Book summary
From perception to communication
Robin Cooper

The aim of this book is to characterize a notion of type which will cover both linguistic and non-linguistic action and to lay the foundations for a theory of action based on these types. We will argue that a theory of language based on action allows us to take a perspective on linguistic content which is centered on interaction in dialogue and that this is importantly different to the traditional view of natural languages as being essentially similar to formal languages such as logics developed by philosophers or mathematicians. At the same time we will argue that the tremendous technical advances made by the formal language view of semantics can be incorporated into the action-based view and that this can lead to important improvements both of intuitive understanding and empirical coverage. In this enterprise we use types rather than possible worlds as commonly employed in studies of the semantics of natural language. Types as more tractable than possible worlds and give us more hope of understanding the implementation of semantics both on machines and in biological brains.

Part I of the book (Chapters 1–3) deals with a theory of types related to perception and action and shows a way of presenting a theory of grammar within a theory of action. Part II (Chapters 4–8) then looks at a number of central issues in semantics from a dialogical point of view and argues that there are advantages to looking at some old puzzles from this perspective.

In Chapter 1 we introduce a notion of perception of an object or situation as making a judgement that it is of a type. We present basic notions of the theory of types which will be developed in the book, TTR, a type theory with records, which builds to a great extent on ideas taken from the type theory of Per Martin-Löf. The overall approach presented here owes much to the theory of situations and situation semantics presented by Barwise and Perry in the 1980's.

In Chapter 2 we introduce some basic notions of a theory of action based on these types and apply the theory of types to basic notions of information update in dialogue. Here we build on seminal work on dialogue analysis by Jonathan Ginzburg and also related computational implementation by Staffan Larsson leading to the information state update approach to dialogue systems. We have adapted these ideas in a way that allows us to pursue the questions of grammar and semantics that we take up in the remainder of the book.

In Chapter 3 we show how syntax and semantics can be embedded in the theory of action characterized in Chapters 1 and 2. Grammatical rules are regarded in terms of affordances which license us to draw conclusions about the types to be associated with speech events on the basis of speech events previously perceived. This is in contrast to a formal language view where language is seen as a set of analyzed strings of symbols associated with meanings of some kind. The philosophical ground of the action-based approach goes back to the relational theory of meaning introduced in Barwise and Perry's situation semantics which focusses on the relation between utterance situations and described situations. It also enables us to develop a formal approach to dialogical theories of language such as those developed by the psychologist Herb Clark and the linguist Per Linell. A recent theory to which the ideas in this chapter are related is that of Dynamic Syntax (DS).

The theory of types that we employ gives us two notions which will be important in the

development of semantics in Part II. The first is the notion of *intensionality*. Types in TTR are intensional in that the identity of a type is not established in terms of the set of witnesses of that type. That is, types are not *extensional* in the way that sets are in a standard set theory. The second notion has to do with the facts that the types themselves are treated as objects that can enter into relations and be used to construct new types. In our terms, an important enabling factor for human language is that we not only can perceive objects and situations in terms of types and act on these perceptions but that we can also reason about and act on the types themselves, for example, in ascribing them to other agents as beliefs or making a plan to achieve a goal by creating an event of a certain type. The types become cognitive *resources* which we can exploit in our communicative activity. In Part II we look at a number of examples of this.

In Chapter 4 we examine reference by uses of proper names and occurrences of pronouns which are not bound by quantifiers. In order to account for this we need a notion of *parametric content*, which is to say that the content of an utterance depends on a context belonging to a certain type. For example, an utterance of the proper name *Sam* requires a context in which there is an individual named "Sam". But where in her resources should a dialogue participant look for such a context? One obvious place is the conversational gameboard that we introduced in Chapter 2. Another place is the visual scene, or more generally the ambient situation which the agent can perceive by different sense modalities. This we also represent as a resource using a type – that is, the type for which the ambient situation would be a witness if the agent's perception is correct. Yet another place to look is the agent's long term memory. This resource is also modelled as a type representing how the world would be if the agent's memory or beliefs are correct. The fact that we are reasoning about the extent to which the context type associated with the utterance matches the types modelling the agent's relevant resources enables us to talk about cases where there are names of non-existent objects (that is, the agent's resource types do not exactly match the world) or where a single object in the world corresponds to two objects in the resources or *vice versa* (another way in which there can be a mismatch between reality and an agent's resources). The proposal to represent different aspects of mental states in terms of record types is closely related to (and inspired by) similar proposals for representing mental states using discourse representation structures.

In Chapter 5 we look at frames associated with common nouns. The idea of frames goes back to early work on frame semantics by Fillmore and also psychological work on frames by Barsalou. We will construe frames as situations (modelled as records in TTR). We will argue that frame types are an additional kind of resource which is exploited in natural language semantics. A common noun like *dog*, in addition to being associated with the property of being a dog, can also be associated with a type of situation (a frame type) which is common for dogs, for example, where the dog has a name, an age and various other attributes we commonly attribute to dogs. We will argue that such a frame can play an important role in interpreting utterances such as *the dog is nine* in the sense of "the dog is nine years old". Some nouns, such as *temperature*, seem to represent frame level predicates, following an analysis suggested by Sebastian Löbner in order to account for the analysis of utterances like *the temperature is rising* where it is not the case that some particular temperature is rising (say, 30°) but that different situations (frames) with different temperatures are being compared. Nouns which normally predicate of individuals can be coerced to predicate of frames. An example is the noun *ship* in an example originally

discussed by Manfred Krifka: *four thousand ships passed through the lock* which can either mean that four thousand distinct ships passed through the lock or that there were four thousand ship-passing-through-the-lock events some of which may have involved the same ship. We argue that in order to interpret such examples you need to have as a resource an appropriate frame type associated with the noun *ship*.

In Chapter 6 we explore phenomena in natural language which are standardly referred to as *modal* and *intensional*. We argue that types as we conceive them are better placed to deal with these phenomena than the possible worlds that are used in standard formal semantics. In standard formal semantics propositions are regarded as sets of possible worlds. For example, the proposition corresponding to *a boy hugged a dog* is the set of all logically possible worlds in which a boy hugged a dog is true. What we substitute for this is the type of situations in which a boy hugged a dog. At an intuitive level these notions are quite similar. They both represent mathematical objects which allow for many different possibilities as long as the fact that a boy hugged a dog is held constant across them. One important difference is that sets of possible worlds are extensional sets whereas as our types are intensional. Thus it is possible for us to have two distinct types which have exactly the same witnesses.

In Chapter 7 we look at generalized quantifiers from the perspective of dialogic interaction. Traditionally generalized quantifiers are treated as sets of sets or sets of properties and the work of Barwise and Cooper on generalized quantifiers built on this idea. Barwise and Cooper also introduced the auxiliary notion of witness set for quantifiers under the heading "Processing quantified statements". In this chapter we turn things around and make the characterization of witness sets the primary notion in defining quantifiers. This makes it more straightforward to account for the anaphoric possibilities relating to quantified expressions in dialogue. We often use quantified statements in dialogue when we have inadequate information to determine their truth. This is particularly true of determiners like *every* and *most* when talking about large sets. We suggest that this phenomenon can be analyzed by estimating a probability based on the evidence presented in our cognitive resources (long-term memory or beliefs as discussed in Chapters 4 and 6).

In Chapter 8 we give an account of how TTR types can be used to talk of content which is underspecified. The idea is to exploit the notion of types which can have several witnesses as "underspecifications" of those witnesses. Rather than associating contents with utterances as we have done in the earlier chapters, we associate *types* of contents with utterances. We show how earlier ideas about the treatment of underspecification of quantifier scope and anaphora can be accommodated in this view.

I have tried to write this book so that it might be accessible to readers who have a basic knowledge of set theory and logic. I have also separated out the technical development by marking it with shading without compromising the technical detail which is necessary for a thorough formal development. In this way I hope that it may be accessible to graduate students and even in some cases advanced undergraduates. The book should be of interest to linguists, philosophers, psychologists, cognitive scientists and computational linguists with an interest in language as action and the formal modelling of cognition and natural language understanding.