Abstract

We argue that the mathematics developed for the semantics of computer languages can be fruitfully applied to problems in human communication and action.

Keywords: reasoning about action, identity of action, category theory

1. INTRODUCTION

Let us start with the philosophy of language. There is a well known programme, due to Davidson [7, 12], the goal of which is to give an axiomatic semantic theory from which can be derived the observable facts of human language use. This programme is deliberately abstract, in that it ignores any evidence from neurophysiology or, indeed, any questions of mechanism: it starts from the overt, publicly observable facts about behaviour in the hope that those, together with the constraint of being able to construct an axiomatic theory explaining them, will yield an interesting semantics of human language. This programme has been extraordinarily influential: in particular, its analysis of action [5] has influenced a great deal of work in linguistics (see, for example, [15]).

We argue, in [23], that Davidson’s programme is strikingly similar to the theory of the semantics of programming languages: both are concerned with the construction of theories which account for the overt facts of their languages – on the one hand, the facts of human linguistic behaviour, and, on the other hand, the observable behaviour of computing devices – and both deliberately ignore facts about implementation. There are, in fact, further similarities between these areas, and these will become evident in the course of this paper.

There is, however, one striking difference: the philosophers’ theories are much less sophisticated than the computer scientists’. The semantics of programming languages is mathematically sophisticated because it has to be: it must guide the design of programming languages, and it must support reasoning about programmes, and it thereby has access to a very large amount of complex data. Furthermore, the data is relatively unproblematic, and whether an analysis succeeds or fails is likewise relatively unproblematic. None of this is holds for the activity of philosophers or linguists. Consequently, whereas the computer scientists have a wide range of powerful mathematical techniques available – domain theory, for example, and the theory of fibred categories – there is relatively little available on the philosophical side. The philosopher’s formal theories are quite fragmentary, and most of it is done, using naïve set theory, without much thought for what the real-world semantics of that set theory might be.

We will argue, then, that there is a great deal to be gained by applying the methods of programming language semantics to the analysis of natural language. What will emerge is an account of contexts for human communication based on the theory of fibred categories: in the important special case of language about action, we get a theory of contexts for reasoning about action.

This is not merely of philosophical interest: these issues have important applications. With the advent of mobile devices, the context of use has become much more important than it used to be for traditional desktop machines. Such contexts are largely given by the situation of the user in the real world, and so a theory such as ours, which is concerned with these issues, becomes
important. Similarly, empirical studies [9] show that context is important in social networking software, and so we may well want to reproduce, in the virtual world, those aspects of the context of human communication which are important in the real world. And, finally, there are applications in the semantic web: qualitative reasoning about the human world has become a central topic there.

The strategy of this paper, then, will be as follows: it will be a case study, showing how the methods of fibred category theory can be applied to reasoning about action. We will start with a treatment of reasoning about action due to Reiter, and formalise the notion of context involved in that treatment. This will, surprisingly, involve equality between the results of actions: this allows us to revisit some of Davidson's fundamental work on actions, and give a new formalisation of it. Our treatment – by comparison with Davidson's – will not involve questionable set-theoretic assumptions, but will specify, very concretely, what the appropriate contexts for reasoning about action are: they are tuples of equations between the outcomes of actions. Finally, we will show how these results can be implemented in the semantic web, and we will conclude with some remarks about how this approach can be extended.

2. REASONING ABOUT ACTION

2.1. Reiter

Reiter [19], [17] works within an artificial intelligence formalism known as the situation calculus: the basic notions of this formalism are situations and fluents. Situations can be thought of as states of the world in which actions are performed: fluents are propositions which can be true or false in those situations. In order to use the situation calculus, then, we have to say what the actions do. That is, we have to say, for each fluent, how it is affected by actions. Reiter handles this with what he calls regression [16]: given an action \( s \xrightarrow{\alpha} t \) between situations \( s \) and \( t \), and given a fluent \( P \) at \( t \) the regression problem is to find a fluent \( P' \) at \( s \) which will be true iff \( P \) is true at \( t \).

One might think that another notion would be more fundamental: after all, the traditional AI scenario is that one knows about the present and wants to predict the future. This concept is known as progression: that is, given an action \( s \xrightarrow{\alpha} t \), and given a theory \( \Xi \) describing the situation \( s \), find the theory \( \Xi' \) describing \( t \). Reiter does, in fact, define this notion: however, his insight was that regression is technically simpler to work with, and that progression can be defined in terms of it [18].

There is a straightforward category-theoretic interpretation of regression: it is simply pullback in a fibred category. Given an action \( \alpha: s \to t \), and a fluent \( Q \) in the situation \( t \), its regression to the situation \( s \) will be the pullback \( \alpha^*Q \). So we construct a fibred category, the objects of whose base are – let us say – situations, the morphisms between those objects are actions, and whose fibres are boolean algebras of the fluents in those situations. Progression, according to Reiter's definition, will be a left adjoint to pullbacks by actions.

There is a logic associated to this fibred category: it is typed by the objects of the base category, and the propositions with a given type will be the fluents in that situation. That is, given a fluent \( P \) in a situation \( t \), we can write it, in a typed idiom, \( P: t \); actions will correspond to typed functions, and pullback will correspond to substitution. Left adjoints to pullbacks will correspond to existential quantification [11]. We also need more structure in the base. If we want to reason about actions, we want to be able to compare situations – to compare, for example, the situation in which I get the bus to college from the one in which I walk. So that naturally leads to reasoning about tuples, i.e. cartesian products, of situations. And similarly, we will want to compare the extent to which two actions are similar or not, which leads to the idea of equalisers.

So we can construct this logic, prove cut elimination (which yields proof search algorithms), and prove soundness and completeness with respect to a suitable semantics in terms of fibred
categories: details are given in [24]. This system solves several problems which seem difficult, or impossible, to solve in Reiter’s original system (see, in particular, [24, Section 4]). The problem seems to be that Reiter uses an untyped second-order system in which situations are individuals: so, when he wants to say that, for example, a proposition is about a particular situation – something which can be expressed naturally in our typed system – he cannot put the untyped genie back into the bottle.

2.2. An Explicit Construction

This is all very well, and is, in its own right, a usable system. However, it is still rather abstract, and one could very well argue that, although it uses contexts, it does not really tell us what they are. So we will now remedy that defect. First a well-known result:

**Remark 1.** If we start with an arbitrary category $\mathcal{C}$, and add finite products and equalisers, we form the free category with finite limits over $\mathcal{C}$.

So, the process that we describe on p. 2 – of starting with a category of situations and actions, and then adding cartesian products and equalisers – is, in fact, the free finite limit category over the original category of situations and actions. So we need to find a more explicit representation of this construction.

There are numerous constructions of free finite limit categories in the literature (for example, [3]). We want to be able to calculate with our system, and this corresponds to working in the internal language of these categories: a straightforward treatment of this internal language leads to a syntax characterised, in Peter Freyd’s words, by “remarkable inelegance” [8, p. 6]. So we need to proceed less directly. We use a construction due to Palmgren and Vickers [14]; they define a system which gives a tractable “internal language” for finite limit categories. (Their construction does, in fact, do rather more than they claim, and we will come to that later).

Suppose, then, that we are given a category $\mathcal{C}_0$: we want to construct the free cartesian category on it. We first define a typed logic of equalities between partially defined terms: we use a typed equality symbol, $\equiv$, between terms of the same type. Informally, $\alpha(x) \equiv \beta(x)$ will be true for those $x$ for which the terms $\alpha$ and $\beta$ are both defined, and their values are equal. We abbreviate $\alpha \equiv \alpha$ to $\alpha \downarrow \alpha \downarrow \{	heta(x)\}$ will be true for those $x$ for which $\alpha$ is defined.

**Definition 1** (Palmgren and Vickers). Given a category $\mathcal{C}_0$, define a typed logic as follows:

- **types** are tuples of objects of $\mathcal{C}_0$
- **variables** for each type, we suppose that there is a stock of variables of that type
- **functions** between types are given by tuples of function symbols corresponding to the morphisms of $\mathcal{C}_0$: the free variables of these functions will have the types of the sources of the morphisms of $\mathcal{C}_0$, and the result types of the functions will have types corresponding to the tuples of objects of $\mathcal{C}_0$.
- **predicates** for each type, there is one binary predicate symbol, $\cdot \equiv \cdot$, between terms of that type
- **axioms** are given by the rules in Table 1, where $\theta(x)$ is a conjunction of partial equations with free variables in the tuple $x$.

**Example 1.**

1. Given an object $A$ of $\mathcal{C}_0$, then the tuple $(\text{Id}_A(x : A), \text{Id}_A(x : A))$ is a function from $A$ to $(A, A)$, known as the diagonal, $\Delta_A$
2. Given objects $A$ and $B$ of $\mathcal{C}_0$, then the twist map on $A, B$ is the function $(\text{Id}_B(y : B), \text{Id}_A(x : A))$.

**Theorem 1** (Palmgren and Vickers). For a given category, construct the category $\mathcal{C}$ whose objects are of the form

$$\{x : S|\theta(x)\},$$

where $x$ is a tuple of variables of type $S$, and $\theta(x)$ is a conjunction of partial equations of that type, and whose morphisms

$$\{x : S|\zeta(x)\} \rightarrow \{y : T|\eta(y)\}$$
are given by equivalence classes under equality of terms $\alpha(x : S) : T$ such that 

$$\Theta, \zeta(x) \vdash \alpha(x) \downarrow \land \eta(\alpha(x))$$

where $\Theta$ is a collection of axioms expressing the convergence of all morphisms of $C_0$. Then $C$ is the free category with finite limits on $C_0$.

### 2.2.1. Informal Semantics

We can give the above construction an informal semantics as follows. Reasoning about action has two sides, which we will, following philosophical terminology, call the intensional and the extensional. The intensional side is the agent’s view of actions: what actions are performed, in what sequence, and so on. It is this view of actions which is sometimes referred to as the “knowledge level” [13].

However, as well as their intensional aspect, actions also have an effect on the world. This is the extensional side of action and it will be important also to talk about it. We will represent the extensional side of actions by propositional assertions about states. If, like Reiter, we use classical logic, “the way the world is” can be described by assigning truth values to propositions: that is, by what is called, in logical jargon, a possible world, and we can, therefore, think of the effect of an action as a function from possible worlds to possible worlds.

Extensions and intensions will be related as follows. States encode intensional information, and such information will, in general, only yield partial knowledge of the world: thus, each (intensional) state will, in general, correspond to several different possible worlds. However, the agent’s epistemic state will be part of the world, so that each possible world will correspond to a unique such state. So, each state will have, associated to it, a set of possible worlds, and these sets of possible worlds will be disjoint. We thus get a well-defined functor from extensional possible worlds to intensional states: in our fibrational view, this will correspond to the projection functor of the fibration.

Our encoding of the category with finite limits in terms of partial equations can be extended to fibrations over such categories, and, in these terms, it goes as follows. Objects of the original category correspond, as we have seen, to sets of possible worlds: the variables of the equations are to be understood as ranging over those possible worlds. An object of the finite limit category is of the form

$$\{x | \vartheta(x)\}$$
and will thus correspond to a set of tuples of possible worlds, subject to the constraints given by
the equations $\vartheta(x)$. And the morphisms of the finite limit category will be tuples of morphisms of
the original category, and these morphisms will send possible worlds to possible worlds (they are,
after all, actions, and that is what actions do: they change the world).

**Example 2.** The philosopher Donald Davidson (following Austin [1]) considers the following
pattern of reasoning.

‘I didn’t know that it was loaded’ belongs to one standard pattern of excuse. I do not deny that I pointed
the gun and pulled the trigger, nor that I shot the victim. My ignorance explains how it happens that I
pointed the gun and pulled the trigger intentionally, but did not shoot the victim intentionally. [6, p. 109]

The misfortune in Davidson’s scenario stems directly from the gap between the intensional and
the extensional. That is, the agent had performed an action which, from his point of view (that is,
intensionally) was merely the action of pulling the trigger; however, extensionally, and unknown to
the agent, this action turned out to be identical with that of firing the gun.

So, from this point of view, being in a particular context for reasoning about action amounts to
this: it amounts to knowing that, in certain situations, particular pairs of actions have the same
effect on the world. In Davidson’s scenario, the agent does not know that pulling the trigger has
the same effect as shooting the gun. In our system, we can encode this piece of knowledge in
terms of a partial equality: if $f$ encodes pulling the trigger, and $g$ encodes shooting the gun, then
$f(x) \equiv g(x)$ is a proposition which is true of some worlds but false of others.

This is fairly remarkable. We started from a sequent calculus formalisation of a typed version
of Reiter’s system, and, for the sake of the reasoning we wanted to perform, we used a base
category with finite limits. And we then used a concrete presentation of finite limit categories, and
translated it back into our terms: we found that it had a very natural translation in terms of contexts
for reasoning about action. Now there are reasons for this coincidence, but we need a little more
formalism before we can say what they are.

We should note also that results like this need both an abstract and a concrete side. Abstractly,
we need to know that the base category is a category with finite limits; we need to know this in
order to write down the sequent calculus for this logic. Concretely, however, we need a concrete
presentation of these abstractly defined objects: and we need this in order to give an illuminating,
common-sense reading of the calculus.

### 2.3. Partiality

There is still an aspect of Reiter’s work that we have not used. A given action will only be possible
in certain circumstances, and a treatment of reasoning about action ought to be able to say when
an action will succeed. This will handled by our semantics of actions as partial functions: an action
will only succeed in the possible worlds in which the corresponding function is defined. So we can
naturally extend our concrete presentation to this case: we use Palmgren and Vickers’ system of
partial equational logic, and merely relax the totality condition on the base category.

**Definition 2.** Given a category of situations and actions $\mathcal{C}_0$, construct the category $\mathcal{C}$ whose
objects are of the form

$$\{x : S | \vartheta(x)\},$$

where $x$ is a tuple of variables of type $S$, and $\vartheta(x)$ is a conjunction of partial equations of that
type, and whose morphisms

$$\{x : S \mid \zeta(x)\} \rightarrow \{y : T \mid \eta(y)\}$$

are written $\langle \alpha(x : S) | \vartheta(x) \rangle$. This morphism is define where $\alpha$ is defined, and where, in addition, $\vartheta$
is true.

We extend the partial equality $\equiv$ to these morphisms by

$$\langle \alpha | \vartheta \rangle \equiv \langle \alpha' | \vartheta' \rangle := \alpha \equiv \alpha' \land \vartheta \land \vartheta'$$
There is a partial ordering on the sets of morphisms with given source and target: if we have \( \langle \alpha | \vartheta \rangle \sqsubseteq \langle \alpha' | \vartheta' \rangle \) iff, whenever the first is defined, the second is also defined, and their values are equal: that is, iff

\[ \langle \alpha | \vartheta \rangle \downarrow \vdash \langle \alpha | \vartheta \rangle \equiv \langle \alpha' | \vartheta' \rangle \]

This extends our informal interpretation to the case of partial maps: we now want, then, an abstract structure which this definition instantiates. Note that what we have defined is a bicategory: the homsets are partially ordered sets, rather than simply sets. The required structure has been defined by Carboni [4] (see also [20]).

**Definition 3** (Carboni). A bicategory of partial maps is a locally posetal bicategory together with a monoidal structure \( \otimes, I \), together with, for each object \( A \), 1-cells \( \Delta_A : A \rightarrow A \otimes A \), \( \nabla_A : A \otimes A \rightarrow A \), and \( !_A : A \rightarrow I \), such that:

1. For each \( A \), \( \langle \Delta_A, !_A \rangle \) makes \( A \) into a coalgebra, and this is the unique coalgebra structure on \( A \)
2. For any \( f : A \rightarrow B \), we have

\[ \Delta_B f = (f \otimes f) \Delta_A \quad \quad \text{and} \quad \quad !_B f \sqsubseteq !_A \]

3. For any \( A \), \( \Delta_A \vdash \nabla_A \),
4. For any \( f, g : A \rightarrow B \), \( \nabla_B (f \otimes g) \Delta_B \sqsubseteq f \), and
5. For any \( A \), \( \Delta_A \nabla_A = (\text{id}_A \otimes \nabla_A)(\Delta_A \otimes \text{id}_A) \)

We say that a 1-cell \( f : A \rightarrow B \) of such a bicategory is total if \( !_B f = !_A \).

Carboni also proves that the subcategory of total 1-cells has trivial bicategorical structure, and that it is a category with finite limits.

We can show that what our construction defines is a bicategory of partial maps:

**Theorem 2.** The construction of Definition 2 defines the free bicategory of partial maps over \( C_0 \).

In order to talk about predicates, we now need to define fibrations over such a bicategory. We need first to settle what the pullback of a predicate along a partially defined morphism should be: there are two choices, either to prolong the pullback by \( \top \) where the morphism is not defined, or to prolong it by \( \bot \). The former gives pullbacks that preserve \( \top \), the latter gives pullbacks that preserve \( \bot \). Since we want our pullbacks to have left adjoints, we will choose to prolong pullbacks by \( \top \). This in turn means that, if we have 1-cells \( f \) and \( g \), and a predicate \( P \) over their target, then

\[ f \sqsubseteq g \quad \Rightarrow \quad g^* P \vdash f^* P \quad (1) \]

Abstractly, the definition of a fibration over a bicategory of partial maps can be adapted from Hermida [10], and offers few surprises: we want the obvious naturality conditions, together with (1), and we want the projections of homsets to be fibrations of posets.

**Theorem 3.** We can show the following:

1. Using Hermida’s definition of 2-fibrations over bicategories [10], there is a natural equivalence between 2-fibrations over a bicategory of partial maps and fibrations over the associated finite limit category
2. A fibration over the finite limit category has boolean algebras for fibres, and boolean algebra morphisms for pullbacks, if the 2-fibration over the cartesian bicategory has boolean algebras for fibres and \( \wedge, \top, \vee \)-preserving morphisms for pullbacks, and if, in addition, \( !_A \bot = \bot \) for any \( A \).
3. 2-fibrations over a cartesian bicategory have left adjoints to the pullbacks iff the corresponding fibrations over the associated finite limit category have.

We can define a sequent calculus for categories 2-fibred in boolean algebras over cartesian bicategories: the rules for the classical connectives are unremarkable, and the rules for the
non-classical ones (i.e. pullbacks and their left adjoints) are given in Table 2. We can prove soundness, completeness, and (with somewhat more complex but equivalent rules) we can prove cut elimination for this calculus.

**Theorem 4.** We can prove soundness and completeness for our sequent calculus: the semantics is 2-fibrations in boolean algebras over bicategories of partial maps, where the pullbacks have left adjoints and where we have $!^* \bot = \bot$, for any object $A$.

**Theorem 5.** We can prove cut elimination for our calculus.

These results, although technical, give us the abstract structures for reasoning about partially defined actions: they also show that, when the actions are total, our fibrations reduce to the ones defined in the previous section. They also show that the contexts and morphisms constructed out of the Palmgren-Vickers logic play a privileged role in these structures, too: that is, they are the free cartesian bicategories. So here, too, we can think of the natural contexts for reasoning about action as being given by sets of partial equations.

But more than this is true: we can explain the privileged role of equality.

**Definition 4.** Given a 2-fibration in Boolean algebras over a cartesian bicategory, and two 1-cells $f, g : A \to B$ in the base, define the an element of the fibre over $A$ by

$$f \simeq g := \neg((\nabla_B(f \otimes g)\Delta A)^* \bot)$$

where $\bot$ is the bottom element of the fibre over $B$.

Note that we have here defined a proposition $f \simeq g$ in our system, rather then (as previously) using a partial equality between terms to construct models for our system.

**Theorem 6.** $f \simeq g$ allows substitution: that is, the following rule is admissible.

$$\frac{A[\Gamma \vdash \Delta]}{A[\Gamma[f^*/g^*] \vdash \Delta[f^*/g^*]]}$$

So equalities play a prominent role in concrete implementations of this logic because they come from the abstract structure of the base: it is, then, no coincidence that equality plays such a prominent role in the Palmgren-Vickers models.

**3. DAVIDSON**

Donald Davidson analyses his example as follows:

'I didn’t know that it was loaded’ belongs to one standard pattern of excuse. I do not deny that I pointed the gun and pulled the trigger, nor that I shot the victim. My ignorance explains how it happens that I pointed the gun and pulled the trigger intentionally, but did not shoot the victim intentionally. . . . The logic of this sort of excuse includes, it seems, at least this much structure: I am accused of doing $b$, which is deplorable. I admit I did $a$, which is excusable. My excuse for doing $b$ rests upon my claim that I did not know that $a = b$. [6, p. 109]

Davidson, then, is arguing for two things:

1. equalities between actions are meaningful, and
2. we use these equalities in common-sense reasoning about action.

We can modify this somewhat. Davidson construes actions as individuals (although their ontology remains somewhat mysterious), and equalities between actions that he has in mind are equalities between those individuals: in our terms, actions would be functions, and the equalities would be equalities between functions considered as first-class objects.

However, our system handles the above reasoning very well, and, although equality plays a central role in it, the equalities are those between values of functions rather than between the functions themselves: furthermore, these equalities seem to be definable in any system which expresses the rather weak requirements of Reiter’s theory of action. So it seems that the equalities of our system are the ones that we need, and that, with them, we can handle the sort of reasoning which Davidson has in mind.

4. IMPLEMENTATION

It is worth pointing out that we can implement the reasoning in Davidson’s example using standard tools for the Semantic Web. Owl, the Web ontology language, is based on description logic, which is basically a syntactically sugared multimodal logic: we have translated the concepts there into Owl, and proved the necessary entailments and non-entailments using a standard proof tool for description logic. The files are available at http://www.dcs.qmul.ac.uk/~graham/excuses.

5. FURTHER DEVELOPMENT

The above is, as we have said, only a case study: there are many areas in which it should be developed further. The first is that it does not handle the causal aspect of action at all: that is, the fact that actions are part of the causal fabric of the world, and that they cause, and are caused by, other events. Some treatment of this would be interesting: there seems to be good evidence that causal reasoning and linear logic are quite closely related [22] [21], and it would be natural to extend the above fibrational treatment to some sort of monoidal fibration. Secondly, we have been using classical logic (basically, because Reiter did): we should develop an intuitionistic version, because human reasoning about action uses finite evidence. Preliminary investigations look interesting. Finally, implementation in description logic is interesting, and also suggests extensions to the standard logics of the semantic web (the problem is that description logic has good facilities for restricting objects by logical expressions – in our terms, \( \{ x \mid \vartheta(x) \} \) – but not for restricting 1-cells, which, in our terms, would be \( \langle \alpha(x) \mid \vartheta(x) \rangle \)).

REFERENCES


