example, the proposition expressed by the second, uttered on a given occasion, but not that expressed by the first, would entail that a speaker exists, that I am now designating something, etc. The same objection can be made to treating demonstratives and indexicals as token-reflexive terms, that is, as being implicitly metalinguistic. The truth of the proposition expressed by (1) on a given occasion of utterance does not entail that any words have been uttered, any more than the truth of '2 + 2 = 4' as uttered on a given occasion. The semantic function of demonstratives is to pick out an object; part of their sense is given by a rule for determining the referent or at least a characterization of what determines the referent. But the rule is not semantically equivalent to the original expression, for if it were, then the semantic role of the indexical would go beyond simply picking out an object for predication. Indeed, it might be suggested that grasping the sense of a demonstrative simply consists in our acquiring the ability to determine systematically which object is in question on its various occasions of use (*ceteris paribus,* as always), that is, not in anything explicitly representable in other words but in a certain skill.

We can give a similar story in the case indexicals such as 'you', 'he', 'yesterday', etc. These differ from pure demonstratives like 'this' and 'that' in that their function in the language puts restrictions on what kinds of objects they can refer to. Here too, we should resist the temptation to provide synonymous expressions, such as 'the object I am now speaking to' for 'you'. In these cases as well, it is clear that we will get different entailment relations. Rather, we should think of their contribution to the semantics of an expression as their referent, and perhaps a sortal, and think of the rule for picking out the referent or constraining what the appropriate referent can be as determined by the sense but not as giving a synonym.

(v) How should we conceive of the relation between direct reference in thought and in language in general? From what we have said above, it should be clear that there is a considerable gap between the semantics of an sentence a person utters and the content of a person's thought about the object

which any directly referring term in the sentence he utters picks out. If the picture which I have sketched in the preceding sections is correct, then all reference by an individual in thought to spatio-temporal particulars is grounded in his ability to refer directly to himself. The semantics of most indexical terms in a natural language, however, do not involve any reference to the person who uses them. Not only would this seem excessively ego-centric, it is not required for the purposes of talking about objects around us. All that is required is that different parties to a conversation are able to pick out the same object the speaker picks out and refers to using an indexical expression or other directly referring term. Since each individual himself secures reference to spatio-temporal particulars in a different way from anyone else, it is necessary that referring terms in a public language abstract from this, for the sense and truth conditions of sentences in a public language must be graspable by more than one person. Since different people of necessity take different referential routes in thought to the same objects, we want the semantics of a public language to focus just on the object picked out. Thus, of necessity, the semantics of an expression used in a public language must underdetermine the truth conditional content of the thought of someone who uses it. If my account is correct, the only indexicals that reflect directly an underlying feature of thought are the first person pronoun and an indexical for referring to the present time. Even here, of course, the semantic function is to pick out the object. The speaker does it directly. The auditor by description.

(vi) Do directly referring expressions in language show that some thought contents are relational in character? After the preceding discussion, we can answer this question very briefly. Arguments that there are singular thoughts about the world around us that are object involving rely on taking the content of sentences used to characterize a speaker's attitude in the that-clause of a propositional attitude to give the content of his thought.[15] For example, in a sentence of the form,

(S)   Bob believes that *that* is an apple,

---

[15] See John McDowell "Singular Thought and the Extent of Inner Space," in *Subject, Thought and Context,* ed. by John McDowell and Philip Pettit (Oxford: Oxford University Press, 1986) and Gareth Evans, *Varieties of Reference* (Oxford: Clarendon Press, 1982).

the fact that the semantic value of 'that' is simply the object which it denotes on this occasion of use is taken to show that Bob has an object involving belief. But from our present perspective this can be seen to be a mistake. What 'that' refers to in the embedded sentence depends upon the referential intentions of the speaker. How Bob picks out the apple is independent of how the speaker does. That the speaker uses a directly referring term in characterizing Bob's attitude no more shows that Bob has an object involving thought than the speaker's use of a directly referring term shows that he has a thought with an apple as a constituent. We use sentences like (S) because very often we are not interested in how an individual picks out an object but only in what he thinks about it. The only singular thoughts, that is, object involving thoughts, are thoughts that individuals have about themselves. (Similar deflationary readings can be given of predicates which are apparently directly referring as well, though because the treatment would be more involved, I will not undertake it here.)[16]

7. Conclusion

To conclude, let me summarize some of the results we have reached.

1. Directly referring terms are a commonplace in language.

2. Nothing follows directly about whether a direct reference theory for thought is correct from a direct reference theory being correct for some terms of natural languages.

---

[16] Similar objections to externalist theories have been raised by Brian Loar (see e.g., "Social Content and Psychological Content," in *Contents of Thought,* ed. by Robert Grimm and Daniel Merrill (Tucson: University of Arizona Press, 1988)), though I believe Loar makes too many concessions to the externalist; the point with respect to singular terms has also be urged before, e.g., by Searle in "Referential and Attributive," in his *Expression and Meaning,* (Cambridge: Cambridge University Press, 1979) and in *Intentionality.*

3. A direct reference theory for thought is forced on us by our holding four things: that the universe may be spatio-temporally symmetrical, that we do not know whether it is or not, and that we both uniquely-refer and know we uniquely-refer to spatio-temporal particulars.

4. The most conservative solution to the problem of how unique-reference to spatio-temporal particulars is possible is that we can refer directly to the self-at-a-time. This is sufficient in principle for referring to spatio-temporal particulars around us, and if we are to know that we refer sometimes to the self, it is necessary.

5. A theory of direct reference for thought, if we argue for it as in section 5, requires that we have maximally direct knowledge that we refer to the self on some occasions.

6. Both direct reference to the self and the direct knowledge we have that we refer to the self must be taken as primitive relations. We know that we bear these relations to the self because they are required to explain the possibility of our knowingly uniquely-referring to spatio-temporal particulars.

7. If we refer to spatio-temporal particulars by means of referring directly to the self-at-a-time, then any language we speak which includes terms that refer to spatio-temporal particulars must contain directly referring terms.

8. A language which contains directly referring terms does not have to contain a first person pronoun or any other specialized indexicals.

9. It is in principle possible for a language to contain no indexical expressions at all, but it is a practical impossibility for finite beings like ourselves who need to be able to uniquely-refer to an indefinitely large number of particulars.

10. Indexicals should not be thought of as reducible to a canonical form or to be intertranslatable: it is not required for their linguistic function that they be strictly synonymous with any other expressions in a language, and sentences in which we substitute the rules that can be given to characterize how to determine the referents of indexical expressions for those expressions do not have the same entailment relations as the original.

11. It is a mistake to read off from the semantics of a sentence in a that-clause of a propositional attitude attribution the content of the thought of the person to whom an attitude is attributed.

12. Consequently, there is no reason found in ordinary usage to think that there are any singular thoughts other than thoughts about the self.