

Type theory and language
From perception to linguistic communication

Robin Cooper

Draft, February 27, 2013
PLEASE DO NOT QUOTE

Contents

Acknowledgements	v
1 From perception to intensionality	1
1.1 Perception as type assignment	1
1.2 Modelling type systems in terms of mathematical objects	3
1.3 Situation types	5
1.4 The string theory of events	13
1.5 Doing things with types	20
1.6 Modal type systems	20
1.7 Intensionality: propositions as types	23
2 Information exchange	27
2.1 Speech events	27
2.2 Signs	31
2.3 Information exchange in dialogue	33
2.4 Resources	50

3	Grammar	51
3.1	Syntax	51
3.2	Semantics	59
4	A semantic benchmark	73
4.0.1	Lexicon	75
4.0.2	Constraints on lexical predicates	81
4.0.3	Syntactic rules with compositional semantics	83
5	Coordination and learning	95
A	Type theory with records	97
A.1	Basic types	97
A.2	Complex types	98
A.2.1	Predicates	98
A.2.2	Systems of complex types	99
A.3	Function types	100
A.4	List types	100
A.5	Set types	101
A.6	Join types	101
A.7	Meet types	101
A.8	Models and modal systems of types	102
A.9	Strings and regular types	103

<i>CONTENTS</i>	iii
A.10 The type <i>Type</i> and stratification	104
A.11 Record types	106
A.12 Merges of record types	111
B Grammar rules	115
C Dialogue rules	117
D First order logic	119
D.1 Interpreting first order logic in TTR	119
D.2 Examples of first order logic interpretations	121
D.3 Examples of inference in first order logic	122
E Intensional logic	125
E.1 Interpreting intensional logic in TTR	125
E.2 Examples of intensional logic interpretations	130
E.3 Examples of inference in intensional logic	130

Acknowledgements

This research was supported in part by the following projects: Records, types and computational dialogue semantics, Vetenskapsrådet, 2002-4879, Library-based grammar engineering, Vetenskapsrådet, 2005-4211, and Semantic analysis of interaction and coordination in dialogue (SAICD), Vetenskapsrådet, 2009-1569.

Chapter 1

From perception to intensionality

1.1 Perception as type assignment

Kim is out for a walk in the park and sees a tree. She knows that it is a tree immediately and does not really have to think anything particularly linguistic, such as “Aha, that’s a tree”. As a human being with normal visual perception, Kim is pretty good at recognizing something as a tree when she sees it, provided that it is a fairly standard exemplar, and the conditions are right: for example, there is enough light and she is not too far away or too close. We shall say that Kim’s perception of a certain object, *a*, as a tree involves the ascription of a type *Tree* to *a*. In terms of modern type theory (as in Martin-Löf (1984); Nordström *et al.* (1990)), we might say that Kim has made the *judgement* that *a* is of type *Tree* (in symbols $a : Tree$).

Objects can be of several types. An object *a* can be of type *Tree* but also of type *Oak* (a subtype of *Tree*, since all objects of type *Oak* are also of type *Tree*) and *Physical Object* (a supertype of *Tree*, since all objects of type *Tree* are of type *Physical Object*). It might also be of an intuitively more complicated type like *Objects Perceived by Kim* which is neither a subtype nor a supertype of *Tree* since not all objects perceived by Kim are trees and not all trees are perceived by Kim.

There is no perception without some kind of judgement with respect to types of the perceived object. When we say that we do not know what an object is, this normally means that we do not have a type for the object which is narrow enough for the purposes at hand. I trip over something in the dark, exclaiming “What’s that?”, but my painful physical interaction with it through my big toe tells me at least that it is a physical object, sufficiently hard and heavy to offer resistance to my toe. The act of perceiving an object is perceiving it *as* something. You cannot perceive something without ascribing some type to it, even if it is a very general type such as *thing* or *entity*.

Recognizing something as a tree may be immediate and not involve conscious reasoning. Recognizing a tree as an aspen, an elm or a tree with Dutch elm disease may involve closer inspection and some conscious reasoning about the shape of the leaves or the state of the bark. For humans the relating of objects to certain types can be the result of a long chain of reasoning involving a great deal of conscious effort. But whether the perception is immediate and automatic or the result of a conscious reasoning process, from a logical point of view it still seems to involve the ascription of a type to an object.

The kind of types we are talking about here correspond to pretty much any useful way of classifying things and they correspond to what might be called properties in other theories. For example, in the classical approach to formal semantics developed by Montague (1974) and explicated by Dowty *et al.* (1981) among many others, properties are regarded not as types but as functions from possible worlds and times to (the characteristic functions of) sets of entities, that is, the property *tree* would be a function from possible worlds and times to the set of all entities which are trees at that world and time. Montague has types based on a version of Russell's (1903) simple theory of types but they were "abstract" types like *Entity* and *Truth Value* and types of functions based on these types rather than "contentful" types like *Tree*. Type theory for Montague was a way of providing basic mathematical structure to the semantic system in a way that would allow the generation of interpretations of infinitely many natural language expressions in an orderly fashion that would not get into problems with logical paradoxes. The development of type theory which we will undertake here can be regarded as an enrichment of an "abstract" type theory like Montague's with "contentful" types. We want to do this in a way that allows the types to account for content and relate to cognitive processing such as perception. We want our types to have psychological relevance and to correspond to what Gibson (1986) might call *invariants*, that is, aspects that we can perceive to be the same when confronted with similar objects or the same object from a different perspective. In this respect our types are similar to notions developed in situation theory and situation semantics (Barwise and Perry, 1983; Barwise, 1989).

Gibson's notion of attunement is adopted by Barwise and Perry. The idea is that certain organisms are attuned to certain invariants while others are not. Suppose that Kim perceives a cherry tree with flowers and that a bee alights on one of the flowers. One assumes that the bee's experience of the tree is very different from Kim's. It seems unlikely that the bee perceives the tree as a tree in the sense that Kim does and it is not at all obvious that the bee perceives the tree in its totality as an object. Different species are attuned to different types and even within a species different individuals may vary in the types to which they are attuned. This means that our perception is limited by our cognitive apparatus – not a very surprising fact, of course, but philosophically very important. If perception involves the assignment of types to objects and we are only able to perceive in terms of those types to which we are attuned, then as Kant (1781) pointed out we are not actually able to be aware of *das Ding an sich* ("the thing itself"), that is, we are not able to be aware of an object independently of the categories (or types) which are available to us through our cognitive apparatus.

1.2 Modelling type systems in terms of mathematical objects

In order to make our theory precise we are going to create models of the systems we propose as mathematical objects. This represents one of the two main strategies that have been employed in logic to create rigorous theories. The other approach is to create a formal language to describe the objects in the theory and define rigorous rules of inference which explicate the properties of the objects and the relations that hold between them. At a certain level of abstraction the two approaches are doing the same thing – in order to characterize a theory you need to say what objects are involved in the theory, which important properties they have and what relations they enter into. However, the two approaches tend to get associated with two different logical traditions: the *model theoretic* and *proof theoretic* traditions.

The philosophical foundation of type theory (as presented, for example, by Martin-Löf (1984)) is normally seen as related to intuitionism and constructive mathematics. It is, at bottom, a proof-theoretic discipline rather than a model-theoretic one (despite the fact that model theories have been provided for some type theories). However, it seems that many of the ideas in type theory that are important for the analysis of natural language can be adopted into the classical set theoretic framework familiar to linguists from the classical model-theoretic canon of formal semantics starting from Montague (1974).

Your theory is not very interesting if it does not make *predictions*, that is, by making certain assumptions you can infer some conclusions. This gives you one way to test your theory: see what you can conclude from premises that you know or believe to be true and then test whether the conclusion is actually true. If you can show that your theory allows you to predict some conclusion and its negation, then your theory is *inconsistent*, which means that it is not useful as a scientific theory. One way to discover whether a theory is consistent or not is to formulate it very carefully and explicitly so that you can show mathematical properties of the system and any inconsistencies will appear.

From the informal discussion of type theory that we have seen so far it is clear that it should involve two kinds of entity: the types and the objects which are of those types. (Here we use the word “entity” not in the sense that Montague did, that is, basic individuals, but as an informal notion which can include both objects and types.) This means that we should characterize a type theory with two domains: one domain for the objects of the types and another domain for the types to which these objects belong. Thus we see types as theoretical entities in their own right, not, for example, as collections of objects. Diagrammatically we can represent this as in Figure 1.1 where object a is of type T_1 .

A system of basic types consists of a set of types which are *basic* in the sense that they are not analyzed as *complex* objects composed of other objects in the theory. Each of these types is associated with a set of objects, that is, the objects which are of the type, that is the *witnesses* for the type. Thus if T is a type and A is the set of objects associated with T , then a is of type

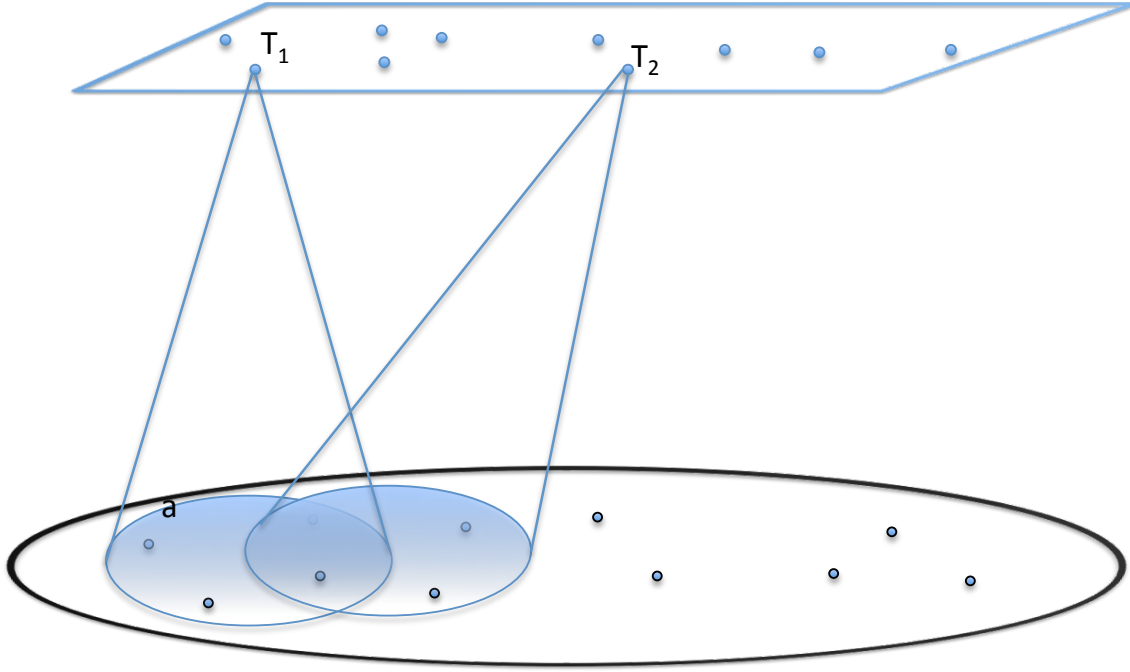


Figure 1.1: System of basic types

T (in symbols, $a : T$) just in case $a \in A$. We require that any object a which is a witness for a basic type is not itself one of the types in the system. A type may be *empty* in the sense that it is associated with the empty set, that is, there is nothing of that type.

Notice that we are starting with the types and associating sets of objects with them. This means that while there can be types for which there are no witnesses, there cannot be objects which do not belong to a type. This relates back to our claim in Section 1.1 that we cannot perceive an object without assigning a type to it.

Notice also that the sets of objects associated with types may have members in common. Thus it is possible for objects to belong to more than one type. This is important if we want to have basic types *Elm*, *Tree* and *Physical Object* and say that a single object a belongs to all three types as discussed in Section 1.1.

An extremely important property of this kind of type system is that there is nothing which prevents two types from being associated with exactly the same set of objects. In standard set theory the notion of set is *extensional*, that is sets are defined by their membership. You cannot have two distinct sets with the same members. The choice of defining types as entities in their own right rather than as the sets of their witnesses, means that they can be *intensional*, that is, you can have

more than one type with the same set of witnesses. This can be important for the analysis of natural language words like *groundhog* and *woodchuck* which (as I have learned from the literature on natural language semantics) are the same animal. In this case one may wish to say that you have two different words which correspond to the same type, rather than two types with the same *extension* (that is, set of witnesses). Such an analysis is less appealing in the case of *unicorn* and *centaur*, both mythical animals corresponding to types which have an empty extension. If types were extensional, there would only be one empty type (just as there is only one empty set in set theory). In the kind of possible world semantics espoused by Montague the distinction between *unicorn* and *centaur* was made by considering their extension not only in the actual world (where both are empty) but also in all possible worlds, since there will be some worlds in which the extensions are not the same. However, this kind of possible worlds analysis of intensionality fails when you have types whose extensions cannot possibly be different. Consider *round square* and *positive number equal to 2 – 5*. The possible worlds analysis cannot distinguish between these since their extensions are both empty no matter which possible world you look at.

Finally, notice that there may be different systems of basic types, possibly with different types and different objects. One way of exploiting this would be to associate different systems with different organisms as discussed in Section 1.1. (Below we will see different uses of this for the analysis of types which model the cognitive system of a single agent.) Thus properly we should say that an object a is of type T with respect to a basic systems of types \mathbf{TYPE}_B , in symbols, $a :_{\mathbf{TYPE}_B} T$. However, we will continue to write $a : T$ in our informal discussion when there is no danger of confusion.

The definition of a system of basic types is made precise in Appendix A.1.

What counts as an object may vary from agent to agent (particularly if agents are of different species). Different agents have what Barwise (1989) would call different *schemes of individuation*. There appears to be a complex relationship between the types that an agent is attuned to and the parts of the world which the agent will perceive as an object. We model this in part by allowing different type systems to have different objects. In addition we will make extensive use in our systems of a basic type *Ind* for “individual” which corresponds to Montague’s notion of “entity”. The type *Ind* might be thought of as modelling a large part of an agent’s scheme of individuation in Barwise’s sense. However, this clearly still leaves a great deal to be explained and we do this in the hope that exploring the nature of the type systems involved will ultimately give us more insight into how individuation is achieved.

1.3 Situation types

Kim continues her walk in the park. She sees a boy playing with a dog and notices that the boy gives the dog a hug. In perceiving this event she is aware that two individuals are involved and

that there is a relation holding between them, namely hugging. She also perceives that the boy is hugging the dog and not the other way around. She sees that a certain action (hugging) is being performed by an agent (the boy) on a patient (the dog). This perception seems more complex than the classification of an individual object as a tree in the sense that it involves two individual participants and a relation between them as well as the roles those two individuals play in the relation. While it is undoubtedly more complex than the simple classification of an object as a tree, we want to say that it is still the assignment of a type to an object. The object is now an event and she classifies the event as a hugging event with the boy as agent and the dog as patient. We shall have complex types which can be assigned to such events.

Complex types are constructed out of other entities in the theory. As we have just seen, cognitive agents, in addition to being able to assign types to individual objects like trees, also perceive the world in terms of states and events where objects have properties and stand in relations to each other – what Davidson (1967) called events and Barwise and Perry (1983) called situations. We introduce types which are constructed from predicates (like ‘hug’) and objects which are arguments to this predicate like a and b . We will represent such a constructed type as $\text{hug}(a,b)$ and we will sometimes call it a *p*type to indicate that it is a type whose main constructor is a predicate. What would an object belonging to such a type be? According to the type-theoretic approach introduced by Martin-Löf it should be an object which constitutes a proof that a is hugging b . For Martin-Löf, who was considering mathematical predicates, such proof objects might be numbers with certain properties, ordered pairs and so on. Ranta (1994) points out that for non-mathematical predicates the objects could be events as conceived by Davidson (1967, 1980). Thus $\text{hug}(a,b)$ can be considered to be an event or a situation type. In some versions of situation theory Barwise (1989); Seligman and Moss (1997), objects (called *infons*) constructed from a relation and its arguments was considered to be one kind of situation type. Thus one view would be that ptypes are playing a similar role in type theory to the role that infons play in situation theory.

What kind of entity are predicates? The notion is made precise in Appendix A.2.1. The important thing about predicates is that they come along with an *arity*. The arity of a predicate tells you what kind of arguments the predicate takes and what order they come in. For us the arity of a predicate will be a sequence of types. The predicate ‘hug’ as discussed above we can think of as a two-place predicate both of whose arguments must be of type *Ind*, that is, an individual. Thus the arity of ‘hug’ will be $\langle \text{Ind}, \text{Ind} \rangle$. The idea is that if you combine a predicate with arguments of the appropriate types in the appropriate order indicated by the arity then you will have a type. Thus if $a : \text{Ind}$ and $b : \text{Ind}$ then $\text{hug}(a,b)$ will be a type, intuitively the type of situation where a hugs b .

It may be desirable to allow some predicates to combine with more than one assortment of argument types. Thus, for example, one might wish to say that the predicate ‘believe’ can combine with two individuals just like ‘hug’ (as in *Kim believes Sam*) or with an individual and a “proposition” (as in *Kim believes that Sam is telling the truth*). Similarly the predicate ‘want’ might be both a two-place predicate for individuals (as in *Kim wants the tree*) or a two-place predicate

between individuals and “properties” (as in *Kim wants to own the tree*). We shall have more to say about “propositions” and “properties” later. For now, we just note that we want to allow for the possibilities that predicates can be *polymorphic* in the sense that there may be more than one sequence of types which characterize the arguments they are allowed to combine with. The sequences need not even be of the same length (consider *Kim walked* and *Kim walked the dog*). We thus allow for the possibility that these pairs of natural language examples can be treated using the same polymorphic predicate. Another possibility, of course, is to say that the English verbs can correspond to different (though related) predicates in the example pairs and not allow this kind of predicate polymorphism in the type theory. We do not take a stand on this issue but merely note that both possibilities are available. If predicates are to be considered polymorphic then the arity of a predicate can be considered to be a set of sequences of types.

Predicates can be considered as functions from sequences of objects matching their arity to types. As such they would be a *dependent type*, that is, an entity which returns a type when provided with an appropriate object or sequence of objects. However, we have not made this explicit in Appendix A.2.1.

A system of complex types (made precise in Appendix A.2.2) adds to a system of basic types a collection of types constructed from a set of predicates with their arities, that is, it adds all the types which you can construct from the predicates by combining them with objects of the types corresponding to their arities according to the types in the rest of the system. The system also assigns a set of objects to all the types thus constructed from predicates. Many of these types will be assigned the empty set. Intuitively, if we have a type $\text{hug}(c,d)$ and there are no situations in which c hugs d then there will be nothing in the extension of $\text{hug}(c,d)$, that is, it will be assigned the empty set in the system of complex types. Notice that the intensionality of our type system becomes very important here. There may be many individuals x and y for which $\text{hug}(x,y)$ is empty but still we would want to say that the types resulting from the combination of ‘hug’ with the various different individuals corresponds to different types of situations. There are thus two important functions in a system of complex types: one, which we call A , which comes from the system of basic types embedded in the system and assigns extensions to basic types and the other, which we call F , which assigns extensions to types constructed from predicates and arguments corresponding to the arity of the predicates. We have chosen the letters A and F because they are used very often in the characterization of models of first order logic. A model for first order logic is often characterized as a pair $\langle A, F \rangle$ where A is the domain and F a function which assigns denotations to the basic expressions (constants and predicates) of the logic. In a slight variation on classical first order logic A may be a sorted domain, that is the domain is not a single set but a set divided into various subsets, corresponding to *sorts*. For us, A characterizes assignments to basic types and thus provides something like a sorted domain in first order model theory. In first order logic F gives us what we need to know to determine the truth of expressions like $\text{hug}(a,b)$ in first order logic. Thus F will assign to the predicate ‘hug’ a set of ordered pairs telling us who hugs whom. Our F also give us the information we need in order to tell who stands in a predicate relation. However, it does this, not by assigning a set of ordered n -tuples to each predicate, but by assigning sets of witnesses (or “proofs”) to each type constructed from

a predicate with appropriate arguments. The set of ordered pairs assigned to ‘hug’ by the first order logic F corresponds to the set of pairs of arguments $\langle x, y \rangle$ for which the F in a complex system of types assigns a non-empty set. For this reason we call the pair $\langle A, F \rangle$ a *model* within the type system, even though it is not technically a model in the sense of model theory for logic. The correspondence will become important below, however.

Kim sees this situation where a (the boy) hugs b (the dog) and perceives it to be of type $\text{hug}(a,b)$. However, there are intuitively other types which she could assign to this situation other than the type of situation where a hugs b which is represented here. For example, a more general type, which would be useful in characterizing all situations where hugging is going on between any individuals, is that of “situation where one individual hugs another individual”. Another type of situation she might use is that of “situation where a boy hugs a dog”. This is a more specific type than “situation where one individual hugs another individual” but still does not tie us down to the specific individuals a and b as $\text{hug}(a,b)$ does.

There are at least two different ways in type theory to approach these more general types. One is to use Σ -types such as (1).

- (1) a. $\Sigma x:Ind. \Sigma y:Ind. \text{hug}(x,y)$
 b. $\Sigma x:Boy. \Sigma y:Dog. \text{hug}(x,y)$

In general $\Sigma x:T_1. T_2(x)$ will have as witnesses any ordered pair the first member of which is a witness for T_1 and the second member of which is a witness for $T_2(x)$. Thus this type will be non-empty (“true”) just in case there is something a of type T_1 such that there is something of type $T_2(a)$. This means that Σ -types correspond to existential quantification. A witness for (1a) would be $\langle a, \langle b, s \rangle \rangle$ where $a:Ind$, $b:Ind$ and $s:\text{hug}(a,b)$. If there is such a witness then some individual hugs another individual and conversely if some individual hugs another individual there will be a witness for this type. Σ -types are exploited for the semantics of natural language by Ranta (1994) among others.

Another approach to these more general types is to use *record types* such as (2).

- (2) a. $\left[\begin{array}{ll} x & : \text{Ind} \\ y & : \text{Ind} \\ c & : \text{hug}(x,y) \end{array} \right]$
 b. $\left[\begin{array}{ll} x & : \text{Boy} \\ y & : \text{Dog} \\ c & : \text{hug}(x,y) \end{array} \right]$

We make the notion of record type precise in Appendix A.11. Record types consist of sets of fields such as $[x:Ind]$ and $[c:hug(x,y)]$. Fields themselves are pairs consisting of a *label* such as ‘x’ or ‘c’ in the first position (before the ‘:’ in our notation) and a type in the second position. You cannot have more than one field with the same label in a record type. The witnesses of record types are *records*. These are also sets of fields, but in this case the fields consist of a label and an object belonging to a type. A record, r , belongs to a record type, T , just in case r contains fields with the same labels as those in T and the objects in the fields in r are of the type with the corresponding label in T . The record may contain additional fields with labels not mentioned in the record type with the restriction there can only be one field within the record with a particular label. Thus both (3a) and (3b) are records of type (2a).

$$(3) \quad \begin{array}{ll} \text{a.} & \left[\begin{array}{lcl} x & = & a \\ y & = & b \\ c & = & s \end{array} \right] \quad \text{where } a:Ind, b:Ind \text{ and } s:hug(a,b) \\ \\ \text{b.} & \left[\begin{array}{lcl} x & = & d \\ y & = & e \\ c & = & s' \\ z & = & f \\ w & = & g \end{array} \right] \quad \text{where } d:Ind, e:Ind, s':hug(d,e) \text{ and} \\ & \quad f \text{ and } g \text{ are objects of some type} \end{array}$$

Note that in our notation for records we have ‘=’ between the two elements of the field whereas in record types we have ‘:’. Note also that when we have types constructed from predicates in our record types and the arguments are represented as labels as in (2a) this means that the type is *dependent* on what objects you choose for those labels in the object of the record type. Thus in (3a) the type of the object labelled ‘c’ is $hug(a,b)$ whereas in (3b) the type is $hug(d,e)$. Actually, the notation we are using here for the dependent types is a convenient simplification of what is needed as we explain in Appendix A.11.

Record types and Σ -types are very similar in an important respect. The type (2a) will be non-empty (“true”) just in case there are individuals x and y such that x hugs y . Thus both record types and Σ -types can be used to model existential quantification. In fact record types and Σ -types are so similar that you would probably not want to have both kinds of types in a single system and we will not use Σ -types. We have chosen to use record types for a number of reasons:

fields are unordered The Σ -types in (4) are distinct, although there is an obvious equivalence which holds between them.

- (4) a. $\Sigma x:Ind.\Sigma y:Ind.hug(x,y)$
 b. $\Sigma y:Ind.\Sigma x:Ind.hug(x,y)$

They are not only distinct types but they also have distinct sets of witnesses. The object $\langle a, \langle b, s \rangle \rangle$ will be of type (4a) just in case $\langle b, \langle a, s \rangle \rangle$ is of type (4b). In contrast, since we are regarding record types (and records) as *sets* of fields, (5a,b) are variant notations for the same type.

- (5) a. $\left[\begin{array}{lcl} x & : & Ind \\ y & : & Ind \\ c & : & hug(x,y) \end{array} \right]$
 b. $\left[\begin{array}{lcl} y & : & Ind \\ x & : & Ind \\ c & : & hug(x,y) \end{array} \right]$

labels Record types (and their witnesses) include labelled fields which can be used to access “components” of what is being modelled. This is useful, for example, when we want to analyze anaphoric phenomena in language where pronouns and other words refer back to parts of previous meanings in the discourse. They can also be exploited in other cases where we want to refer to “components” of utterances or their meanings as in clarification questions.

discourse representation The labels in record types can play the role of discourse referents in discourse representation structures (DRSs, Kamp and Reyle, 1993) and record types of the kind we are proposing can be used to model DRSs.

dialogue game boards Record types have been exploited to model dialogue game boards or information states (see in particular Ginzburg, 2012).

feature structures Record types can be used to model the kind of feature structures that linguists like to use (as, for example, in linguistic theories like Head Driven Phrase Structure Grammar, HPSG, Sag *et al.*, 2003). Here the labels in record types correspond to attributes in feature structures.

frames Record types can also be used to model something very like the kinds of frames discussed in frame semantics (Fillmore, 1982, 1985; Ruppenhofer *et al.*, 2006). Here the labels in record types correspond to roles (frame elements).

For discussion of some of the various uses to which record types can be put see Cooper (2005). We will take up all of the uses named here as we progress.

Another way of approaching these more general types in type theory is to use *contexts*. In (6) we take *True* to be the type of non-empty types.

- (6) a. $x : Ind, y : Ind \vdash \text{hug}(x,y) : True$
 b. $x : Boy, y : Dog \vdash \text{hug}(x,y) : True$

(6a,b) mean in a context where x and y are individuals or a boy and a dog respectively the type $\text{hug}(x,y)$ is non-empty. This notation is normally taken to mean universal quantification over the parameters or variables in the context (i.e. sequence of parametric type judgements) to the left of ‘ \vdash ’. Thus they would mean that for any two individuals or pair of a boy and a dog, the first hugs the second. However, we can also devise ways for thinking of existential quantification over the variables of the context, e.g. for some boy, x , and some dog, y , the type $\text{hug}(x,y)$ is non-empty. We can also think of the contexts as being objects belonging to types in our type theory. Records and record types give us a way of doing this. Thus, for example, (2a) models the type of context which might be represented as the sequence of parametric type judgements given in (7).

- (7) $x : Ind, y : Ind, c : \text{hug}(x,y)$

As in the comparison with Σ -types there is a difference in that the judgements in a standard type theory context are ordered whereas the fields in a record type are unordered. This means that technically (8) is a distinct context from (7) even though there is an obvious equivalence between them.

- (8) $y : Ind, x : Ind, c : \text{hug}(x,y)$

They correspond to the same record type, however. Since we will use record types to model type theoretic contexts and records to model instantiations of contexts we will not introduce a separate notion of context.

Thus we use record types to replace both the Σ -types and contexts that one often finds in standard versions of type theory.

The introduction of predicates and ptypes raises some new questions. We said above that the arity of ‘hug’ is $\langle Ind, Ind \rangle$. However, when we look at (2b) where the types labelled with ‘x’ and ‘y’ are *Boy* and *Dog* we see that there is nothing explicit here that requires that the two arguments of ‘hug’ are of type *Ind*. One obvious way to achieve this would be to require that *Boy* and *Dog* are subtypes of *Ind*, that is, that any object of type *Boy* is also of type *Ind* and similarly for *Dog*. However, now that we have introduced predicates there is nothing to stop us having two predicates ‘boy’ and ‘dog’ with arity $\langle Ind \rangle$. Thus we could have the record type (9).

$$(9) \quad \left[\begin{array}{ll} x & : \quad Ind \\ c_{boy} & : \quad boy(x) \\ y & : \quad Ind \\ c_{dog} & : \quad dog(y) \\ c_{hug} & : \quad hug(x,y) \end{array} \right]$$

How do we choose between a type like (9) where common nouns like *boy* and *dog* correspond to one-place predicates and a type like (2b) where common nouns correspond to basic types? One advantage is that (9) explicitly represents that the arity of ‘hug’ is fulfilled. Another advantage is that many, if not all, nouns in natural languages will in more detailed analyses correspond to predicates of more than one argument. Consider, for example, the fact that boys grow into men. The same individual can be a boy at one time and a man at a later time. Thus ‘boy’ should be a predicate of at least two arguments with arity $\langle Ind, Time \rangle$. In fact if we are going to deal with tense and aspect in natural language we will probably want to add time arguments to most if not all of our predicates and thus allow ourselves record types like (10).

$$(10) \quad \left[\begin{array}{ll} e\text{-time} & : \quad Time \\ x & : \quad Ind \\ c_{boy} & : \quad boy(x, e\text{-time}) \\ y & : \quad Ind \\ c_{dog} & : \quad dog(y, e\text{-time}) \\ c_{hug} & : \quad hug(x, y, e\text{-time}) \end{array} \right]$$

where ‘e-time’ stands for “event time”. Here we have required that the times in all the predicate fields be the event time but this is not always the case. Consider (11).

(11) The minister smoked pot in his youth

Here the time of the pot-smoking event most likely precedes the time of the pot-smoking individual being a minister. We will thus use our basic types for basic ontological categories like

individual and *time* and use predicates for words that occur in natural language. Predicates can be n -ary whereas our types will always be unary. Note that a ptype like $\text{hug}(a,b,t)$ is constructed from a ternary predicate ‘hug’ but the type itself is a unary type of situations. Thus we might have the judgement $s : \text{hug}(a,b,t)$.

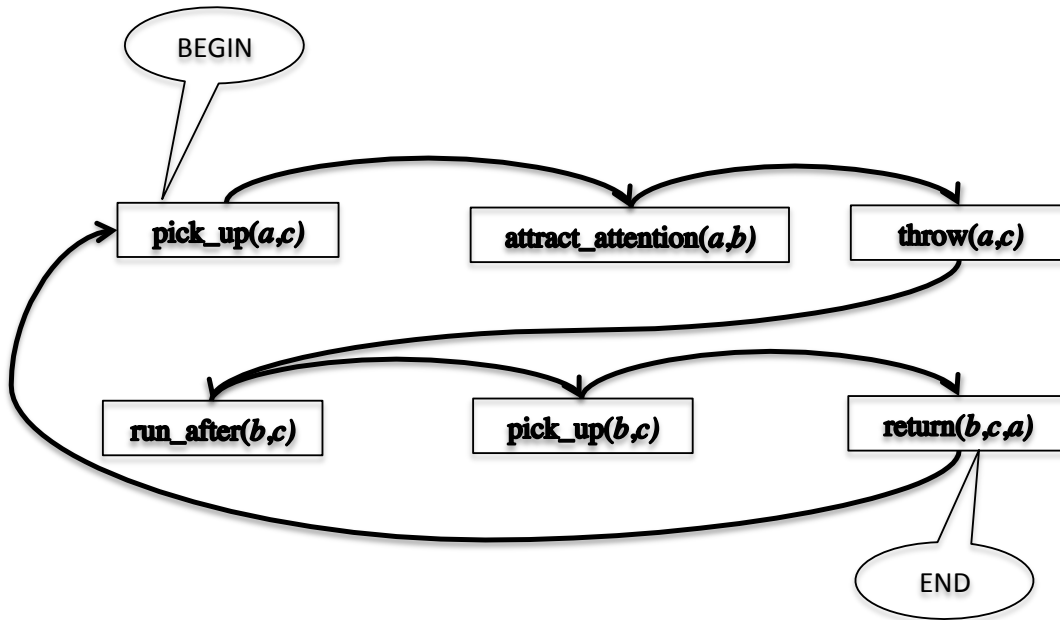
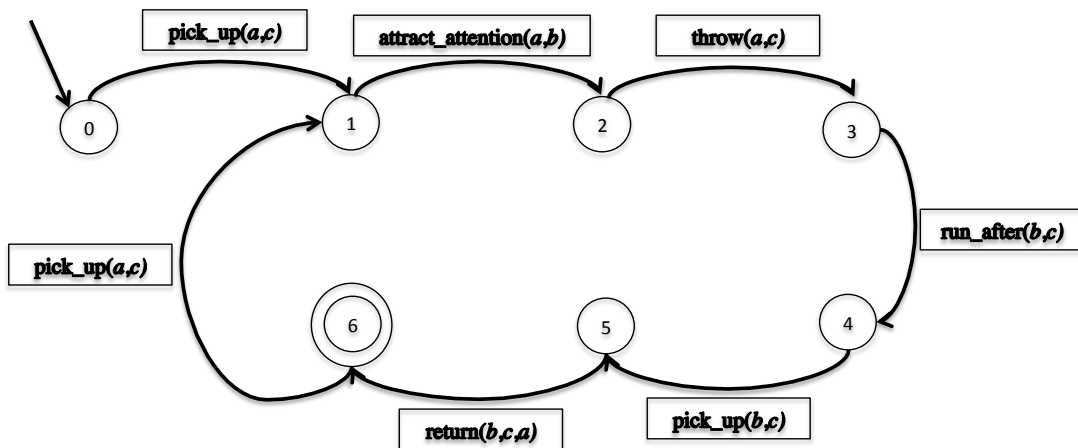
1.4 The string theory of events

Kim stands and watches the boy and the dog for a while. They start to play fetch.¹ This is a moderately complex game in that it consists of a number of components which are carried out in a certain order. The boy picks up a stick, attracts the attention of the dog (possibly shouting “Fetch!”), and throws the stick. The dog runs after the stick, picks it up in his mouth and brings it back to the boy. This sequence can be repeated arbitrarily many times. One thing that becomes clear from this is that events do not happen in a single moment but rather they are stretched out over intervals of time. So if we were to have a type of event (that is, a kind of situation) $\text{play_fetch}(a,b,c,t)$ where a is a human, b is a dog and c is a stick, then t should not be a moment of time but a time *interval* starting with the time of the beginning of the event and ending with the time of the end of the event. But we could also say something about the series of subevents that we have identified. So we might draw an informal diagram something like Figure 1.2 (ignoring time for the moment).

In an important series of papers including Fernando (2004, 2006, 2008, 2009), Fernando introduces a finite state approach to event analysis where events are analyzed in terms of finite state automata something like what we have represented in Figure 1.3. Such an automaton will recognize a string of sub-events. The idea is that our perception of complex events can be seen as strings of punctual observations similar to the kind of sampling we are familiar with from audio technology and digitization processing in speech recognition. Thus events can be analyzed as strings of smaller events. What we mean by a string is made precise in Appendix A.9. Any object of any type can be part of a string. Any two objects (including strings themselves), s_1 and s_2 , can be *concatenated* to form a string $s_1 \frown s_2$. An important property of concatenation is *associativity*, that is if we concatenate s_1 with s_2 and then concatenate the result with s_3 we get the same string that we would obtain by concatenating s_2 with s_3 and then concatenating s_1 with the result. In symbols: $(s_1 \frown s_2) \frown s_3 = s_1 \frown (s_2 \frown s_3)$. For this reason we normally write $s_1 \frown s_2 \frown s_3$ (without the parentheses) or simply $s_1 s_2 s_3$ if it is clear from the context that we mean this to be string concatenation.

Now let us build further on the types that we have introduced so far to include string types. For any two types, T_1 and T_2 , we can form the type $T_1 \frown T_2$. This is the type of strings $a \frown b$ where $a : T_1$ and $b : T_2$. The concatenation operation on types (just like that on objects) is associative

¹[http://en.wikipedia.org/wiki/Fetch_\(game\)](http://en.wikipedia.org/wiki/Fetch_(game)), accessed 10th Oct 2011.

Figure 1.2: `play_fetch(a,b,c)`Figure 1.3: `play_fetch(a,b,c)` as a finite state machine

so we do not use parentheses when more than one type is involved, e.g. $T_1 \frown T_2 \frown T_3$.

Let us return to Kim watching the boy, a , playing fetch with the dog, b , using the stick, c over a time interval t . She perceives the event as being of type $\text{play_fetch}(a,b,c,t)$. But what does it mean to be an event of this type? Given our concatenation types we can build a type which corresponds to most of what we have sketched in Figure 1.2, namely (12) (again ignoring time for the sake of simplification).

$$(12) \quad \text{pick_up}(a,c) \frown \text{attract_attention}(a,b) \frown \text{throw}(a,c) \frown \text{run_after}(b,c) \frown \text{pick_up}(b,c) \frown \text{return}(b,c,a)$$

(12) is a type corresponding to everything we have represented in Figure 1.2 except for the arrow which loops back from the end state to the start state. We will refer to this type as $T_{\text{fetch}(a,b,c)}$. In order to get the loop into the event type we will use a kind of type which introduces a Kleene-+. In standard notations for strings s^+ stands for a string consisting of one or more occurrences of s .² We will adopt this into types by saying that for any type T there is also a type T^+ which is the type of strings of objects of type T containing one or more members. (See Appendix A.9 for a more precise definition.) The type $T_{\text{fetch}(a,b,c)}^+$ will, then, give us a type corresponding to the complete Figure 1.2 since it will be the type consisting of strings of one or more events of type $T_{\text{fetch}(a,b,c)}$.

In (12) we simplified the string type by excluding time. Now we will consider what is needed to put time back in. Each of the subevents represented there is associated with a time interval, a different time interval for each subevent. The type for time intervals, which we will abbreviate as *TimeInt* is (13).

$$(13) \quad \left[\begin{array}{ll} \text{start} & : \text{Time} \\ \text{end} & : \text{Time} \\ \text{c}_{<} & : \text{start} < \text{end} \end{array} \right]$$

where *Time* is the type of time points and $<$ is the relation “earlier than” defined on the witnesses of this type.³ In order to get time intervals associated with each subevent we will treat the subevent types as record types rather than ptypes. Thus as a first step towards characterizing the right type we have (14).

²This notation was introduced by the mathematician Stephen Kleene.

³We use the infix notation $t_1 < t_2$ rather than the official notation $<(t_1, t_2)$.

$$\begin{aligned}
(14) \quad & \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{pick_up}}:\text{pick_up}(a,c,\text{e-time}) \end{array} \right] \frown \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{attract_attention}}:\text{attract_attention}(a,b,\text{e-time}) \end{array} \right] \frown \\
& \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{throw}}:\text{throw}(a,c,\text{e-time}) \end{array} \right] \frown \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{run_after}}:\text{run_after}(b,c,\text{e-time}) \end{array} \right] \frown \\
& \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{pick_up}}:\text{pick_up}(b,c,\text{e-time}) \end{array} \right] \frown \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{return}}:\text{return}(b,c,a,\text{e-time}) \end{array} \right]
\end{aligned}$$

This will get us a time interval for each subevent but it does not require any relationship between the time intervals of the different subevents whereas, of course, each subevent should occur at a later time than the previous one. In order to achieve this we will introduce a more restricted notion of type concatenation especially for record types which have a field [e-time:TimeInt] as in (15).

$$\begin{aligned}
(15) \quad & 1. \text{ If } T_1 \text{ and } T_2 \text{ are subtypes of } \text{[e-time:TimeInt]}^+ \text{ then} \\
& \quad T_1 \frown^< T_2 \text{ is a type.} \\
& 2. a : T_1 \frown^< T_2 \text{ iff } a = a_1 \frown a_2, a_1 : T_1, a_2 : T_2 \text{ and} \\
& \quad \text{last}(a_1).\text{e-time.end} < \text{first}(a_2).\text{e-time.start}
\end{aligned}$$

Here $\text{first}(s)$ and $\text{last}(s)$ pick out the first and last elements respectively of the string s . We need to make a similar adjustment to the Kleene-+ type constructor.

$$\begin{aligned}
(16) \quad & 1. \text{ If } T \text{ is a subtype of } \text{[e-time:TimeInt]} \text{ then } T^{+<} \text{ is a} \\
& \quad \text{type} \\
& 2. a : T^{+<} \text{ iff } a = x_1 \frown \dots \frown x_n, n > 0 \text{ and for } i, j, 1 \leq \\
& \quad i < j \leq n, x_i : T, x_j : T \text{ and } x_i \frown x_j : T \frown^< T
\end{aligned}$$

[The details of temporal concatenation above should be included in the appendix.]

So now we can represent the event type for playing fetch as (17).

$$\begin{aligned}
(17) \quad & \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{pick_up}}:\text{pick_up}(a,c,\text{e-time}) \end{array} \right] \frown_{<} \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{attract_attention}}:\text{attract_attention}(a,b,\text{e-time}) \end{array} \right] \frown_{<} \\
& \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{throw}}:\text{throw}(a,c,\text{e-time}) \end{array} \right] \frown_{<} \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{run_after}}:\text{run_after}(b,c,\text{e-time}) \end{array} \right] \frown_{<} \\
& \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{pick_up}}:\text{pick_up}(b,c,\text{e-time}) \end{array} \right] \frown_{<} \left[\begin{array}{l} \text{e-time:TimeInt} \\ \text{c}_{\text{return}}:\text{return}(b,c,a,\text{e-time}) \end{array} \right] \frown_{<}^{+}
\end{aligned}$$

What happens when Kim perceives an event as being of this type? She makes a series of observations of events, assigning them to types in the string type. Note that the ptypes in each of the types can be further broken down in a similar way. This gives us a whole hierarchy of perceived events which at some point have to bottom out in basic perceptions which are not further analyzed. In order to recognize an event as being of this type Kim does not need to perceive a string of events corresponding to each of the types in the string types. She may, for example, observe the boy waving the stick to attract the dog's attention, get distracted by a bird flying overhead for a while, and then return to the fetch event at the point where the dog is running back to the boy with the stick. This still enables her to perceive the event as an event of fetch playing because she has seen such events before and learned that such events are of the string type in (17). It suffices for her to observe enough of the elements in the string to distinguish the event from other event types she may have available in her knowledge resources. Suppose, for example, that she has just two event string types available that begin with the picking up of a stick by a human in the company of a dog. One is (17). The other is one that leads to the human beating the dog with the stick. If she only observes the picking up of the stick she cannot be sure whether what she is observing is a game of fetch or a beating. However, as soon as she observes something in the event string which belongs only to the fetch type of string she can reasonably conclude that she is observing an event of the fetch type. She may, of course, be wrong. She may be observing an event of a type which she does not yet have available in her resource of event types, in which case she will need to learn about the new event type and add it to her resources. However, given the resources at her disposal she can make a prediction about the nature of the rest of the event. One could model her prediction making ability in terms of a function which maps a situation (modelled as a record) to a type of predicted situation, for example (18).

$$(18) \quad \lambda r: \left[\begin{array}{l} x:Ind \\ c_{human}:human(x) \\ y:Ind \\ c_{dog}:dog(y) \\ z:Ind \\ c_{stick}:stick(z) \\ e-time:TimeInt \\ e: \left[\begin{array}{l} e-time:TimeInt \\ c_{\leq}:\uparrow e-time.start \leq e-time.start \\ c_{pick_up}:pick_up(\uparrow x, \uparrow z, e-time) \end{array} \right] \end{array} \right] \frown < \left[\begin{array}{l} e-time:TimeInt \\ c_{\leq}:e-time.start \leq \uparrow e-time.start \\ c_{attract_attention}:attract_attention(\uparrow x, \uparrow y, e-time) \end{array} \right] \left(\begin{array}{l} e-time:TimeInt \\ c_{\leq}:r.e-time.start \leq e-time.start \\ e:play_fetch(r.x, r.y, r.z, e-time) \end{array} \right)$$

Here a ‘ \uparrow ’ prefixed to a path-name consisting of labels such as ‘e-time.start’ or ‘x’ means that the path-name is to start in the next higher record type in which the current record type is embedded. This notation is an abbreviation for something much less readable (but more precise) which is given in Appendix A.11. [???? This notation needs to be added to the appendix.]

The kind of function of which (18) is an instance is a function of the general form (19).

$$(19) \quad \lambda a : T_1(T_2[a])$$

where we use the notation $T_2[a]$ to represent the fact that T_2 depends on a . The nature of this dependence in (18) is seen in the occurrences of r in the body of the function, for example, ‘play_fetch($r.x, r.y, r.z, e-time$)’. Such a function maps an object of some type (represented by T_1) to a type (represented by $T_2[a]$). The type that results from an application of this function will depend on what object it is applied to – that is, we have the possibility of obtaining different types from different objects. In type theory such a function is often called a *dependent type*. These functions will play an important role in much of what is to come later in this book. They will show up many times in what appear at first blush to be totally unrelated phenomena. We want to suggest, however, that all of the phenomena we will describe using such functions have their origin in our basic cognitive ability to make predictions on the basis of partial observation of objects and events.

What happens when Kim does not observe enough of the event to be able to predict with any certainty that the complete event will be a game of fetch? One theory would be that she can only make categorical judgements, and that she has to wait until she has seen enough so that there is only one type that matches in the collection of situation types in her resources. Another theory would be one where she predicts a disjunction of the available matching types when there

is more than one that matches. One might refine this theory so that she can choose one of the available types but assign it a probability based on the number of matching types. If n is the number of matching types the probability of any one of them might be $\frac{1}{n}$. This assumes that each of the types is equally likely to be realized. It would be natural to assume, however, that the probability which Kim assigns to any one of the matching types would be dependent on her previous experience. Suppose, for example, that she has seen 100 events of a boy picking up a stick in the company of a dog, 99 of those events led to a game of fetch and only one led to the boy beating the dog. One might then assume that when she now sees the boy pick up the stick she would assign a 99% probability to the type of fetch events and only 1% probability to the boy beating the dog. That is, the probability she assigns to an event of a boy picking up a stick leading to a game of fetch is the result of dividing the number of instances of a game of fetch she has already observed by the sum of the number of instances she has observed of any types whose initial segment involves the picking up of a stick. In more general terms we can compute the probability which an agent A assigns at time t to a predicted type T_{pr} given an observed type T_{obs} , $P_{A,t}(T_{pr} \mid T_{obs})$, in the case where T_{pr} is a member of the set of alternatives which can be predicted from T_{obs} according to A 's resources at t , $\text{alt}_{A,t}(T_{obs})$, by the following formula:

$$P_{A,t}(T_{pr} \mid T_{obs}) = \frac{|\{T_{pr}\}^{A,t}|}{\sum_{T_{alt} \in \text{alt}_{A,t}(T_{obs})} |\{T_{alt}\}^{A,t}|}$$

where $\{T\}^{A,t}$ is the set of objects of type T observed by A prior to t . If T_{pr} is not a member of $\text{alt}_{A,t}(T_{obs})$, that is not one of the alternatives, we say that $P_{A,t}(T_{pr} \mid T_{obs}) = 0$.

While this is still a rather naive and simple view of how probabilities might be assigned it is not without interest, as shown by the following points:

Probability distributions It will always provide a probability distribution over sets of alternatives, that is,

$$\sum_{T_{pr} \in \text{alt}_{A,t}(T_{obs})} P_{A,t}(T_{pr} \mid T_{obs}) = 1$$

Alternatives We have assumed a notion of alternatives based on types of completed events for which the observed event is an initial segment but other notions of alternativeness could be considered and perhaps even combined.

Relativity of probability assignments The notion of probability is both agent and time relative. It represents the probability which an agent will assign to a type when observing a given situation at a given time. Two agents may assign different probabilities depending on the resources they have available and their previous experiences using those resources.

Learning Relevant observations will update the probability distributions an agent will assign to a given set of alternatives since the probability is computed on the basis of previous observations of the alternative types.

Kim is not alone in being able to draw conclusions based on partial observations of an event. The dog can do it too. As soon as the boy has raised the stick and attracted the dog's attention the dog is excitedly snapping at the stick and starting to run in the direction in which the boy seems to be about to throw. The dog also seems to be attuned to string types of events just as Kim is and also able to make predictions on the basis of partial observations. The types to which a dog is attuned will not be the same as those to which humans can be attuned and this can certainly lead to miscommunication between humans and dogs. For example, there may be many reasons why I would go to the place where outdoor clothes are hanging and where the dog's lead is kept. Many times it will be because I am planning to take the dog out for a walk, but not as often as the dog appears to think, judging from the excitement he shows any time I go near the lead. It is difficult to explain to the dog that I am just looking for a receipt that I think I might have left in my coat pocket. But the basic mechanism of being able to assemble types of events into string types of more complex events and make predictions on the basis of these types seems to be common to both humans and dogs and a good number of other animals too. Perhaps simple organisms do not have this ability and can only react to events that have already happened, but not to predicted outcomes.

This basic inferential ability is thus not parasitic on the ability to communicate using a human language. It is, however, an ability which appears to be exploited to a great extent in our use of language as we will see in later chapters. In the remaining sections of this chapter we will look at some aspects of the type theory which seem more likely to correspond to cognitive abilities which only humans have.

1.5 Doing things with types

The boy and the dog have to coordinate and interact in order create an event of the game of fetch.
@@

1.6 Modal type systems

Kim continues her walk still thinking about the boy and the dog. She thinks, "Was the boy standing too close to the pond? Suppose he had fallen in. If he had been my son, I wouldn't have let him play just there." An important aspect of human cognition is that we are not only able

observe things as they are but also to conceive of alternatives which go beyond the completion of observed events in the way discussed in Section 1.4. We can not only observe objects and perceive them to be of certain types we can also consider possibilities in which they belong to different types and perhaps do not belong to the type we have observed. We have managed to unhook type judgements from direct perception. While the seeds of this ability can be seen in the kind of event perception and prediction discussed above in that it gives us a way to consider types which have not yet been realized, it is at least one step further in cognitive evolution to be able to consider alternative type assignments which do not correspond to completions of events already perceived.

This leads us to construct *modal type systems* with alternative assignments of objects to types.⁴ Figure 1.4 provides an example of a modal system of basic types with two possibilities, one where the extensions of types T_1 and T_2 overlap and another possibility where they do not.

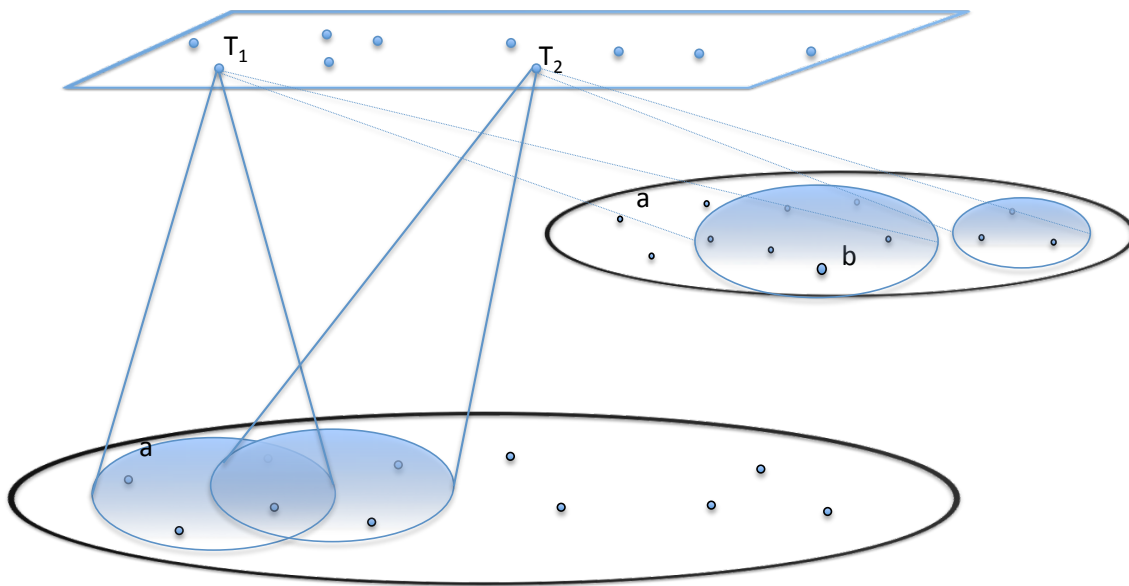


Figure 1.4: Modal system of basic types

The object a is of type T_1 in the first possibility but not in the second possibility. There is an object, b , of type T_1 in the second possibility. b does not exist at all in the first possibility. In the figure we just show two possibilities but our general definition in Appendix A.1 allows for there to be any number of possibilities, including infinitely many.

Given this apparatus we define four simple modal notions:

⁴The term *modal* is taken from modal logic. See Hughes and Cresswell (1968) for a classic introduction. A modern introduction is to be found in Blackburn *et al.* (2001).

(necessary) equivalence Two types are (necessarily) equivalent just in case the extension of one type is identical with that of the other type in all the possibilities. While the different possibilities may provide different extensions for the types, it will always be the case that in any given possibility the two types will have the same extension.

subtype One type is a subtype of another just in case whatever possibility you look at it is always the case that the extension of the first type is a subset of the extension of the second. We can also say that the first type “entails” the second, that is, any object which is of the first type will also be of the second type, no matter which possibility you are considering.

necessity The notion of necessity we characterize for a type could be glossed as “necessarily realized” or “necessarily instantiated”. A type will be necessary just in case there is something of the type in all the possibilities.

possibility This notion corresponds to “possibly realized” or “possibly instantiated”. A type will be possible just in case there is some possibility according to which it has a non-null extension.

These notions are made precise in Appendix A.1. Note that all of these notions are relativized to the modal system you are considering and the possibilities it offers. We may think of the family of assignments \mathcal{A} as providing a modal base (cf. Kratzer) or alternatives (in the sense of ???). For these kinds of applications we may wish to consider very small families of assignments corresponding to the knowledge we have. Alternatively, we may want to consider strong logical variants of these modal notions where we consider all the logical possibilities, for example, all possible assignments of extensions to types.

So far we have talked about modal systems of basic types. Modal systems of complex types, where we introduce ptypes, create a minor complication. What ptypes that are present in a system depends on what objects there are of the types that are used in the arities of the predicates. Thus if we have some predicate r with arity $\langle Ind, Ind \rangle$ and a possibility where the set assigned to Ind is $\{a, b\}$ then according to that possibility the ptypes formed with r will be $r(a, a)$, $r(a, b)$, $r(b, a)$ and $r(b, b)$. In a possibility where Ind is assigned a different set the set of available ptypes will be different. It is an important feature of type theories with types constructed from predicates that the collection of such types depends on what objects are available as arguments to the predicates. This makes type theory very different from a logical language such as predicate calculus where the notion of well-formedness of syntactic expressions containing predicates is defined independently of what is provided by the model as denotations of arguments to the predicate.

This leads us (in Appendix A.8) to define two variants of each of our modal notions: *restrictive* variants which are only defined for types which exist in all possibilities and *inclusive* variants which require that the modal definition holds for all the possibilities in which the types exist and disregards those in which the types do not exist. For example, a type is *necessary_r* (that is, “restrictively necessary”) just in case the type is available in all possibilities and has a non-empty

set of witnesses in all possibilities. It is *necessary_i* (“inclusively necessary”) just in case in all the possibilities in which the type is provided it has a non-empty set of witnesses. It is clear that if a type is *necessary_r* it will also be *necessary_i* but there may be types which are *necessary_i* but not *necessary_r* (if the type is not provided in all possibilities). A similar relationship between the restrictive and inclusive notions holds for all the modal notions we have discussed.

There may be significant classes of modal type systems in which the types available in the different possibilities do not vary. This could be achieved by requiring that the types used in the arities of predicates always have the same witnesses in all the possibilities. This seems feasible if we restrict the types used in predicate arities to basic ontological categories such as individual or time point. It seems reasonable to consider modal systems in which an individual in one possibility will be an individual in any other possibility, for example. It seems reasonable to say that we wish to consider possibilities where, for example, Kim is a man rather than a woman, but not possibilities where Kim is a point in time rather than an individual. However, the notion “basic ontological category” is a slippery one and we do not want to be forced to make commitments about that.

In the definition of a system of complex types in section A.2.2 we call the pair of an assignment to basic types and assignment to ptypes, $\langle A, F \rangle$, a *model* because of its similarity to first order models.⁵ The model provides an interface between the type theoretical system and a domain external to the type theory. The natural domain to relate to the type theory is that of individuals and situations, that is the kind of things we can perceive or at least consider as possibilities. However, we may want to use models which relate to our perceptual apparatus, as in Larsson (2011), rather than directly to the world. This can also be the key for relating the type theory to a dynamically changing world where the models representing our perceived possibilities are not fixed. We will return to this in Chapter 5.

1.7 Intensionality: propositions as types

Kim continues to think about the boy and the dog as she walks along. It was fun to see them playing together. They seemed so happy. The boy obviously thought that the dog was a good playmate. Kim is not only able to perceive events as being of certain types. She is able to recall and reflect on these types. She is able to form attitudes towards these types: it was fun that the boy and the dog were playing but a little worrying that they were so close to the pond. This means that the types themselves seem to be arguments to predicates like ‘fun’ and ‘worrying’. This seems to be an important human ability – not only to be able to take part in or observe an event and find it fun or worrying but to be able to reflect independently of the actual occurrence of the event that it or in general similar events are fun or worrying. This is a source of great richness

⁵For a more detailed discussion of the relationship between this and first order models as used in the interpretation of first order logic see Cooper (fthc).

in human cognition in that it enables us to consider situation types independently of their actual instantiation.⁶ This abstraction also enables us to consider what attitudes other individuals might have. For example, Kim believes that the boy thought that the dog was a good playmate. She is able to ascribe this belief to the boy. Furthermore, we are able to reflect on Kim's state of mind where she has a belief concerning the type of situation where the boy thinks that the dog was a good playmate. And somebody else could consider of us that we have a certain belief about Kim concerning her belief about the boy's belief. There is in principle no limit to the depth of recursion concerning our attitudes towards types.

We propose to capture this reflective nature of human cognition by making the type theory technically *reflective* in the sense that we allow types themselves to be objects which can belong to other types. In classical model theoretic semantics we think of *believe* as corresponding to a relation between individuals and propositions. In our type theory, however, we are subscribing to the “propositions as types” view which comes to us via Martin-Löf (1984) but has its origins in intuitionistic logic [????]. Propositions are true or false. Types of situations such as $\text{hug}(a,b)$ correspond to propositions in the sense that if they are non-empty then the proposition is true. If there is nothing of this type then it is false. The reasoning is thus that we do not need propositions in our system as separate semantic objects if we already have types. We can use the types to play the role of propositions. To believe a type is to believe it to be non-empty. From the point of view of a type theory for cognition in which we connect types to our basic perceptual ability, this provides a welcome link between our perceptual ability and our ability to entertain propositions (that is, to consider whether they are true or false).

A predicate like ‘believe’ which represents that an individual has an attitude (of belief) to a certain type should thus have an arity which requires its arguments to be an individual and a type. That is, we should be able to construct the type $\text{believe}(c, \text{hug}(a,b))$ corresponding to *c believes that a hugs b*. We thus create *intensional* type systems where types themselves can be treated as objects and belong to types. Care has to be taken in constructing such systems in order to avoid paradoxes. We use a standard technique known as stratification Turner (2005). We start with a basic type system and then add higher order levels of types. Each higher order includes the types of the order immediately below as objects. In each of these higher orders n there will be a type of all types of the order $n - 1$ but there is no ultimate “type of all types” – such a type would have to have itself as an object. This is made precise in Appendix A.10. For more detailed discussion see Cooper (fthc). Figure 1.5 represents an intensional modal type system where we indicate just the initial three orders of an infinite hierarchy of type orders.

⁶This richness also has its downside in that we often become so engaged in our internal cognitive abstraction that it can be difficult to be fully present and conscious of our direct perception of the world – for example, worrying about what might happen in the future rather than enjoying the present.

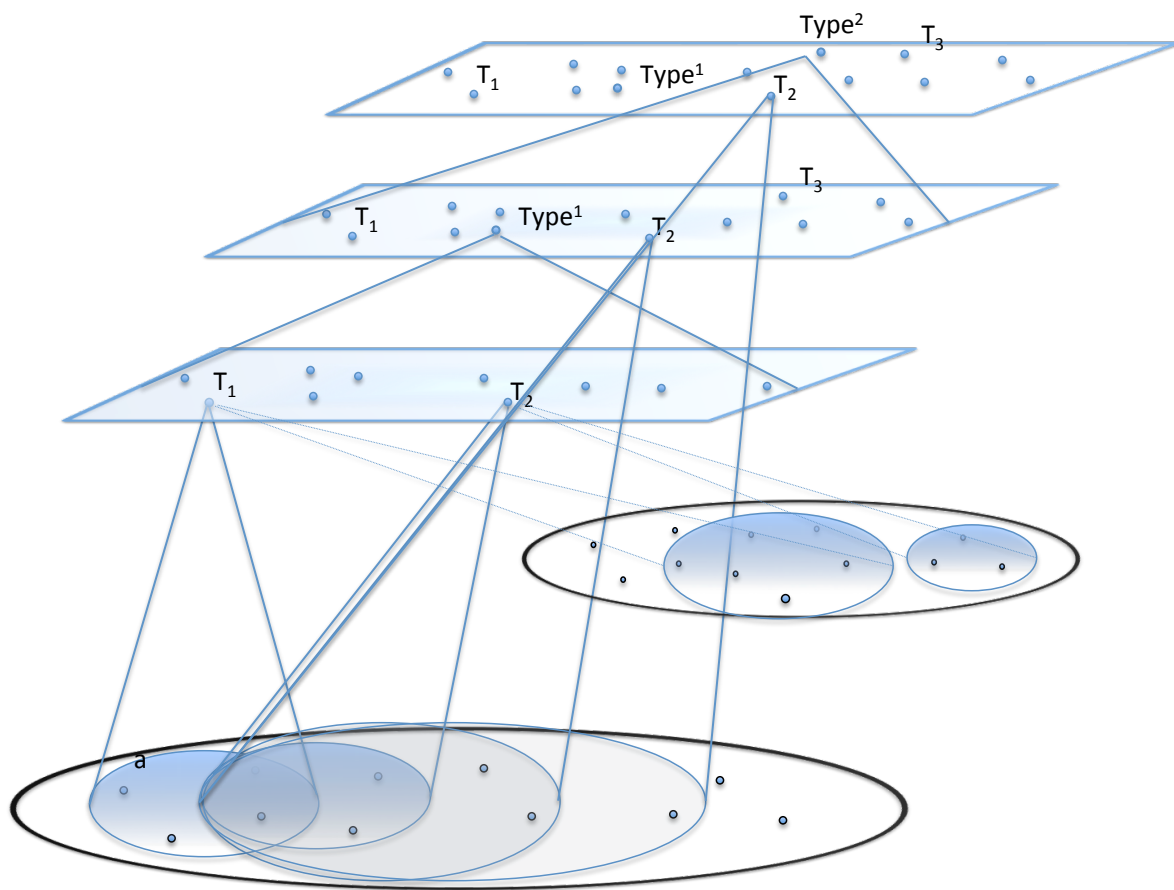


Figure 1.5: Intensional modal type system

Chapter 2

Information exchange

2.1 Speech events

In Chapter 1 we talked about the perception of events such as a boy and a dog playing fetch. We imagined Kim walking through the park and perceiving various kinds of events. Suppose that she meets a friend in the park and they start to have a conversation. A conversation is a kind of event involving language which seems to be uniquely human. The kind of dialogue involved in a conversation enables humans to exchange information in a way that is more complex and more abstracted from currently occurring events than other animals seem capable of. Nevertheless, we will argue that the basic mechanisms of dialogue involve assigning types to events in way that we discussed in Chapter 1. The events involved are *speech events*.

Consider the kind of event type prediction that we considered in Chapter 1. Suppose that Kim sees the boy playing fetch with the dog and the boy is standing close to the lake with his back to it. As the dog runs towards him with the stick he takes a step backwards. “No,” says Kim, seeing that the boy is about to fall in the lake. “Watch out,” she shouts to the boy who takes a step forward just in time and narrowly misses falling in the lake. Her utterance of *no* represents a negative attitude towards a predicted outcome. This kind of negation is discussed briefly in Cooper and Ginzburg (2011a,b) where examples are given of cases where *no* is a response to a completed event and where it is used as an attempt to prevent the predicted outcome. This latter exploits the fact that agents cannot only perceive and classify events according to the types to which they are attuned but can also intervene and prevent a predicted outcome. Kim’s linguistic utterance of *watch out* is used in this way. While Kim is using words of English this is not yet completely linguistic interaction. A dog, sensing danger, will begin to bark and this can have the effect of preventing a predicted outcome. It is a kind of inter-agent communication nevertheless in that it is an intervention in the flow of events which involves predicting and changing the behaviour of

- (20) John: Hello doctor.
 Anon 1: Hello.
 Well Mr [last or full name], what can I [do for you today]₁?
 John: [Er, it's]₁ a wee problem I've had for a ⟨pause⟩ say about a year now.
 Anon 1: Mhm.
 John: It's er my face.
 And my skin.
 I seem to get an awful lot of, it's like
 Anon 1: Aha.
 John: dry flaky skin.
 Anon 1: Yeah.
 John: And I get it on my forehead, [down here]₂
 Anon 1: [I can see]₂

BNC file G43, sentences 1–13

another agent. In this sense it is similar to human dialogue, although human dialogue is normally a much more abstract affair, involving predicting and influencing the other agent's linguistic behaviour and the attitudes and beliefs which the other agent has concerning certain types.

Dialogues themselves are events and, just like other events, can be regarded as strings of smaller events. Consider the dialogue excerpt (20) from the British National Corpus which is the beginning of a consultation between a patient (John) and a doctor (Anon 1).

We might assign the whole dialogue of which this is a part to a genre type for patient doctor consultation.¹ The genre type could be seen as an event type which, like the type for the game of fetch discussed in Chapter 1, can be broken down into a string of subevent types such as greeting (realized here by the exchange *Hello doctor/Hello*), establishing the patient's symptoms (realized here by the remainder of (20)), making a diagnosis, prescribing treatment and so on. Events belonging to these subtypes can be further broken down into strings of turns which further can be broken down into strings of utterances of phrases. In turn phrase utterances are constituted by strings of word utterances which in turn can be regarded as strings of phoneme events. Notice that the temporal relationships between the elements of these strings is more varied than we accounted for in Chapter 1. In dialogue utterances may temporally overlap each other (as indicated in (20) by the notation [...]_n). When we consider adjacent phoneme events in a string overlap becomes the norm (referred to as coarticulation in phonetics). Although we did not take it up in Chapter 1, temporal overlap in event strings is not restricted to speech events. For example, in the game of fetch it is quite often the case that the dog will start running after the stick before the human has finished throwing it. Perceiving temporally overlapping events is part of our basic perceptual

¹For a recent discussion of genre in the kind of framework that we are describing see Ginzburg (2010, 2012).

apparatus.

We will work on developing a type for speech events, *SEvent*. Speech events take time and therefore a time interval will be involved, the “speech time”. Crucial also is the “phonology”, that is the type of the sound produced in the event. This corresponds to the phonology feature in HPSG (Sag *et al.*, 2003). We will say that the phonology is of type *Phon*. For simplicity we might assume that *Phon* is an abbreviation for *Word*⁺ that is a non-empty string of words.² Here *Word* would be a basic type to which word forms of the language belong. A more accurate proposal would be that *Phon* is *Phoneme*⁺³ where *Phoneme* is the basic type to which phonemes of the language belong. This would still be a simplification and an abstraction from the actual events that are being classified, however. A phoneme is rather to be regarded as belonging to a complex type of acoustic and articulatory event. We will keep the simplifying assumptions here as we do not aim to give a detail account of phonology. We also need a condition that links the phonology to the speech time. Thus a proposal for the type *SEvent* might be (21).

$$(21) \quad \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{phon} & : \text{Phon} \\ \text{e} & : \text{uttered_at}(\text{phon}, \text{e-time}) \end{array} \right]$$

where *TimeInt* is the type of time intervals as defined on page 15. To this we might usefully add the speech location as in (22).

$$(22) \quad \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{e-loc} & : \text{Loc} \\ \text{phon} & : \text{Phon} \\ \text{e} & : \text{uttered_at_loc}(\text{phon}, \text{e-time}, \text{e-loc}) \end{array} \right]$$

It might seem natural to add roles of speaker and audience, given what we know about speech act theory (Searle, 1969). Thus we might consider *SEvent* to be the type in (23).

$$(23) \quad \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{e-loc} & : \text{Loc} \\ \text{sp} & : \text{Ind} \\ \text{au} & : \text{Ind} \\ \text{phon} & : \text{Phon} \\ \text{e} & : \text{utter}(\text{sp}, \text{phon}, \text{au}, \text{e-time}, \text{e-loc}) \end{array} \right]$$

²If we want to be more grammatically sophisticated we might want to allow silent speech events by allowing empty phonologies, that is, we say that *Phon* is the type *Word*^{*}.

³Or *Phoneme*^{*}.

However, while many speech events may be considered to be of this type, not all will. Of course, longer speech events like dialogues will not have a single speaker or audience. But even shorter chunks corresponding perhaps to single speech acts do not always have a single speaker or audience. For example, consider split utterances as discussed by Purver *et al.* (2010) who give the example (24).

- (24) A: I heard a shout. Did you
B: Burn myself? No, luckily.

Here we probably want to consider the utterance of *Did you . . . burn myself?* as a speech event on which *A* and *B* collaborate. Otherwise it might be hard to explain how *you* can be interpreted as the subject of *burn*. We have a single predication split across two speakers. Similarly, speakers can address different audiences within the same predicate structure as in (25).

- (25) You [pointing] work with you [pointing] and you [pointing]
work on your own.

Nevertheless, we might consider that the majority of speech events would belong to the more restricted type (23).

We might expect that both (22) and (23) would be subtypes of (21) intuitively. However, technically this will not necessarily be the case because we have chosen different predicates for the event conditions labelled ‘e’. We could fix this in a number of ways. We could stipulate that $\text{uttered_at_loc}(\pi, t, l)$ is a subtype of $\text{uttered_at}(\pi, t)$ and that $\text{utter}(a, \pi, b, t, l)$ is a subtype of $\text{uttered_at_loc}(\pi, t, l)$, or we could define witnesses for these types so that the subtyping is required. Alternatively, we could redefine the speech event types so that the subtyping is required, as in (26).

$$\begin{aligned}
(26) \quad & \text{a. } \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{phon} & : \text{Phon} \\ \text{e} & : [\text{c}_{\text{uttered_at}} : \text{uttered_at}(\text{phon}, \text{e-time})] \end{array} \right] \\
& \text{b. } \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{e-loc} & : \text{Loc} \\ \text{phon} & : \text{Phon} \\ \text{e} & : \left[\begin{array}{l} \text{c}_{\text{uttered_at}} : \text{uttered_at}(\text{phon}, \text{e-time}) \\ \text{c}_{\text{uttered_loc}} : \text{uttered_loc}(\text{phon}, \text{e-loc}) \end{array} \right] \end{array} \right] \\
& \text{c. } \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{e-loc} & : \text{Loc} \\ \text{sp} & : \text{Ind} \\ \text{au} & : \text{Ind} \\ \text{phon} & : \text{Phon} \\ \text{e} & : \left[\begin{array}{l} \text{c}_{\text{uttered_at}} : \text{uttered_at}(\text{phon}, \text{e-time}) \\ \text{c}_{\text{uttered_loc}} : \text{uttered_loc}(\text{phon}, \text{e-loc}) \\ \text{c}_{\text{utter}} : \text{utter}(\text{sp}, \text{phon}, \text{au}) \end{array} \right] \end{array} \right]
\end{aligned}$$

While (26) is possibly the most directly explicit way of requiring the subtyping it is probably the least elegant. We will use *SEvent* below to represent the most general of the types, (21).

2.2 Signs

We interpret many speech events as being associated with a semantic content, but not all. When John in (20) says *It's a wee problem I've had for a, say, about a year now* he is using the speech event to refer to another situation - a situation in which he has dry skin for a period of a year. This is what Barwise and Perry (1983) would refer to as the *described situation* which is distinct from the speech situation. In contrast the doctor's utterance of *Hello* in (20) does not tell us anything about a described situation external to the current conversation, although it does give us information about where we are in the conversation (the beginning) and indicate that the doctor is paying attention. We shall say that the former utterance with a type of described situation and call this the *content* of the utterance. A situation type is an appropriate content for a declarative sentence used to make an assertion. The contents of phrases within such a sentence such as *a wee problem* or *about a year* will be objects which can be combined to produce such a type. The contents of other kinds of speech acts, for example, associated with questions like the doctor's utterance of *what can I do for you today?* will be objects based on situation types, in the case of this question a function which maps actions to a situation type. (See Ginzburg (2012) for a discussion of the kind of treatment of questions we have in mind.)

We can think of this association of content with a speech event in similar terms to prediction of event completion discussed in Section 1.4 of Chapter 1. At least in the case of declarative assertions it is a mapping from an observation of a situation to a type of situation. In the case of the event completion the result of the mapping was a type for the completion of the event so far observed. In the case of the speech event we are relating the observation to a type of situation which is entirely distinct from the speech event. The association is less immediate and more abstract but the underlying mechanism, associating the observation of a situation of a given type with another type and drawing the conclusion that the second type must be non-empty, is the same. We could represent the association by a function of the form (27), corresponding to (19) in Chapter 1.

$$(27) \quad \lambda s : T_{SpEv}(T_{Cnt}[s])$$

This represents a mapping from a speech event s of a given type T_{SpEv} to a type T_{Cnt} which is the content of the speech event. The type T_{Cnt} can depend on s (for example, the type of the described situation may be related to the time or place of the utterance, and so on).

de Saussure (1916) called the association between speech and content a sign and this notion has been taken up in modern linguistics in Head Driven Phrase Structure Grammar (HPSG, Sag *et al.*, 2003). In HPSG a sign is regarded technically as a feature structure and our notion of record type corresponds to a feature structure. One way in which our type system differs from HPSG is that we have both records and record types where HPSG has just feature structures. We will consider a sign to be a record representing a pairing of a speech event and a type representing the content. One advantage of considering a sign as a record rather than a function as in (27) is that there is no directionality in a record as there is in a function. Thus the record can be associated with either interpretation (from speech event to content) or generation (from content to speech event). We can make a straightforward relationship between a function such as (27) and a record type (28).

$$(28) \quad \left[\begin{array}{ll} \text{s-event} & : T_{SpEv} \\ \text{cnt}=T_{Cnt} & : Cnt \end{array} \right]$$

(28) is a type of signs. Notice that the ‘cnt’-field in (28) is a manifest field corresponding to the fact that the function in (27) returns the type T_{Cnt} , not an object of type T_{Cnt} . This means that the ‘cnt’ field in (28) requires that the type itself is in the ‘cnt’ field in a record of the type, that is, in the sign. We have introduced the type Cnt , the type of contents. For the moment we will say that Cnt is the type *RecType*, that is, that contents are record types. This is because, for the moment, we will restrict our attention to declarative sentences. When we come to look at constituents of sentences and speech acts other than assertions we will need to expand Cnt to include other kinds of entities as well. Restricting our attention first to complete declarative sentences is similar to

starting with propositional logic before moving on to more complex analysis. The type *Sign* of signs in general is given in (29).

$$(29) \quad \left[\begin{array}{ll} \text{s-event} & : \text{SEvent} \\ \text{cnt} & : \text{Cnt} \end{array} \right]$$

A record of this type, a sign, will pair a speech event with a content. We will refine our definition of *Sign* as we progress.

2.3 Information exchange in dialogue

We start by considering simple dialogues such as (30) which might occur between two people one of whom is instructing the other about simple facts or between a user and a system where the user is adding simple facts to a database using a natural language interface.

- (30) User: Dudamel is a conductor
 System: Aha
 User: Beethoven is a composer
 System: OK

The job of the dialogue partner identified as “System” is to record the facts in memory and confirm to the dialogue partner identified as “User” that this has happened. It seems straightforward to think of the user’s utterances in (30) as corresponding to signs as described in Section 2.2. For example, the user’s first utterance could be regarded as corresponding to a sign of the type in (31).

$$(31) \quad \left[\begin{array}{ll} \text{s-event:} & \left[\begin{array}{ll} \text{e-time} & : \text{TimeInt} \\ \text{phon} & : \text{“Dudamel is a conductor”} \\ \text{e} & : \text{uttered_at(phon,s-event.e-time)} \end{array} \right] \\ \text{cnt} = [\text{e} & : \text{conductor(dudamel, s-event.e-time)}] \end{array} \right] : \text{RecType}$$

Here “Dudamel is a conductor” is a convenient abbreviation for (32).

- (32) “Dudamel”[^]“is”[^]“a”[^]“conductor”

where for any word w , “ w ” is the type of event where w is uttered. “Dudamel is a conductor” is thus a type of string of events of word utterances and is thus a subtype of *Phon*, given our assumptions in Section 2.1. The content is that Dudamel is a conductor at the time of utterance, that is at the event time of the speech event.

This sign type seems to give us what we need in order to explain how an utterance of *Dudamel is a conductor* can convey the information that Dudamel is a conductor. If both dialogue participants have this sign type among their resources then the User knows that in order to convey this content she has to make an utterance which witnesses the appropriate speech event type. The System knows that on observing a speech event of this type the corresponding content should be recorded.

Things are not as straightforward, however, for the acknowledgements *Aha* and *OK* expressed by the system. It is not obvious whether these utterances are to be regarded as signs at all. Certainly a speech event is involved but one might question what content they have. One suggestion would be that the content of *Aha* uttered after an assertion by the other dialogue partner would be the same as the content of that assertion. Thus the system is expressing the same content as the user. This may or may not be true. But such an analysis seems to be missing a central point about what is going on in this dialogue, namely that the user is making an assertion and the system is acknowledging that the content has been accepted and duly processed. In order to account for this kind of fact Ginzburg in a large body of work has developed the notion of a dialogue gameboard, most recently formulated in terms of TTR in Ginzburg (2012); Ginzburg and Fernández (2010). In the computational dialogue systems literature this has given rise to the Information State Update (ISU) approach (Larsson and Traum, 2001; Larsson, 2002) which is also described in Ginzburg and Fernández (2010). We can regard the gameboard as a record which keeps track of important facts about the dialogue so far. Each dialogue participant will have among their resources a record type which represents their understanding of the current dialogue gameboard. Following Larsson (2002) we place this under the label ‘shared’ in the information state type and also have a field with the label ‘private’ representing information about the state of the dialogue which is not shared with other dialogue participants. This will include, for example, plans for what to say next. In Figure 2.1 we give a schematic view of the information state types associated with each of the dialogue participants in the first exchange in (30).

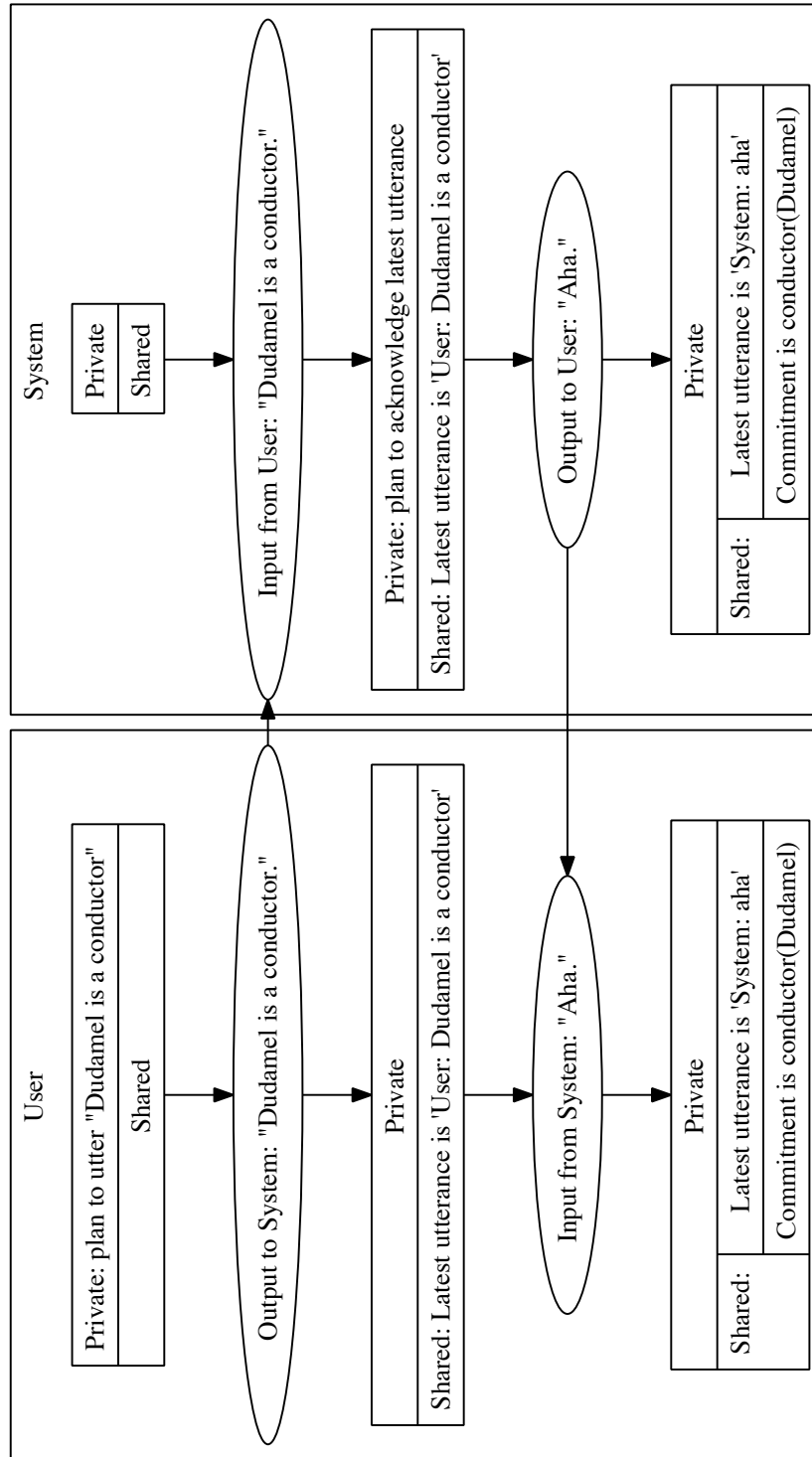


Figure 2.1: Dialogue management: "Dudamel is a conductor"

This assumes ideal communication. There is lots that could go wrong which could have the consequence that the two agents become misaligned and an important part of this framework is to provide a basis for the description of miscommunication as well as communication. (See Ginzburg (2012) for more discussion of this.)

We treat the dialogue information states represented by the square boxes as records as in (33).

$$(33) \quad \left[\begin{array}{lcl} \text{private} & = & \left[\begin{array}{lcl} \text{agenda} & = & A \end{array} \right] \\ \text{shared} & = & \left[\begin{array}{lcl} \text{latest-utterance} & = & LU \\ \text{commitments} & = & C \end{array} \right] \end{array} \right]$$

What kinds of objects should A , LU and C be?

We will define a type *Move* for the type of dialogue moves which will play an important role in characterizing these objects. A should be a list of move types (that is subtype of *Move*) which the agent plans to instantiate. Move types will be record types, that is, they belong to the type *RecType*. A is thus a list of record types and will have type $[RecType]$ since for any type T , $[T]$ is the type of lists all of whose members are of type T (see Appendix A.4).

LU should tell us what move (or moves⁴) has just been carried out. But we will need more information than this. We will need information about what (the agent thinks) was actually said. For this we will use a chart, i.e. a set of edges between vertices representing hypotheses about parts of the utterance, that is, sign types associated with parts of the utterance. The move should be predictable from the chart by a process of move-interpretation for which we will use the predicate ‘m-interp’. Thus LU should itself be of the type (34).

$$(34) \quad \left[\begin{array}{lcl} \text{move} & : & Move \\ \text{chart} & : & Chart \\ e & : & \text{m-interp}(\text{chart}, \text{move}) \end{array} \right]$$

We will shortly say more about the type *Move*. The type *Chart* we will say more about in Chapter 3.

The commitments field has normally been considered as a set of facts or propositions (Ginzburg, 2012; Larsson, 2002). Here we will treat them as a single record type, i.e. a member of the type *RecType*. Using a single type will make it more straightforward to deal with issues like consistency and anaphora [????].

⁴See Larsson (2002) for a proposal where dialogue contributions involve several moves. For now we will make the simplifying assumption that utterances are associated with a single move.

Thus information states can belong to the type (35).

$$(35) \left[\begin{array}{l} \text{private:} [\text{agenda:} [\text{RecType}]] \\ \text{shared:} \left[\begin{array}{l} \text{latest-utterance:} \left[\begin{array}{l} \text{move:} \text{Move} \\ \text{chart:} \text{Chart} \\ \text{e:m-interp}(\text{shared.latest-utterance.chart}, \\ \text{shared.latest-utterance.move}) \end{array} \right] \\ \text{commitments:} \text{RecType} \end{array} \right] \end{array} \right]$$

(35) is, however, no quite general enough. It requires that there always will be a latest utterance. At the beginning of a dialogue this will not be the case and we need a way of representing that there is no previous utterance. We will use a special type *Nil* for this. The only object of this type is ‘nil’. If you like you can think of ‘nil’ as the empty set and *Nil* as a type whose only witness is the empty set. At the beginning of a dialogue there will not be any shared commitments either. However, commitments is a record type and record types are sets of fields (see Appendix A.11) including the empty set. Therefore, it will be natural to use the empty record type ‘[]’ for the commitments at the beginning of a dialogue. Therefore the only adjustment we need to make to (35) in order to include dialogue initial information states is to the shared.latest-utterance field as in (36).

$$(36) \left[\begin{array}{l} \text{private:} [\text{agenda:} [\text{RecType}]] \\ \text{shared:} \left[\begin{array}{l} \text{latest-utterance:} \left[\begin{array}{l} \text{move:} \text{Move} \\ \text{chart:} \text{Chart} \\ \text{e:m-interp}(\text{shared.latest-utterance.chart}, \\ \text{shared.latest-utterance.move}) \end{array} \right] \vee \text{Nil} \\ \text{commitments:} \text{RecType} \end{array} \right] \end{array} \right]$$

(36) uses a join type (Appendix A.6). For any two types T_1 and T_2 you can form the join (or disjunction) $T_1 \vee T_2$. $a : T_1 \vee T_2$ just in case either $a : T_1$ or $a : T_2$.

Dialogue moves are a type of event in which an actor (normally speaker) is related to an intended audience, an illocutionary force (such as ‘assert’) and a content (that is, for our present purposes, a record type such as $[e:\text{conductor}(\text{Dudamel})]$).

We will take the type *Move* to be an abbreviation for (37).

$$(37) \left[\begin{array}{ll} \text{actor} & : \text{Ind} \\ \text{audience} & : \{\text{Ind}\} \\ \text{self} & : \text{Ind} \\ \text{c}_{\text{ego}} & : \text{ego}(\text{self}) \\ \text{orientation} & : \text{eq}(\text{self}, \text{actor}) \vee \text{member}(\text{self}, \text{audience}) \\ \text{i-force} & : \text{IForce} \\ \text{content} & : \text{RecType} \\ \text{e} & : \text{move}(\text{actor}, \text{audience}, \text{i-force}, \text{content}) \end{array} \right]$$

$\{\text{Ind}\}$ is the type of a set of individuals (Appendix A.5). $\text{eq}(\text{self}, \text{actor}) \vee \text{member}(\text{self}, \text{audience})$ is a join type (Appendix A.6). Something will be of this type just in case it is either of type ‘ $\text{eq}(\text{self}, \text{actor})$ ’ (that is, a situation where ‘self’ is identical with ‘actor’) or of type ‘ $\text{member}(\text{self}, \text{audience})$ ’ (that is, a situation where ‘self’ is a member of the set ‘audience’). Thus either the move is one made by self or one which is directed to self. (Note that this allows self to be both the actor and a member of the audience, that is, it allows for people talking to themselves.)

We place a condition on ‘self’ that it be the “ego”, that is, that it involves a situation which is of the type $\text{ego}(\text{self})$. The idea is that this is a type *de se* by which we mean that a judgement that a situation s is of the type $\text{ego}(a)$ requires that a is the individual making the judgement. We do not have a formal treatment of this. We have relativized our judgements in Appendix A to systems of types involving assignments to the basic types but not to the agent making the judgement. We will not develop this further here but merely note that the perception of oneself as the actor or the audience of a dialogue move is an essentially *de se* phenomenon. For recent discussion of and a survey of the literature on *de se* phenomena see Ninan (2010); Hanks (2012).

We will be able to read partial information about a dialogue move from certain aspects of a speech-event. For example, an utterance of *ok* may tell us that the dialogue move is of type (38).

$$(38) \left[\begin{array}{ll} \text{actor} & : \text{Ind} \\ \text{audience} & : \{\text{Ind}\} \\ \text{self} & : \text{Ind} \\ \text{c}_{\text{ego}} & : \text{ego}(\text{self}) \\ \text{orientation} & : \text{member}(\text{self}, \text{audience}) \\ \text{i-force=acknowledge} & : \text{IForce} \\ \text{content} & : \text{RecType} \\ \text{e} & : \text{move}(\text{actor}, \text{audience}, \text{i-force}, \text{content}) \end{array} \right]$$

We will have to look to the previous utterance to find the content of an utterance of *ok*. An utterance of *Dudamel is a conductor* may give us information about the illocutionary force (assertion) and the content ($[\text{e:conductor}(\text{Dudamel})]$). Note, however, that we only get such a fully specified content if we have a unique individual ‘Dudamel’ whom we associate with utterances

of the name *Dudamel*. If the resources we have available do not give us such an individual associated with *Dudamel* then we only get the information that somebody named *Dudamel* is a conductor, that is, we may only get partial information about the content the speaker intended to communicate. Consider the example *Strauss is a composer*. There are at least two famous composers named *Strauss* (and also some more not so famous ones). If our available resources give us two people associated with the name *Strauss* we will not know which of them is being referred to. Representing contents as record types will enable us to handle this content underspecification. However, in order to do this we will need to abandon our current simplifying “propositional logic” assumption that sentences come as unanalyzed wholes associated with their contents. This we will do in Chapter 3.

Why did we not include syntactic information in the move type? The reason is that we want to be able to say that a speaker is carrying out the same move independently of which language they are speaking. Thus, if the user says to the system “*Dudamel is a conductor*” or (in Swedish) “*Dudamel är dirigent*” she will in both cases have carried out the move involving the assertion of the content $[e:\text{conductor}(\text{Dudamel})]$. This abstraction will be important, for example, if we want to change language in the middle of a dialogue, as people sometimes do.⁵ At the same time it is the normal case to continue a dialogue in the same language and thus we need to note which language was used in the previous utterance, i.e. keep track of what was actually said. This information will be in the chart which is part of the latest move. In the chart there will be more information about what was actually said which will be important when it comes to dealing with parts of the utterance for things like clarification and anaphora. But this again requires us to abandon our current simplifying “propositional logic” assumption.

Not surprisingly, when we are dealing with an agenda, a plan for future action, we have got ourselves into a situation where we need types rather than the objects. The things that are on the agenda list are not actual events, but rather *types* of events planned for the future. Normally the types occurring on the agenda will be subtypes of *Move*, though we may wish to include types of events like looking something up in a database, i.e. non-speech events.

We will assume that agents do not have complete information about the information state, that is, they reason in terms of types of information state. The basic intuition behind our reasoning about information state updates can be expressed as in (39).

$$(39) \quad \text{If } r_i : T_i, \text{ then } r_{i+1} : T_{i+1}(r_i)$$

That is, given that we believe that the current information state is of type T_i (recall that we can come to this belief without having any belief about which specific information state is involved), then we can conclude that the next information state is of type T_{i+1} which can depend on the current information state. According to this, we can have a hypothesis about the type of the next

⁵This phenomenon is known as code-switching (Bullock and Toribio, 2009).

information state even though we may not know exactly what the current information state is. Exactly which type the next information state belongs to depends, though, on the exact nature of the current information state. Thus the dependency in our types provides us with an additional means for representing underspecification.

This basic rule of inference corresponds to a function from records to record types, a function of type $(T_i \rightarrow RecType)$, that is, the same kind of function we were using for event prediction in Chapter 1. Such a function is of the form (40).

$$(40) \quad \lambda r : T_i(T_{i+1}[r])$$

Things are a little more complicated than this, however, because this only represents the change from one information state to another, whereas in fact this change is triggered by a speech event which bears an appropriate relation to the current information state represented by r . Thus we are actually interested in functions from the current information state to a function from events to the new information state, as in (41).

$$(41) \quad \lambda r : T_i(\lambda e : T_e[r](T_{i+1}[r, e]))$$

We will call such functions *update functions*.⁶ Let us consider the update function which the user could use in order to update her information state after her own utterance of *Dudamel is a conductor*. This is modelled on the kind of integration rules discussed in Larsson (2002).

$$(42) \quad \lambda r : \left[\begin{array}{l} \text{private} : \left[\text{agenda} : {}_{ne}[RecType] \right] \\ \lambda u : \left[\begin{array}{l} \text{move} : \text{fst}(r.\text{private}.\text{agenda}) \wedge [\text{i-force} = \text{assert} : IForce] \\ \text{chart} : Chart \\ \text{e} : \text{m-interp}(\text{chart}, \text{move}) \end{array} \right] \\ \left(\begin{array}{l} \text{private} : \left[\text{agenda} = \text{rst}(r.\text{private}.\text{agenda}) : [RecType] \right] \\ \text{shared} : \left[\begin{array}{l} \text{latest-utterance} : \left[\begin{array}{l} \text{move} = u.\text{move} : Move \\ \text{chart} = u.\text{chart} : Chart \\ \text{e} = u.e : \text{m-interp}(\text{shared}.\text{latest-utterance}.\text{chart}, \\ \text{shared}.\text{latest-utterance}.\text{move}) \end{array} \right] \end{array} \right] \end{array} \right) \end{array} \right]$$

This function maps information states (records), r , which have a non-empty agenda to a function that maps events to a type of information state. (See Appendix A.4 for an account of non-

⁶This is one of a number of ways of characterizing update in this kind of framework. One might for instance think of the type of the speech event as being part of the current information state. Also instead of using an update function one can use a record type with a ‘preconditions’-field and an ‘effect’-field. Both Ginzburg (2012) and Larsson (2002) have this kind of approach.

empty list types.) It thus requires that the current information state (the first argument to the function) have a non-empty agenda. The second argument to the function (represented by e) requires the move associated with the speech-event to be of the first type on the agenda in r , the current information state, and also to have the illocutionary force ‘assert’ (see Appendix A.7 for a discussion of meet types, that is, conjunction). It also requires that the chart associated with this utterance can be interpreted as a move of that type. The requirements on the arguments to the function represent the preconditions. The type that results from applying the function to its arguments represents the effect of the update. This type requires the agenda to be the tail of r ’s agenda (that is, if we think of the agenda of the current state as a stack, it *pops* the agenda). It furthermore requires the latest-utterance field to contain the move and chart of the event. The idea is that this function should be used to predict the type of the next information state on the basis of the current information state and the observed event. That is, if we believe the current information state to be of the domain type of the update function and we observe an event of the required type then we reason that the updated information state should be of the type resulting from applying the function to the current information state.

We will now examine how such an update function could be used to reason about an update. Let us suppose that the user considers the current information state to be of type:

$$(43) \quad \left[\begin{array}{l} \text{private:} \\ \text{shared:} \end{array} \left[\begin{array}{l} \text{agenda=} \\ \text{latest-utterance:Nil} \\ \text{commitments=[]:RecType} \end{array} \left[\begin{array}{l} \left[\begin{array}{l} \text{actor=User:Ind} \\ \text{audience=\{System\}:\{Ind\}} \\ \text{self=User:Ind} \\ \text{c}_{ego}:ego(\text{self}) \\ \text{orientation:eq(self,actor)} \\ \text{i-force=assert:IForce} \\ \text{content=[e:conductor(Dudamel)] :RecType} \\ \text{e:move(actor, audience, i-force, content)} \end{array} \right] \\ \text{:RecType} \end{array} \right] \right] \right]$$

This represents that the user intends to assert that Dudamel is a conductor represented by the record type $[e:\text{conductor}(\text{Dudamel})]$ where we ignore matters of tense for simplicity. The user also believes that there was no previous utterance and no commitments, i.e. that the planned utterance will be dialogue initial.

Suppose now that the user utters *Dudamel is a conductor* and judges this utterance event u_1 to be a event of type (44).

$$(44) \quad \left[\begin{array}{ll} & \left[\begin{array}{l} \text{actor=User:Ind} \\ \text{audience}=\{\text{System}\}:\{\text{Ind}\} \\ \text{self=User:Ind} \\ \text{c}_{\text{ego}}:\text{ego}(\text{self}) \\ \text{orientation:eq}(\text{self},\text{actor}) \\ \text{i-force}=\text{assert:IForce} \\ \text{content}=[\text{e:conductor}(\text{Dudamel})]:\text{RecType} \\ \text{e:move}(\text{actor}, \text{audience}, \text{i-force}, \text{content}) \end{array} \right] \\ \text{move} & : \\ \text{chart} & : \text{Chart} \\ \text{e} & : \text{m-interp}(\text{chart},\text{move}) \end{array} \right]$$

The user will have more information about the nature of the chart (that is, about what was actually said and how it might be analyzed) than we have represented but we will leave this underspecified for now.

Clearly in the user's judgement the utterance u_1 fulfils the requirements placed on it by (42) since the move interpretation associated with it is of the type which occurs at the head of the agenda. Note that we are reasoning with this function without actually providing it with an argument since we only have a (hypothesized) type of the current information state, not the actual information state. The crucial judgement is that type of the current information state is a subtype of the domain type of the function. This is sufficient to allow us to come to a conclusion about the type of the new information state.

According to the update function the next information state must be of the type (45).

$$(45) \quad \left[\begin{array}{ll} \text{private:} & [\text{agenda}=[]:\text{RecType}] \\ \text{shared:} & \left[\begin{array}{l} \text{latest-utterance:} \left[\begin{array}{l} \text{move}=u_1.\text{move:Move} \\ \text{chart}=u_1.\text{chart:Chart} \\ \text{e}=u_1.\text{e:m-interp}(\text{shared.latest-utterance.chart}, \\ \text{shared.latest-utterance.move}) \end{array} \right] \end{array} \right] \end{array} \right]$$

But we know more about the new information state than what is expressed by the type which results from the update function. Everything we know about the current information state which remains unchanged by the function must be carried over from the current information state. This is related to the frame problem introduced by McCarthy and Hayes (1969).⁷ We shall handle this performing an *asymmetric merge* (see Appendix A.12) of the type we have for the current information state with the type resulting from the update function. The asymmetric merge of two types T_1 and T_2 is represented by $T_1 \sqcup T_2$. If one or both of T_1 and T_2 are non-record types then $T_1 \sqcup T_2$ will be T_2 . If they are both record types, then for any label ℓ which occurs in both T_1

⁷For a recent overview of the frame problem see Shanahan (2009).

and T_2 , $T_1 \sqcup T_2$ will contain a field labelled ℓ with the type resulting from the asymmetric merge of the corresponding types in the ℓ -fields of the two types (in order). For labels which do not occur in both types, $T_1 \sqcup T_2$ will contain the fields from T_1 and T_2 unchanged. In this informal statement we have ignored complications that arise concerning dependent types in record types. This is discussed in Appendix A.12. Our notion of asymmetric merge is related to the notion of priority unification (Shieber, 1986).

Let us see how this works with our example. We have assumed that the type under consideration for the current information state, T_{curr} , is (43) and computed that the predicted type of the updated information state, T_{pr} , is (45). Therefore we need to compute $T_{curr} \sqcup T_{pr}$, that is, (46).

$$(46) \quad \left[\begin{array}{l} \text{private:} \left[\begin{array}{l} \text{agenda=} \left[\begin{array}{l} \text{actor=User:Ind} \\ \text{audience}=\{\text{System}\}:\{\text{Ind}\} \\ \text{i-force}=\text{assert:IForce} \\ \text{content}=[\text{e:conductor(Dudamel)}]:\text{RecType} \\ \text{e:move(actor, audience, i-force, content)} \end{array} \right] \end{array} \right] \\ \text{shared:} \left[\begin{array}{l} \text{latest-utterance:Nil} \\ \text{commitments}=[]:\text{RecType} \end{array} \right] \end{array} \right] :[\text{RecType}] \\ \sqcup \left[\begin{array}{l} \text{private:} \left[\begin{array}{l} \text{agenda}=[]:[\text{RecType}] \end{array} \right] \\ \text{shared:} \left[\begin{array}{l} \text{latest-utterance:} \left[\begin{array}{l} \text{move}=u_1.\text{move:Move} \\ \text{chart}=u_1.\text{chart:Chart} \\ \text{e}=u_1.\text{e:m-interp}(\text{shared.latest-utterance.chart}, \\ \text{shared.latest-utterance.move}) \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

A straightforward way to think of the asymmetric merge of two record types is in terms of the paths in each of them. Both T_{curr} and T_{pr} contain paths ‘private.agenda’. The types at the end of the respective paths, however, are distinct singleton types. (Recall that manifest fields $[\ell=a:T]$ are a convenient notation for $[\ell:T_a]$ where T_a is a restriction of the type T whose only witness is a .) Therefore we include the complete path from the second type in the result of the asymmetric merge. In the case of the path ‘shared.latest-utterance’ we have the non-record type Nil compared with a record type in T_{pr} and therefore we choose the record type in the result since that is what occurs in the second argument of the asymmetric merge operation. Finally, the path ‘shared.commitments’ occurs in the first type but not in the second and therefore it occurs in its form from the first type in the result of the asymmetric merge. The result is given in (47) which represents the type of the new information state which has been computed as a result of the update.

$$(47) \quad \left[\begin{array}{l} \text{private:} \\ \text{shared:} \end{array} \left[\begin{array}{l} \text{agenda}=[]:[RecType] \\ \text{latest-utterance:} \left[\begin{array}{l} \text{move}=u_1.\text{move}:Move \\ \text{chart}=u_1.\text{chart}:Chart \\ \text{e}=u_1.\text{e:m-interp}(\text{shared.latest-utterance.chart}, \\ \text{shared.latest-utterance.move}) \end{array} \right] \\ \text{commitments}=[]:RecType \end{array} \right] \right]$$

Why has the field ‘shared.commitments’ not been updated after the user has asserted that Dudamel is a conductor? This is because the audience has not yet confirmed that they have understood and accepted the move. We assume that our agents are *cautious* and do not assume that commitments are shared until the dialogue participant(s) they are addressing have confirmed acceptance. This interaction is known as grounding and is discussed (among other places) in Traum (1994) and Larsson (2002).

We shall call the update function (42) **IntegrateOwnAssertion** following the style of Larsson (2002) although this does not correspond exactly to any of Larsson’s particular update rules. This then can be used to account for the state that the user is in after asserting that Dudamel is a conductor.

We now need an update function that will account for the effect of this utterance on another dialogue participant. For this we will define a function **IntegrateOtherAssertion** which allows

an agent to integrate a move which it perceives to be an assertion.

$$\begin{aligned}
 (48) \quad & \lambda r: [\text{private} : [\text{agenda} : [RecType]]] \\
 & \lambda u: \left[\begin{array}{l} \text{move} : \left[\begin{array}{l} \text{actor} : Ind \\ \text{audience} : \{Ind\} \\ \text{self} : Ind \\ c_{ego} : ego(self) \\ \text{orientation} : member(self, audience) \\ i\text{-force} = \text{assert} : IForce \\ \text{content} : RecType \end{array} \right] \\ \text{chart} : Chart \\ e : m\text{-interp}(chart, move) \end{array} \right] \\
 & \left(\begin{array}{l} \text{private: agenda=} \left[\begin{array}{l} \text{actor: } Ind \\ \text{audience} = \{u.move.actor\} : \{Ind\} \\ \text{self: } Ind \\ c_{ego}: ego(self) \\ \text{orientation: } eq(self, actor) \\ i\text{-force} = \text{acknowledge: } IForce \\ \text{content} = u.move.content: RecType \end{array} \right] : [RecType] \\ \quad \quad \quad | r.private.agenda \\ \text{shared: latest-utterance:} \left[\begin{array}{l} \text{move} = u.move: Move \\ \text{chart} = u.chart: Chart \\ e = u.e: m\text{-interp}(shared.latest-utterance.chart, \\ \quad \quad \quad shared.latest-utterance.move) \end{array} \right] \end{array} \right)
 \end{aligned}$$

If an agent uses (48) to update then the new information state will contain a move type on the agenda which involves acknowledging the content of the assertion by the other dialogue partner. This update function is also cautious in that it does not yet update the shared commitments since the acknowledgement is only scheduled on the agenda but has not yet been performed.

If an agent performs an acknowledge-event (“ok”) and it can integrate it with the update function **IntegrateOwnAcknowledgement** which will finally perform an update of shared commitments. Before we define this update function we will examine what needs to happen in order to update the commitments.

Suppose that in the dialogue so far it has been established that Dudamel is a conductor and that this is represented by the record type (49).

$$(49) \quad [e : \text{conductor}(Dudamel)]$$

Suppose further that the latest move has the content that Beethoven is a composer, namely (50).

(50) $[e : \text{composer(Beethoven)}]$

One obvious way to combine them would be to merge them, that is, (51a) which is identical with (51b) which in turn is identical with (51c), given the definition in Appendix A.12 which requires that the merge of any two record types which are not both record types is identical with the meet of the two types.

(51) a. $[e : \text{conductor(Dudamel)}] \wedge [e : \text{composer(Beethoven)}]$
 b. $[e : \text{conductor(Dudamel)} \wedge \text{composer(Beethoven)}]$
 c. $[e : \text{conductor(Dudamel)} \wedge \text{composer(Beethoven)}]$

For the simple storing of information represented by predicates and names represented in (30) this might be sufficient. It makes the claim that all the information is collected into one eventuality. In more narrative dialogues referring to separate events which we may wish to be able to refer back to this would be an inadequate solution, however. It would be better if we have a way of keeping the labels ‘e’ separate so that they don’t clash, for example in (52a) which is identical with (52b)

(52) a. $[e_1 : \text{conductor(Dudamel)}] \wedge [e_2 : \text{composer(Beethoven)}]$
 b. $\left[\begin{array}{l} e_1 : \text{conductor(Dudamel)} \\ e_2 : \text{composer(Beethoven)} \end{array} \right]$

The potential problems of label clash become very clear if we consider the types in (53a) corresponding to *a boy hugged a dog* and *a girl stroked a cat*. (53a) is identical with (53b) and has a single individual which is both a girl and a boy stroking another individual which is both a dog and a cat.

$$\begin{aligned}
 (53) \quad & \text{a. } \left[\begin{array}{ll} x & : \text{Ind} \\ c_{\text{boy}} & : \text{boy}(x) \\ y & : \text{Ind} \\ c_{\text{dog}} & : \text{dog}(y) \\ e & : \text{hug}(x,y) \end{array} \right] \wedge \left[\begin{array}{ll} x & : \text{Ind} \\ c_{\text{girl}} & : \text{girl}(x) \\ y & : \text{Ind} \\ c_{\text{cat}} & : \text{cat}(y) \\ e & : \text{stroke}(x,y) \end{array} \right] \\
 & \text{b. } \left[\begin{array}{ll} x & : \text{Ind} \\ c_{\text{boy}} & : \text{boy}(x) \\ c_{\text{girl}} & : \text{girl}(x) \\ y & : \text{Ind} \\ c_{\text{dog}} & : \text{dog}(y) \\ c_{\text{cat}} & : \text{cat}(y) \\ e & : \text{hug}(x,y) \wedge \text{stroke}(x,y) \end{array} \right]
 \end{aligned}$$

One way to get around this problem is to ensure that whenever you introduce new types you always use fresh labels that have not been used before and then use explicit constraints to require identity in cases where it is required. However, when we come to examine compositional semantics in Chapter 3 we will see that it is quite important to refer to particular labels in our rules of combination. Instead of introducing unique *labels* we will use the power of records introduce unique *paths* when contents are combined. We will use the label ‘prev’ (“previous”). If T_{old} is the content so far and T_{new} is the content we wish to add then the new combined content will be as in (54a). Thus adding the content of *a girl stroked a cat* to that of *a boy hugged a dog* will yield (54b).

$$\begin{aligned}
 (54) \quad & \text{a. } \left[\text{prev} : T_{\text{old}} \right] \wedge T_{\text{new}} \\
 & \text{b. } \left[\begin{array}{ll} \text{prev} & : \left[\begin{array}{ll} x & : \text{Ind} \\ c_{\text{boy}} & : \text{boy}(x) \\ y & : \text{Ind} \\ c_{\text{dog}} & : \text{dog}(y) \\ e & : \text{hug}(x,y) \end{array} \right] \\ x & : \text{Ind} \\ c_{\text{girl}} & : \text{girl}(x) \\ y & : \text{Ind} \\ c_{\text{cat}} & : \text{cat}(y) \\ e & : \text{stroke}(x,y) \end{array} \right]
 \end{aligned}$$

In the case of our example with Dudamel and Beethoven the result will be (55).

$$(55) \quad \left[\begin{array}{ll} \text{prev} & : \left[e : \text{conductor}(\text{Dudamel}) \right] \\ e & : \text{composer}(\text{Beethoven}) \end{array} \right]$$

If we add a further fact to this, say, that Uchida is a pianist we would obtain (56)

$$(56) \quad \left[\begin{array}{l} \text{prev} : \left[\begin{array}{l} \text{prev} : [e : \text{conductor}(\text{Dudamel})] \\ e : \text{composer}(\text{Beethoven}) \end{array} \right] \\ e : \text{pianist}(\text{Uchida}) \end{array} \right]$$

This means that we now have to add additional information if we want to require identity, for example if we want the Beethoven and Uchida eventualities (prev.e and e in (56)) to be identical. We will return to these matters when we deal with anaphora in Chapter 3. Note that this strategy also gives us a straightforward record of the order in which content was added.

The update function **IntegrateOwnAcknowledgement** is given in (57).

$$(57) \quad \lambda r: \left[\begin{array}{l} \text{private} : \left[\begin{array}{l} \text{agenda} : {}_{ne}[\text{RecType}] \\ \text{latest-utterance} : \left[\begin{array}{l} \text{move} : [\text{content} : \text{RecType}] \end{array} \right] \\ \text{commitments} : \text{RecType} \end{array} \right] \\ \text{shared} : \left[\begin{array}{l} \text{move} : \text{fst}(r.\text{private}.\text{agenda}) \wedge [\text{i-force}=\text{acknowledge}:\text{IForce}] \\ \text{chart} : \text{Chart} \\ e : \text{m-interp}(\text{chart}, \text{move}) \end{array} \right] \end{array} \right] \\ \lambda u: \left[\begin{array}{l} \text{private}: \left[\begin{array}{l} \text{agenda}=\text{rst}(r.\text{private}.\text{agenda}):[\text{RecType}] \\ \text{latest-utterance}: \left[\begin{array}{l} \text{move}=u.\text{move}:\text{Move} \\ \text{chart}=u.\text{chart}:\text{Chart} \\ e=u.e:\text{m-interp}(\text{shared}.\text{latest-utterance}.\text{chart}, \\ \text{shared}.\text{latest-utterance}.\text{move}) \end{array} \right] \\ \text{commitments}=[\text{prev}:r.\text{commitments}] \wedge r.\text{latest-utterance}.\text{move}.\text{content}:\text{RecType} \end{array} \right] \end{array} \right] \end{array} \right]$$

This function will

1. update the agenda with the result of removing the first item on the agenda in r , the information state prior to update
2. update the latest utterance with the current utterance (e.g. the utterance of *ok*)
3. update the commitments to be the result of placing the commitments of r under the label ‘prev’ and merging with the content of the move in the latest utterance according to r (e.g. the utterance of *Dudamel is a conductor*, not the utterance of *ok*)

We then need an update function **IntegrateOtherAcknowledgement** which is like **IntegrateOwnAcknowledgment** except that it requires that the move event is directed towards the agent doing the updating and does not update the agenda. This is given in (58).

$$\begin{aligned}
 (58) \quad & \lambda r: \left[\text{shared} : \left[\begin{array}{l} \text{latest-utterance} : \left[\text{move} : \left[\text{content} : \text{RecType} \right] \right] \\ \text{commitments} : \text{RecType} \end{array} \right] \right] \\
 & \quad \lambda u: \left[\begin{array}{l} \text{move} : \left[\begin{array}{l} \text{audience:}\{\text{Ind}\} \\ \text{self:Ind} \\ \text{c}_{\text{ego}}.\text{ego}(\text{self}) \\ \text{orientation:member}(\text{self},\text{audience}) \\ \text{i-force=acknowledge:IForce} \end{array} \right] \\ \text{chart} : \text{Chart} \\ \text{e} : \text{m-interp}(\text{chart},\text{move}) \end{array} \right] \\
 & \quad \left(\text{shared:} \left[\begin{array}{l} \text{latest-utterance:} \left[\begin{array}{l} \text{move}=u.\text{move:Move} \\ \text{chart}=u.\text{chart:Chart} \\ \text{e}=u.\text{e:m-interp}(\text{shared.latest-utterance.chart}, \\ \text{shared.latest-utterance.move}) \end{array} \right] \\ \text{commitments}=\left[\text{prev:r.commitments} \right] \wedge r.\text{latest-utterance.move.content:RecType} \end{array} \right] \right)
 \end{aligned}$$

This update function will, on the basis of a speech event expressing an acknowledgment move and addressed to the individual labelled by ‘self’ add the contents of the *previous* move to commitments.

Note that neither of these update functions place constraints on the content of the acknowledgement but rather update the commitments with the content of the previous utterance. This leaves the way open for giving an analysis of acknowledgements without any propositional content. Their occurrence in a dialogue serves as an instruction to add the content of the previous utterance not the current utterance. Thus we see how dialogue moves can make a procedural contribution (“update the commitments with the content of the previous utterance!”) rather than expressing a propositional content themselves.

We have so far talked of update functions in this chapter, functions which given an information state and an utterance will return a type for an updated information state. Update functions specify something about the state that an agent will be in after the occurrence of a certain type of event. We have not, however, specified what it is that will specify that an agent should carry out an action which gives rise to an event of the appropriate type. Formally, these will also be functions which map an information state of a given type to a new type, the type of event which the agent is to bring about. Thus they too will be functions from object to types (or dependent functions). We will call such functions *action functions*. We will introduce one such function, **ExecTopAgenda**, which takes an information state with a non-empty agenda and returns a type for a move of that type and a chart which can be interpreted as that move. It is given in (59).

(59) $\lambda r: [\text{private} : [\text{agenda} : {}_{ne}[\text{RecType}]]]$
 $(\left[\begin{array}{ll} \text{move} & : \text{fst}(r.\text{private}.\text{agenda}) \\ \text{chart} & : \text{Chart} \\ \text{e} & : \text{m-interp}(\text{chart}, \text{move}) \end{array} \right])$

2.4 Resources

While there is no formal distinction between update functions and action functions they are to be used in different ways. Update functions are to be used as instructions to conclude that there is something of the resulting type. Action functions are to be used as instructions to create something of the resulting type. The formal nature of the functions themselves are the same but they are to be used in different ways. We shall say that they are different kinds of *resources* that are available to an agent. The update and action functions we have discussed in this chapter belong to a general resource for *dialogue management*. We shall see that there are a number of resources which contain both update and action functions and that in general they can be viewed as the two kinds of enthymemes (inferential and imperative) discussed in Breitholtz' work in progress on Aristotelian enthymemes (Breitholtz and Villing, 2008; Breitholtz, 2010; Breitholtz and Cooper, 2011).

We need more resources: signs and move-interpretations of charts containing signs. @@

Appendix A

Type theory with records

Unless otherwise stated this is the version of TTR presented in Cooper (2012).

A.1 Basic types

A *system of basic types* is a pair:

$$\mathbf{TYPE}_B = \langle \mathbf{Type}, A \rangle$$

where:

1. **Type** is a non-empty set
2. A is a function whose domain is **Type**
3. for any $T \in \mathbf{Type}$, $A(T)$ is a set disjoint from **Type**
4. for any $T \in \mathbf{Type}$, $a :_{\mathbf{TYPE}_B} T$ iff $a \in A(T)$

A *modal system of basic types*¹ is a family of pairs:

$$\mathbf{TYPE}_{MB} = \langle \mathbf{Type}, A \rangle_{A \in \mathcal{A}}$$

¹This definition was not present in Cooper (2012).

where:

1. \mathcal{A} is a set of functions with domain **Type**
2. for each $A \in \mathcal{A}$, $\langle \mathbf{Type}, A \rangle$ is a system of basic types

This enables us to define some simple modal notions:

If $\mathbf{TYPE}_{MB} = \langle \mathbf{Type}, A \rangle_{A \in \mathcal{A}}$ is a modal system of basic types, we shall use the notation \mathbf{TYPE}_{MB_A} (where $A \in \mathcal{A}$) to refer to that system of basic types in \mathbf{TYPE}_{MB} whose type assignment is A . Then:

1. for any $T_1, T_2 \in \mathbf{Type}$, T_1 is (necessarily) equivalent to T_2 in \mathbf{TYPE}_{MB} , $T_1 \approx_{\mathbf{TYPE}_{MB}} T_2$, iff for all $A \in \mathcal{A}$, $\{a \mid a : \mathbf{TYPE}_{MB_A} T_1\} = \{a \mid a : \mathbf{TYPE}_{MB_A} T_2\}$
2. for any $T_1, T_2 \in \mathbf{Type}$, T_1 is a subtype of T_2 in \mathbf{TYPE}_{MB} , $T_1 \sqsubseteq_{\mathbf{TYPE}_{MB}} T_2$, iff for all $A \in \mathcal{A}$, $\{a \mid a : \mathbf{TYPE}_{MB_A} T_1\} \subseteq \{a \mid a : \mathbf{TYPE}_{MB_A} T_2\}$
3. for any $T \in \mathbf{Type}$, T is necessary in \mathbf{TYPE}_{MB} iff for all $A \in \mathcal{A}$, $\{a \mid a : \mathbf{TYPE}_{MB_A} T\} \neq \emptyset$
4. for any $T \in \mathbf{Type}$, T is possible in \mathbf{TYPE}_{MB} iff for some $A \in \mathcal{A}$, $\{a \mid a : \mathbf{TYPE}_{MB_A} T\} \neq \emptyset$

A.2 Complex types

A.2.1 Predicates

We start by introducing the notion of a predicate signature.

A *predicate signature* is a triple

$$\langle \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle$$

where:

1. **Pred** is a set (of predicates)

2. **ArgIndices** is a set (of indices for predicate arguments, normally types)
3. *Arity* is a function with domain **Pred** and range included in the set of finite sequences of members of **ArgIndices**.

A *polymorphic predicate signature* is a triple

$$\langle \mathbf{Pred}, \mathbf{ArgIndices}, \mathit{Arity} \rangle$$

where:

1. **Pred** is a set (of predicates)
2. **ArgIndices** is a set (of indices for predicate arguments, normally types)
3. *Arity* is a function with domain **Pred** and range included in the powerset of the set of finite sequences of members of **ArgIndices**.

A.2.2 Systems of complex types

A *system of complex types* is a quadruple:

$$\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathit{Arity} \rangle, \langle A, F \rangle \rangle$$

where:

1. $\langle \mathbf{BType}, A \rangle$ is a system of basic types
2. $\mathbf{BType} \subseteq \mathbf{Type}$
3. for any $T \in \mathbf{Type}$, if $a :_{\langle \mathbf{BType}, A \rangle} T$ then $a :_{\mathbf{TYPE}_C} T$
4. $\langle \mathbf{Pred}, \mathbf{ArgIndices}, \mathit{Arity} \rangle$ is a (polymorphic) predicate signature
- 5.² $P(a_1, \dots, a_n) \in \mathbf{PType}$ iff $P \in \mathbf{Pred}, T_1 \in \mathbf{Type}, \dots, T_n \in \mathbf{Type}, \mathit{Arity}(P) = \langle T_1, \dots, T_n \rangle$
 $(\langle T_1, \dots, T_n \rangle \in \mathit{Arity}(P))$ and $a_1 :_{\mathbf{TYPE}_C} T_1, \dots, a_n :_{\mathbf{TYPE}_C} T_n$

²This clause has been modified since Cooper (2012) where it was a conditional rather than a biconditional.

6. $\mathbf{PType} \subseteq \mathbf{Type}$

7. for any $T \in \mathbf{PType}$, $F(T)$ is a set disjoint from \mathbf{Type}
8. for any $T \in \mathbf{PType}$, $a :_{\mathbf{TYPE}_C} T$ iff $a \in F(T)$

We call the pair $\langle A, F \rangle$ in a complex system of types the *model* because of its similarity to first order models in providing values for the basic types and the ptypes constructed from predicates and arguments. It is this pair which connects the system of types to the non-type theoretical world of objects and situations.

A.3 Function types

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has *function types* if

1. for any $T_1, T_2 \in \mathbf{Type}$, $(T_1 \rightarrow T_2) \in \mathbf{Type}$
2. for any $T_1, T_2 \in \mathbf{Type}$, $f :_{\mathbf{TYPE}_C} (T_1 \rightarrow T_2)$ iff f is a function whose domain is $\{a \mid a :_{\mathbf{TYPE}_C} T_1\}$ and whose range is included in $\{a \mid a :_{\mathbf{TYPE}_C} T_2\}$

A.4 List types

List types were not included in Cooper (2012).

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has *list types* if

1. for any $T \in \mathbf{Type}$, $[T] \in \mathbf{Type}$
2. for any $T \in \mathbf{Type}$,
 - (a) $a \mid L :_{\mathbf{TYPE}_C} [T]$ iff $a :_{\mathbf{TYPE}_C} T$ and $L :_{\mathbf{TYPE}_C} [T]$
 - (b) $[] :_{\mathbf{TYPE}_C} [T]$

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has *non-empty list types* if

1. for any $T \in \mathbf{Type}$, $_{ne}[T] \in \mathbf{Type}$
2. for any $T \in \mathbf{Type}$,
 - (a) $a \mid L :_{\mathbf{TYPE}_C} _{ne}[T]$ iff $a :_{\mathbf{TYPE}_C} T$ and $L :_{\mathbf{TYPE}_C} _{ne}[T]$
 - (b) $[a] :_{\mathbf{TYPE}_C} _{ne}[T]$ iff $a : T$

If $a \mid L :_{\mathbf{TYPE}_C} _{ne}[T]$ for some system of complex types \mathbf{TYPE}_C and type T , then we use $\text{fst}(a \mid L)$ to refer to a and $\text{rst}(a \mid L)$ to refer to L .

A.5 Set types

Set types were not included in Cooper (2012).

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has set types if

1. for any $T \in \mathbf{Type}$, $\{T\} \in \mathbf{Type}$
2. for any $T \in \mathbf{Type}$, $A :_{\mathbf{TYPE}_C} \{T\}$ iff for all $a \in A$, $a :_{\mathbf{TYPE}_C} T$

A.6 Join types

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has join types if

1. for any $T_1, T_2 \in \mathbf{Type}$, $(T_1 \vee T_2) \in \mathbf{Type}$
2. for any $T_1, T_2 \in \mathbf{Type}$, $a :_{\mathbf{TYPE}_C} (T_1 \vee T_2)$ iff $a :_{\mathbf{TYPE}_C} T_1$ or $a :_{\mathbf{TYPE}_C} T_2$

A.7 Meet types

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has meet types if

1. for any $T_1, T_2 \in \mathbf{Type}$, $(T_1 \wedge T_2) \in \mathbf{Type}$
2. for any $T_1, T_2 \in \mathbf{Type}$, $a :_{\mathbf{TYPE}_C} (T_1 \wedge T_2)$ iff $a :_{\mathbf{TYPE}_C} T_1$ and $a :_{\mathbf{TYPE}_C} T_2$

A.8 Models and modal systems of types

A modal system of complex types provides a collection of models, \mathcal{M} , so that we can talk about properties of the whole collection of type assignments provided by the various models $M \in \mathcal{M}$.

A modal system of complex types based on \mathcal{M} is a family of quadruples³:

$$\mathbf{TYPE}_{MC} = \langle \mathbf{Type}_M, \mathbf{BType}, \langle \mathbf{PType}_M, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, M \rangle_{M \in \mathcal{M}}$$

where for each $M \in \mathcal{M}$, $\langle \mathbf{Type}_M, \mathbf{BType}, \langle \mathbf{PType}_M, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, M \rangle$ is a system of complex types.

This enables us to define modal notions:

If $\mathbf{TYPE}_{MC} = \langle \mathbf{Type}_M, \mathbf{BType}, \langle \mathbf{PType}_M, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, M \rangle_{M \in \mathcal{M}}$ is a modal system of complex types based on \mathcal{M} , we shall use the notation \mathbf{TYPE}_{MC_M} (where $M \in \mathcal{M}$) to refer to that system of complex types in \mathbf{TYPE}_{MC} whose model is M . Let $\mathbf{Type}_{MC_{restr}}$ be $\bigcap_{M \in \mathcal{M}} \mathbf{Type}_M$, the “restrictive” set of types which occur in all possibilities, and $\mathbf{Type}_{MC_{incl}}$ be $\bigcup_{M \in \mathcal{M}} \mathbf{Type}_M$, the “inclusive” set of types which occur in at least one possibility. Then we can define modal notions either restrictively or inclusively (indicated by the subscripts r and i respectively):

restrictive modal notions

1. for any $T_1, T_2 \in \mathbf{Type}_{MC_{restr}}$, T_1 is (necessarily) equivalent_r to T_2 in \mathbf{TYPE}_{MC} , $T_1 \approx_{\mathbf{TYPE}_{MC}} T_2$, iff for all $M \in \mathcal{M}$, $\{a \mid a :_{\mathbf{TYPE}_{MC_M}} T_1\} = \{a \mid a :_{\mathbf{TYPE}_{MC_M}} T_2\}$
2. for any $T_1, T_2 \in \mathbf{Type}_{MC_{restr}}$, T_1 is a subtype_r of T_2 in \mathbf{TYPE}_{MC} , $T_1 \sqsubseteq_{\mathbf{TYPE}_{MC}} T_2$, iff for all $M \in \mathcal{M}$, $\{a \mid a :_{\mathbf{TYPE}_{MC_M}} T_1\} \subseteq \{a \mid a :_{\mathbf{TYPE}_{MC_M}} T_2\}$
3. for any $T \in \mathbf{Type}_{MC_{restr}}$, T is necessary_r in \mathbf{TYPE}_{MC} iff for all $M \in \mathcal{M}$, $\{a \mid a :_{\mathbf{TYPE}_{MC_M}} T\} \neq \emptyset$

³This definition has been modified since Cooper (2012) to make \mathbf{PType} and \mathbf{Type} be relativized to the model M .

4. for any $T \in \mathbf{Type}_{MC_{restr}}$, T is possible_r in \mathbf{TYPE}_{MC} iff for some $M \in \mathcal{M}$, $\{a \mid a : \mathbf{TYPE}_{MC_M} T\} \neq \emptyset$

inclusive modal notions

1. for any $T_1, T_2 \in \mathbf{Type}_{MC_{incl}}$, T_1 is (necessarily) equivalent_i to T_2 in \mathbf{TYPE}_{MC} , $T_1 \approx_{\mathbf{TYPE}_{MC}} T_2$, iff for all $M \in \mathcal{M}$, if T_1 and T_2 are members of \mathbf{Type}_M , then $\{a \mid a : \mathbf{TYPE}_{MC_M} T_1\} = \{a \mid a : \mathbf{TYPE}_{MC_M} T_2\}$
2. for any $T_1, T_2 \in \mathbf{Type}_{MC_{incl}}$, T_1 is a subtype_i of T_2 in \mathbf{TYPE}_{MC} , $T_1 \sqsubseteq_{\mathbf{TYPE}_{MC}} T_2$, iff for all $M \in \mathcal{M}$, if T_1 and T_2 are members of \mathbf{Type}_M , then $\{a \mid a : \mathbf{TYPE}_{MC_M} T_1\} \subseteq \{a \mid a : \mathbf{TYPE}_{MC_M} T_2\}$
3. for any $T \in \mathbf{Type}_{MC_{incl}}$, T is necessary_i in \mathbf{TYPE}_{MC} iff for all $M \in \mathcal{M}$, if $T \in \mathbf{Type}_M$, then $\{a \mid a : \mathbf{TYPE}_{MC_M} T\} \neq \emptyset$
4. for any $T \in \mathbf{Type}_{MC_{incl}}$, T is possible_i in \mathbf{TYPE}_{MC} iff for some $M \in \mathcal{M}$, if $T \in \mathbf{Type}_M$, then $\{a \mid a : \mathbf{TYPE}_{MC_M} T\} \neq \emptyset$

It is easy to see that if any of the restrictive definitions holds for given types in a particular system then the corresponding inclusive definition will also hold for those types in that system.

A.9 Strings and regular types

A string algebra over a set of objects O is a pair $\langle S, \frown \rangle$ where:

1. S is the closure of $O \cup \{e\}$ (e is the empty string) under the binary operation ‘ \frown ’ (“concatenation”)
2. for any s in S , $e \frown s = s \frown e = s$
3. for any s_1, s_2, s_3 in S , $(s_1 \frown s_2) \frown s_3 = s_1 \frown (s_2 \frown s_3)$. For this reason we normally write $s_1 \frown s_2 \frown s_3$ or more simply $s_1 s_2 s_3$.

The objects in S are called strings. Strings have length. e has length 0, any object in O has length

1. If s is a string in S with length n and a is an object in O then $s \frown a$ has length $n + 1$. We use $s[n]$ to represent the n th element of string s .

We can define types whose elements are strings. Such types correspond to regular expressions and we will call them *regular types*. Here we will define just two kinds of such types: concatenation types and Kleene-+ types.

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has concatenation types if

1. (a) for any $T_1, T_2 \in \mathbf{Type}$, $(T_1 \frown T_2) \in \mathbf{Type}$
 (b) for any $T_1, T_2, T_3 \in \mathbf{Type}$, $(T_1 \frown T_2) \frown T_3 = T_1 \frown (T_2 \frown T_3)$ ⁴
2. $a : T_1 \frown T_2$ iff $a = x \frown y$, $x : T_1$ and $y : T_2$

\mathbf{TYPE}_C has Kleene-+ types if

1. for any $T \in \mathbf{Type}$, $T^+ \in \mathbf{Type}$
2. $a : T^+$ iff $a = x_1 \frown \dots \frown x_n$, $n > 0$ and for i , $1 \leq i \leq n$, $x_i : T$

Strings are used standardly in formal language theory where strings of symbols or strings of words are normally considered. Following important insights by Tim Fernando Fernando (2004, 2006, 2008, 2009) we shall be concerned rather with strings of events. We use informal notations like ‘ “sam” ’ and ‘ “ran” ’ to represent phonological types of speech events (utterances of *Sam* and *ran*). Thus ‘ “sam” \frown “ran” ’ is the type of speech events which are concatenations of an utterance of *Sam* and an utterance of *ran*.

A.10 The type *Type* and stratification

An *intensional system of complex types* is a family of quadruples indexed by the natural numbers:

$$\mathbf{TYPE}_{IC} = \langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F^n \rangle \rangle_{n \in \mathbf{Nat}}$$

where (using \mathbf{TYPE}_{IC_n} to refer to the quadruple indexed by n):

1. for each n , $\langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F^n \rangle \rangle$ is a system of complex types

⁴This has been added to the definition in Cooper (2012) to make associativity explicit.

2. for each n , $\mathbf{Type}^n \subseteq \mathbf{Type}^{n+1}$ and $\mathbf{PType}^n \subseteq \mathbf{PType}^{n+1}$
3. for each n , if $T \in \mathbf{PType}^n$ and $p \in F^n(T)$ then $p \in F^{n+1}(T)$
4. for each $n > 0$, $Type^n \in \mathbf{Type}^n$
5. for each $n > 0$, $T :_{\mathbf{TYPE}_{IC_n}} Type^n$ iff $T \in \mathbf{Type}^{n-1}$

An intensional system of complex types \mathbf{TYPE}_{IC} ,

$$\mathbf{TYPE}_{IC} = \langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Ariety} \rangle, \langle A, F^n \rangle \rangle_{n \in Nat}$$

has dependent function types if

1. for any $n > 0$, $T \in \mathbf{Type}^n$ and $\mathcal{F} :_{\mathbf{TYPE}_{IC_n}} (T \rightarrow Type^n)$, $((a : T) \rightarrow \mathcal{F}(a)) \in \mathbf{Type}^n$
2. for each $n > 0$, $f :_{\mathbf{TYPE}_{IC_n}} ((a : T) \rightarrow \mathcal{F}(a))$ iff f is a function whose domain is $\{a \mid a :_{\mathbf{TYPE}_{IC_n}} T\}$ and such that for any a in the domain of f , $f(a) :_{\mathbf{TYPE}_{IC_n}} \mathcal{F}(a)$.

We might say that on this view dependent function types are “semi-intensional” in that they depend on there being a type of types for their definition but they do not introduce types as arguments to predicates and do not involve the definition of orders of types in terms of the types of the next lower order.

Putting the definition of a modal type system and an intensional type system together we obtain:⁵

An *intensional modal system of complex types based on* \mathfrak{M} is a family, indexed by the natural numbers, of families of quadruples indexed by members of \mathfrak{M} :

$$\mathbf{TYPE}_{IMC} = \langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Ariety} \rangle, \mathcal{M}_n \rangle_{\mathcal{M} \in \mathfrak{M}, n \in Nat}$$

where:

1. for each n , $\langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Ariety} \rangle, \mathcal{M}_n \rangle_{\mathcal{M} \in \mathfrak{M}}$ is a modal system of complex types based on $\{\mathcal{M}_n \mid \mathcal{M} \in \mathfrak{M}\}$
2. for each $\mathcal{M} \in \mathfrak{M}$, $\langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Ariety} \rangle, \mathcal{M}_n \rangle_{n \in Nat}$ is an intensional system of complex types

⁵This explicit definition was not present in Cooper (2012).

A.11 Record types

In this section we will define what it means for a system of complex types to have record types. The objects of record types, that is, records, are themselves structured mathematical objects of a particular kind and we will start by characterizing them.

A *record* is a finite set of ordered pairs (called *fields*) which is the graph of a function. If r is a record and $\langle \ell, v \rangle$ is a field in r we call ℓ a *label* and v a *value* in r and we use $r.\ell$ to denote v . $r.\ell$ is called a *path* in r .

We will use a tabular format to represent records. A record $\{\langle \ell_1, v_1 \rangle, \dots, \langle \ell_n, v_n \rangle\}$ is displayed as

$$\begin{bmatrix} \ell_1 & = & v_1 \\ \dots & & \\ \ell_n & = & v_n \end{bmatrix}$$

A value may itself be a record and paths may extend into embedded records. A record which contains records as values is called a *complex record* and otherwise a record is *simple*. Values which are not records are called *leaves*. Consider a record r

$$\begin{bmatrix} f & = & \begin{bmatrix} f & = & \begin{bmatrix} ff & = & a \\ gg & = & b \end{bmatrix} \\ g & = & c \end{bmatrix} \\ g & = & \begin{bmatrix} h & = & \begin{bmatrix} g & = & a \\ h & = & d \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Among the paths in r are $r.f$, $r.g.h$ and $r.f.f.ff$ which denote, respectively,

$$\begin{bmatrix} f & = & \begin{bmatrix} ff & = & a \\ gg & = & b \end{bmatrix} \\ g & = & c \end{bmatrix}$$

$$\begin{bmatrix} g & = & a \\ h & = & d \end{bmatrix}$$

and a . We will make a distinction between *absolute paths*, such as those we have already mentioned, which consist of a record followed by a series of labels connected by dots and *relative*

paths which are just a series of labels connected by dots, e.g. $g.h$. Relative paths are useful when we wish to refer to similar paths in different records. We will use *path* to refer to either absolute or relative paths when it is clear from the context which is meant. The set of leaves of r , also known as its *extension* (those objects other than labels which it contains), is $\{a, b, c, d\}$. The bag (or multiset) of leaves of r , also known as its *multiset extension*, is $\{a, a, b, c, d\}$. A record may be regarded as a way of labelling and structuring its extension. Two records are (*multiset*) *extensionally equivalent* if they have the same (multiset) extension. Two important, though trivial, facts about records are:

Flattening. For any record r , there is a multiset extensionally equivalent simple record. We can define an operation of flattening on records which will always produce an equivalent simple record. In the case of our example, the result of flattening is

$$\left[\begin{array}{lcl} f.f.ff & = & a \\ f.f.gg & = & b \\ f.g & = & c \\ g.h.g & = & a \\ g.h.h & = & d \end{array} \right]$$

assuming the flattening operation uses paths from the original record in a rather obvious way to create unique labels for the new record.

Relabelling. For any record r , if $\pi_1.\ell.\pi_2$ is a path π in r , and $\pi_1.\ell'.\pi_2'$ is *not* a path in r (for any π_2'), then substituting ℓ' for the occurrence of ℓ in π results in a record which is multiset equivalent to r . We could, for example, substitute k for the second occurrence of g in the path $g.h.g$ in our example record.

$$\left[\begin{array}{lcl} f & = & \left[\begin{array}{lcl} f & = & \left[\begin{array}{lcl} ff & = & a \\ gg & = & b \end{array} \right] \\ g & = & c \end{array} \right] \\ g & = & \left[\begin{array}{lcl} h & = & \left[\begin{array}{lcl} k & = & a \\ h & = & d \end{array} \right] \end{array} \right] \end{array} \right]$$

A record *type* is a record in the general sense defined above where the values in its fields are types or, in some cases, certain kinds of mathematical objects which can be used to construct types.

A record r is *well-typed* with respect to a system of types **TYPE** with set of types **Type** and a set of labels L iff for each field $\langle \ell, a \rangle \in r$, $\ell \in L$ and either $a :_{\mathbf{TYPE}} T$ for some $T \in \mathbf{Type}$ or a is itself a record which is well-typed with respect to **TYPE** and L .

A system of complex types $\mathbf{TYPE}_C = \langle \mathbf{Type}, \mathbf{BType}, \langle \mathbf{PType}, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F \rangle \rangle$ has record types based on $\langle L, \mathbf{RType} \rangle$, where L is a countably infinite set (of labels) and $\mathbf{RType} \subseteq \mathbf{Type}$, where \mathbf{RType} is defined by:

1. $Rec \in \mathbf{RType}$
2. $r :_{\mathbf{TYPE}_C} Rec$ iff r is a well-typed record with respect to \mathbf{TYPE}_C and L .
3. if $\ell \in L$ and $T \in \mathbf{Type}$, then $\{\langle \ell, T \rangle\} \in \mathbf{RType}$.
4. $r :_{\mathbf{TYPE}_C} \{\langle \ell, T \rangle\}$ iff $r :_{\mathbf{TYPE}_C} Rec$, $\langle \ell, a \rangle \in r$ and $a :_{\mathbf{TYPE}_C} T$.
5. if $R \in \mathbf{RType}$, $\ell \in L$, ℓ does not occur as a label in R (i.e. there is no field $\langle \ell', T' \rangle$ in R such that $\ell' = \ell$), then $R \cup \{\langle \ell, T \rangle\} \in \mathbf{RType}$.
6. $r :_{\mathbf{TYPE}_C} R \cup \{\langle \ell, T \rangle\}$ iff $r :_{\mathbf{TYPE}_C} R$, $\langle \ell, a \rangle \in r$ and $a :_{\mathbf{TYPE}_C} T$.

This gives us non-dependent record types in a system of complex types. We can extend this to intensional systems of complex types (with stratification).

An *intensional system of complex types* $\mathbf{TYPE}_{IC} = \langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F^n \rangle \rangle_{n \in \mathbf{Nat}}$ has record types based on $\langle L, \mathbf{RType}^n \rangle_{n \in \mathbf{Nat}}$ if for each n , $\langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F^n \rangle \rangle$ has record types based on $\langle L, \mathbf{RType}^n \rangle$ and

1. for each n , $\mathbf{RType}^n \subseteq \mathbf{RType}^{n+1}$
2. for each $n > 0$, $RecType^n \in \mathbf{RType}^n$
3. for each $n > 0$, $T :_{\mathbf{TYPE}_{IC_n}} RecType^n$ iff $T \in \mathbf{RType}^{n-1}$

Intensional type systems may in addition contain *dependent* record types.

An *intensional system of complex types* $\mathbf{TYPE}_{IC} = \langle \mathbf{Type}^n, \mathbf{BType}, \langle \mathbf{PType}^n, \mathbf{Pred}, \mathbf{ArgIndices}, \mathbf{Arity} \rangle, \langle A, F^n \rangle \rangle_{n \in \mathbf{Nat}}$ has dependent record types based on $\langle L, \mathbf{RType}^n \rangle_{n \in \mathbf{Nat}}$, if it has records types based on $\langle L, \mathbf{RType}^n \rangle_{n \in \mathbf{Nat}}$ and for each $n > 0$

1. if R is a member of \mathbf{RType}^n , $\ell \in L$ not occurring as a label in R , $T_1, \dots, T_m \in \mathbf{Type}^n$, $R.\pi_1, \dots, R.\pi_m$ are paths in R and \mathcal{F} is a function of type $((a_1 : T_1) \rightarrow \dots \rightarrow ((a_m : T_m) \rightarrow Type^n) \dots)$, then $R \cup \{\langle \ell, \langle \mathcal{F}, \langle \pi_1, \dots, \pi_m \rangle \rangle \rangle\} \in \mathbf{RType}^n$.
2. $r :_{\mathbf{TYPE}_{IC_n}} R \cup \{\langle \ell, \langle \mathcal{F}, \langle \pi_1, \dots, \pi_m \rangle \rangle \rangle\}$ iff $r :_{\mathbf{TYPE}_{IC_n}} R$, $\langle \ell, a \rangle$ is a field in r , $r.\pi_1 :_{\mathbf{TYPE}_{IC_n}} T_1, \dots, r.\pi_m :_{\mathbf{TYPE}_{IC_n}} T_m$ and $a :_{\mathbf{TYPE}_{IC_n}} \mathcal{F}(r.\pi_1, \dots, r.\pi_m)$.

We represent a record type $\{\langle \ell_1, T_1 \rangle, \dots, \langle \ell_n, T_n \rangle\}$ graphically as

$$\left[\begin{array}{ll} \ell_1 & : T_1 \\ \dots & \\ \ell_n & : T_n \end{array} \right]$$

In the case of dependent record types we sometimes use a convenient notation representing e.g.

$$\langle \lambda u \lambda v \text{ love}(u, v), \langle \pi_1, \pi_2 \rangle \rangle$$

as

$$\text{love}(\pi_1, \pi_2)$$

Our systems now allow both function types and dependent record types and allow dependent record types to be arguments to functions. We have to be careful when considering what the result of applying a function to a dependent record type should be. Consider the following simple example:

$$\lambda v_0 : \text{RecType}([c_0 : v_0])$$

What should be the result of applying this function to the record type

$$\left[\begin{array}{ll} x & : \text{Ind} \\ c_1 & : \langle \lambda v_1 : \text{Ind}(\text{dog}(v_1)), \langle x \rangle \rangle \end{array} \right]$$

Given normal assumptions about function application the result would be

$$\left[\begin{array}{ll} c_0 & : \left[\begin{array}{ll} x & : \text{Ind} \\ c_1 & : \langle \lambda v_1 : \text{Ind}(\text{dog}(v_1)), \langle x \rangle \rangle \end{array} \right] \end{array} \right]$$

but this would be incorrect. In fact it is not a well-formed record type since x is not a path in it. Instead the result should be

$$\left[\begin{array}{ll} c_0 & : \left[\begin{array}{ll} x & : \text{Ind} \\ c_1 & : \langle \lambda v_1 : \text{Ind}(\text{dog}(v_1)), \langle c_0.x \rangle \rangle \end{array} \right] \end{array} \right]$$

where the path from the top of the record type is specified. Note that this adjustment is only required when a record type is being substituted into a position that lies on a path within a resulting record type. It will not, for example, apply in a case where a record type is to be substituted for an argument to a predicate such as when applying the function

$$\lambda v_0 : \text{RecType}([c_0 : \text{appear}(v_0)])$$

to

$$\left[\begin{array}{ll} x & : \text{Ind} \\ c_1 & : \langle \lambda v : \text{Ind}(\text{dog}(v)), \langle x \rangle \rangle \\ c_2 & : \langle \lambda v : \text{Ind}(\text{approach}(v)), \langle x \rangle \rangle \end{array} \right]$$

where the position of v_0 is in an “intensional context”, that is, as the argument to a predicate and there is no path to this position in the record type resulting from applying the function. Here the result of the application is

$$\left[\begin{array}{ll} c_0 & : \text{appear} \left(\left[\begin{array}{ll} x & : \text{Ind} \\ c_1 & : \langle \lambda v : \text{Ind}(\text{dog}(v)), \langle x \rangle \rangle \\ c_2 & : \langle \lambda v : \text{Ind}(\text{approach}(v)), \langle x \rangle \rangle \end{array} \right] \right) \end{array} \right]$$

with no adjustment necessary to the paths representing the dependencies.⁶ (Note that ‘ $c_0.x$ ’ is not a path in this record type.)

These matters arise as a result of our choice of using paths to represent dependencies in record types (rather than, for example, introducing additional unique identifiers to keep track of the positions within a record type as has been suggested by Thierry Coquand). It seems like a matter of implementation rather than a matter of substance and it is straightforward to define a path-aware notion of substitution which can be used in the definition of what it means to apply a TTR function to an argument. If f is a function represented by $\lambda v : T(\phi)$ and α is the representation of an object of type T , then the result of applying f to α , $f(\alpha)$, is represented by $\text{Subst}(\alpha, v, \phi, \emptyset)$, that is, the result of substituting α for v in ϕ with respect to the empty path where for arbitrary α, v, ϕ, π , $\text{Subst}(\alpha, v, \phi, \pi)$ is defined as

1. $\text{extend-paths}(\alpha, \pi)$, if ϕ is v
2. ϕ , if ϕ is of the form $\lambda v : T(\zeta)$, for some T and ζ (i.e. don’t do any substitution if v is bound within ϕ)

⁶This record corresponds to the interpretation of *it appears that a dog is approaching*.

3. $\lambda u : T(\text{Subst}(\alpha, v, \zeta, \pi))$, if ϕ is of the form $\lambda u : T(\zeta)$ and u is not v .
4. $\left[\begin{array}{ll} \ell_1 & : \text{Subst}(\alpha, v, T_1, \pi.\ell_1) \\ \dots & \\ \ell_n & : \text{Subst}(\alpha, v, T_n, \pi.\ell_n) \end{array} \right]$, if ϕ is $\left[\begin{array}{ll} \ell_1 & : T_1 \\ \dots & \\ \ell_n & : T_n \end{array} \right]$
5. $P(\text{Subst}(\alpha, v, \beta_1, \pi), \dots, \text{Subst}(\alpha, v, \beta_n, \pi))$, if α is $P(\beta_1, \dots, \beta_n)$ for some predicate P
6. ϕ otherwise

$\text{extend-paths}(\alpha, \pi)$ is

1. $\langle f, \langle \pi.\pi_1, \dots, \pi.\pi_n \rangle \rangle$, if α is $\langle f, \langle \pi_1, \dots, \pi_n \rangle \rangle$
2. $\left[\begin{array}{ll} \ell_1 & : \text{extend-paths}(T_1, \pi) \\ \dots & \\ \ell_n & : \text{extend-paths}(T_n, \pi) \end{array} \right]$ if α is $\left[\begin{array}{ll} \ell_1 & : T_1 \\ \dots & \\ \ell_n & : T_n \end{array} \right]$
3. $P(\text{extend-paths}(\beta_1, \pi), \dots, \text{extend-paths}(\beta_n, \pi))$, if α is $P(\beta_1, \dots, \beta_n)$ for some predicate P
4. α , otherwise

A.12 Merges of record types

If T_1 and T_2 are record types then there will always be a record type (not a meet) T_3 which is necessarily equivalent to $T_1 \wedge T_2$. Let us consider some examples:

$$[f:T_1] \wedge [g:T_2] \approx \left[\begin{array}{l} f:T_1 \\ g:T_2 \end{array} \right]$$

$$[f:T_1] \wedge [f:T_2] \approx [f:T_1 \wedge T_2]$$

We define a function μ which maps meets of record types to an equivalent record type, record types to equivalent types where meets in their values have been simplified by μ and any other types to themselves:

1. If for some $T_1, T_2, T = T_1 \wedge T_2$ then $\mu(T) = \mu'(\mu(T_1) \wedge \mu(T_2))$.
2. If T is a record type then $\mu(T)$ is T' such that for any $\ell, v, \langle \ell, \mu(v) \rangle \in T'$ iff $\langle \ell, v \rangle \in T$.
3. Otherwise $\mu(T) = T$.

$\mu'(T_1 \wedge T_2)$ is defined by:

1. if T_1 and T_2 are record types, then $\mu'(T_1 \wedge T_2) = T_3$ such that
 - (a) for any ℓ, v_1, v_2 , if $\langle \ell, v_1 \rangle \in T_1$ and $\langle \ell, v_2 \rangle \in T_2$, then
 - i. if v_1 and v_2 are $\langle \lambda u_1 : T'_1 \dots \lambda u_i : T'_i(\phi), \langle \pi_1 \dots \pi_i \rangle \rangle$ and $\langle \lambda u'_1 : T''_1 \dots \lambda u'_k : T''_k(\psi), \langle \pi'_1 \dots \pi'_k \rangle \rangle$ respectively, then $\langle \lambda u_1 : T'_1 \dots \lambda u_i : T'_i, \lambda u'_1 : T''_1 \dots \lambda u'_k : T''_k(\mu(\phi \wedge \psi)), \langle \pi_1 \dots \pi_i, \pi'_1 \dots \pi'_k \rangle \rangle \in T_3$
 - ii. if v_1 is $\langle \lambda u_1 : T'_1 \dots \lambda u_i : T'_i(\phi), \langle \pi_1 \dots \pi_i \rangle \rangle$ and v_2 is a type (i.e. not of the form $\langle f, \Pi \rangle$ for some function f and sequence of paths Π), then $\langle \lambda u_1 : T'_1 \dots \lambda u_i : T'_i(\mu(\phi \wedge v_2)), \langle \pi_1 \dots \pi_i \rangle \rangle \in T_3$
 - iii. if v_2 is $\langle \lambda u'_1 : T''_1 \dots \lambda u'_k : T''_k(\psi), \langle \pi'_1 \dots \pi'_k \rangle \rangle$ and v_1 is a type, then $\langle \lambda u'_1 : T''_1 \dots \lambda u'_k : T''_k(\mu(v_1 \wedge \psi)), \langle \pi'_1 \dots \pi'_k \rangle \rangle \in T_3$
 - iv. otherwise $\langle \ell, \mu(v_1 \wedge v_2) \rangle \in T_3$
 - (b) for any ℓ, v_1 , if $\langle \ell, v_1 \rangle \in T_1$ and there is no v_2 such that $\langle \ell, v_2 \rangle \in T_2$, then $\langle \ell, v_1 \rangle \in T_3$
 - (c) for any ℓ, v_2 , if $\langle \ell, v_2 \rangle \in T_2$ and there is no v_1 such that $\langle \ell, v_1 \rangle \in T_1$, then $\langle \ell, v_2 \rangle \in T_3$
2. Otherwise $\mu'(T_1 \wedge T_2) = T_1 \wedge T_2$

$T_1 \wedge T_2$ is used to represent $\mu(T_1 \wedge T_2)$. We call $T_1 \wedge T_2$ the *merge* of T_1 and T_2 .

The following two clauses could be added at the beginning of the definition of μ (after providing a characterization of the subtype relation, \sqsubseteq).

1. if for some $T_1, T_2, T = T_1 \wedge T_2$ and $T_1 \sqsubseteq T_2$ then $\mu(T) = T_1$
2. if for some $T_1, T_2, T = T_1 \wedge T_2$ and $T_2 \sqsubseteq T_1$ then $\mu(T) = T_2$

The current first clause would then hold in case neither of the conditions of these two clauses are met. The definition without these additional clauses only accounts for simplification of meets which have to do with merges of record types whereas the definition with the additional clauses would in addition have the effect, for example, that $\mu(T \wedge T_a) = T_a$ and $\mu(T_1 \wedge (T_1 \vee T_2)) = T_1$ (provided that we have an appropriate definition of \sqsubseteq) whereas the current definition without the additional clauses means that μ leaves these types unchanged.

We define also a notion of *asymmetric merge* of T_1 and T_2 which is defined by a function exactly like μ except that clause 2 of the definition of μ' is replaced by

- 2'. Otherwise $\mu'(T_1 \wedge T_2) = T_2$

We use $T_1 \boxed{\wedge} T_2$ to represent the asymmetric merge of T_1 and T_2 .

These definitions do not in general avoid the formation of ill-formed record types since they allow record types to be replaced with non-record types within a record type thus potentially removing paths that might be included in dependent type fields elsewhere in the resulting type. However, if merging is restricted to either two record types or two non-record types this problem should not occur since all paths from both types will be preserved. In the case of asymmetric merges we can allow the replacement of non-record types by record types without risk. Note that our definition of dependent record types in A.11 allows for dependencies to fields that have conflicting types. Such record types will be well-formed though will not have any witnesses.

Bibliography

- Artstein, Ron, Mark Core, David DeVault, Kallirroi Georgila, Elsi Kaiser and Amanda Stent, eds. (2011) *SemDial 2011* (Los Angeles): Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue.
- Barwise, Jon (1989) *The Situation in Logic*, CSLI Publications, Stanford.
- Barwise, Jon and John Perry (1983) *Situations and Attitudes*, Bradford Books, MIT Press, Cambridge, Mass.
- Blackburn, Patrick, Maarten de Rijke and Yde Venema (2001) *Modal logic* (*Cambridge Tracts in Theoretical Computer Science* 53), Cambridge University Press.
- Breitholtz, Ellen (2010) Clarification Requests as Enthymeme Elicitors, in *Aspects of Semantics and Pragmatics of Dialogue. SemDial 2010, 14th Workshop on the Semantics and Pragmatics of Dialogue* ,.
- Breitholtz, Ellen and Robin Cooper (2011) Enthymemes as Rhetorical Resources, in Artstein *et al.* (2011).
- Breitholtz, Ellen and Jessica Villing (2008) Can Aristotelian Enthymemes Decrease the Cognitive Load of a Dialogue System User?, in *Proceedings of LonDial 2008, the 12th SEMDIAL workshop*.
- Bullock, Barbara E. and Almeida Jacqueline Toribio, eds. (2009) *The Cambridge Handbook of Linguistic Code-Switching*, Cambridge University Press.
- Chierchia, Gennaro and Raymond Turner (1988) Semantics and property theory, *Linguistics and Philosophy*, Vol. 11, No. 3, pp. 261–302.
- Cooper, Robin (2005) Records and Record Types in Semantic Theory, *Journal of Logic and Computation*, Vol. 15, No. 2, pp. 99–112.
- Cooper, Robin (2012) Type Theory and Semantics in Flux, in R. Kempson, N. Asher and T. Fernando (eds.), *Handbook of the Philosophy of Science*, Vol. 14: Philosophy of Linguistics, pp. 271–323, Elsevier BV. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

- Cooper, Robin (fthc) Type Theory and Semantics in Flux, in R. Kempson, N. Asher and T. Fernando (eds.), *Handbook of the Philosophy of Science*, Vol. 14: Philosophy of Linguistics, Elsevier BV. General editors: Dov M. Gabbay, Paul Thagard and John Woods.
- Cooper, Robin and Jonathan Ginzburg (2011a) Negation in Dialogue, in Artstein *et al.* (2011), pp. 130–139.
- Cooper, Robin and Jonathan Ginzburg (2011b) Negative inquisitiveness and alternatives-based negation, in *Proceedings of the Amsterdam Colloquium, 2011*.
- Davidson, Donald (1967) The Logical Form of Action Sentences, in N. Rescher (ed.), *The Logic of Decision and Action*, University of Pittsburgh Press. Reprinted in Davidson (1980).
- Davidson, Donald (1980) *Essays on Actions and Events*, Oxford University Press, New edition 2001.
- Dowty, David (1989) On the semantic content of the notion of ‘Thematic Role’, in G. Chierchia, B. H. Partee and R. Turner (eds.), *Properties, Types and Meanings*, Vol. II: Semantic Issues, pp. 69–130, Kluwer, Dordrecht.
- Dowty, David, Robert Wall and Stanley Peters (1981) *Introduction to Montague Semantics*, Reidel (Springer).
- Fernández, Raquel (2006) *Non-sentential utterances in dialogue: Classification, Resolution and Use*, PhD dissertation, King’s College, London.
- Fernando, Tim (2004) A finite-state approach to events in natural language semantics, *Journal of Logic and Computation*, Vol. 14, No. 1, pp. 79–92.
- Fernando, Tim (2006) Situations as Strings, *Electronic Notes in Theoretical Computer Science*, Vol. 165, pp. 23–36.
- Fernando, Tim (2008) Finite-state descriptions for temporal semantics, in H. Bunt and R. Muskens (eds.), *Computing Meaning, Volume 3 (Studies in Linguistics and Philosophy 83)*, pp. 347–368, Springer.
- Fernando, Tim (2009) Situations in LTL as strings, *Information and Computation*, Vol. 207, No. 10, pp. 980–999.
- Fillmore, Charles J. (1982) Frame semantics, in *Linguistics in the Morning Calm*, pp. 111–137, Hanshin Publishing Co., Seoul.
- Fillmore, Charles J. (1985) Frames and the semantics of understanding, *Quaderni di Semantica*, Vol. 6, No. 2, pp. 222–254.
- Fox, Chris and Shalom Lappin (2005) *Foundations of Intensional Semantics*, Blackwell Publishing.

- Gibson, James J. (1986) *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates.
- Ginzburg, Jonathan (2010) Relevance for Dialogue, in Łupkowski and Purver (2010), pp. 121–129, Polish Society for Cognitive Science.
- Ginzburg, Jonathan (2012) *The Interactive Stance: Meaning for Conversation*, Oxford University Press, Oxford.
- Ginzburg, Jonathan (fthc) *The Interactive Stance: Meaning for Conversation*, Oxford University Press, Oxford.
- Ginzburg, Jonathan and Raquel Fernández (2010) Computational Models of Dialogue, in A. Clark, C. Fox and S. Lappin (eds.), *The Handbook of Computational Linguistics and Natural Language Processing*, Wiley-Blackwell.
- Hanks, Peter W. (2012) First-Person Propositions, *Philosophy and Phenomenological Research*, Vol. 84, No. 2, p. tba.
- Hughes, G.E. and M.J. Cresswell (1968) *Introduction to modal logic*, Methuen and Co., Ltd.
- Kamp, Hans and Uwe Reyle (1993) *From Discourse to Logic*, Kluwer, Dordrecht.
- Kant, Immanuel (1781) *Critik der reinen Vernunft (Critique of Pure Reason)*, Johann Friedrich Hartknoch, Riga, second edition 1787.
- Larsson, Staffan (2002) *Issue-based Dialogue Management*, PhD dissertation, University of Gothenburg.
- Larsson, Staffan (2011) The TTR perceptron: Dynamic perceptual meanings and semantic coordination., in Artstein *et al.* (2011).
- Larsson, Staffan and David R. Traum (2001) Information state and dialogue management in the TRINDI dialogue move engine toolkit, *Natural Language Engineering*, Vol. 6, No. 3&4, pp. 323–340.
- Łupkowski, Paweł and Matthew Purver, eds. (2010) Aspects of Semantics and Pragmatics of Dialogue. SemDial 2010, 14th Workshop on the Semantics and Pragmatics of Dialogue. Poznań: Polish Society for Cognitive Science.
- Martin-Löf, Per (1984) *Intuitionistic Type Theory*, Bibliopolis, Naples.
- McCarthy, J. and P. J. Hayes (1969) Some philosophical problems from the standpoint of artificial intelligence, *Machine Intelligence*, Vol. 4, pp. 463–502.
- Montague, Richard (1974) *Formal Philosophy: Selected Papers of Richard Montague*, Yale University Press, New Haven, ed. and with an introduction by Richmond H. Thomason.

- Ninan, Dilip (2010) *De Se* Attitudes: Ascription and Communication, *Philosophy Compass*, Vol. 5, No. 7, pp. 551–567.
- Nordström, Bengt, Kent Petersson and Jan M. Smith (1990) *Programming in Martin-Löf's Type Theory* (*International Series of Monographs on Computer Science* 7), Clarendon Press, Oxford.
- Prior, Arthur N. (1957) *Time and modality*, Oxford University Press.
- Prior, Arthur N. (1967) *Past, present and future*, Oxford University Press.
- Prior, Arthur N. (2003) *Papers on Time and Tense*, Oxford University Press, new edition, edited by Per Hasle, Peter Øhrstrøm, Torbren Braüner and Jack Copeland.
- Purver, Matthew, Eleni Gregoromichelaki, Wilfried Meyer-Viol and Ronnie Cann (2010) Splitting the *Is* and Crossing the *Yous*: Context, Speech Acts and Grammar, in Łupkowski and Purver (2010), pp. 43–50, Polish Society for Cognitive Science.
- Pustejovsky, James (1995) *The Generative Lexicon*, MIT Press, Cambridge, Mass.
- Pustejovsky, James (2006) Type theory and lexical decomposition, *Journal of Cognitive Science*, Vol. 6, pp. 39–76.
- Ranta, Aarne (1994) *Type-Theoretical Grammar*, Clarendon Press, Oxford.
- Reichenbach, Hans (1947) *Elements of Symbolic Logic*, University of California Press.
- Ruppenhofer, Josef, Michael Ellsworth, Miriam R.L. Petruck, Christopher R. Johnson and Jan Scheffczyk (2006) *FrameNet II: Extended Theory and Practice*. Available from the FrameNet website.
- Russell, Bertrand (1903) *Principles of Mathematics*, Cambridge University Press.
- Sag, Ivan A., Thomas Wasow and Emily M. Bender (2003) *Syntactic Theory: A Formal Introduction*, 2nd edition, CSLI Publications, Stanford.
- de Saussure, Ferdinand (1916) *Cours de linguistique générale*, Payot, Lausanne and Paris, edited by Charles Bally and Albert Séchehaye.
- Searle, John R. (1969) *Speech Acts: an Essay in the Philosophy of Language*, Cambridge University Press.
- Seligman, Jerry and Larry Moss (1997) Situation Theory, in J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*, North Holland and MIT Press.
- Shanahan, Murray (2009) The Frame Problem, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2009 edition, <http://plato.stanford.edu/archives/win2009/entries/frame-problem/>.

- Shieber, Stuart (1986) *An Introduction to Unification-Based Approaches to Grammar*, CSLI Publications, Stanford.
- Thomason, Richmond (1980) A model theory for propositional attitudes, *Linguistics and Philosophy*, Vol. 4, pp. 47–70.
- Traum, David R. (1994) *A Computational Theory of Grounding in Natural Language Conversation*, PhD dissertation, University of Rochester, Department of Computer Science.
- Turner, Raymond (2005) Semantics and Stratification, *Journal of Logic and Computation*, Vol. 15, No. 2, pp. 145–158.