A GENERALISED THEORY OF DEFAULT REASONING - PRELIMINARY REMARKS

Richard Ball

Department of Computer Science
University of Essex
August 1992

Intuitive starting points.

This document gives an account of the main ideas and general direction of my ph.D thesis.

We start from the realisation (which I will argue for) that semi-normal defaults (and other types of prototypical reasoning with exceptions) do not fit into a recursively defined bijectional semantics.

The nearest attempt was adaquate as far as it went (Etherington).

This involved a modification of the semantics at the end (ie cycling back through the default rules). A conditional logic trying to give one-to-one semantics to a semi-normal default would have to use some kind of "virtual" model. That is, since there appears to be no way to decide the truth conditions of one semi- or ab-normal default rule in the absense of the knowledge of how the other ones behave, any attempt to represent such truth conditions would necessarily involve some sort of model theoretic device to wait for the outcome of the other rules. I have no intuitions as to what such a structure would represent.

So if a wff has no denotational semantics, what are we left with? We have only a set of sentences in some language for the agent's theory of the world (this set is always finite), and we have a conviction that these sentences are true. Included in this set of sentences we have (perhaps) a number of default rules, whose meaning we do not know until we set it/them in a system of some kind.

We have the starting set of a reasoning system to be given below. The idea is that the underlying theory is simply justified on itself. Details below.

We wish to explore closely what it is for a derivation to be <u>justified</u> on other wffs.

Consider the intuition that a car (eg) is where one left it until someone either moves it or it is blown up . That it is where we left it is justified on the "consistency" (details later) of the assertions that nobody has moved it, blown it up ...
We will not need a temporal logic, at least not for the basic

case.

We do not agree with Shoham that
1. Incoherent default theories (:MA/¬A) etc. are meaningless. They express a procedural paradox of sorts. So it is legitimate to prefer (and want to capture, or at least illuminate) Reiter's formulation.
2. That we should shift our attention to semantics. We are after something general enough to capture the expressiveness of Reiter (as well as its flavour). Other formulations may be expressible in this way.

Inconsistent theories etc. are valuable. We do not want a structure which alows the whole language to be deduced as soon as inconsistency arises. Furthermore, we will pay more attention than usual to the way that we envisage such a system being used in the event of new information. Particularly, we think of an extension as a reasoning out on a limb, a structure yet to be confirmed.

Note that there is an intuitionistic flavour to what follows, but it is the agent's information (justified on itself) which is formulated cumulatively, not the expressiveness of the object language. The hope is by this means to capture the iterative feel of nonmonotonic reasoning.

Details

First an intuitive idea. We decribe an tree structure over a domain of pairs of sets of wffs <j,d> to be understood as justifications and derivations respectively.

Note that such a pair of wffs is to be taken as the line of a proof theoretic structure. The contents of <d> will determine what pairs will follow in the proof, and whether it branches or not.

For ordinary formulae the process described below will resemble a semantic tableau, but with important differences.

There is a notion of an interim, procedural semantics. The "meaning" of a formula is in a certain way given as the way that formula (with its justification to date) modifies the growth of the agent's information, at whatever stage in that growth the formula is considered; given a justification and a derivation so far, there is a generating rule which determines how the tree is built.

The set of formulae in $\langle j \rangle_i$ is cumulative. i.e. always (in a downward direction on the tree whose root is the bottom of the page) $\langle j \rangle_i \subseteq \langle j \rangle_{i-1}$. Though we are getting ahead of ourselves, it will turn out that if we take these successive justification sets to mark successive "cognitive events", because of their cumulative nature it will be possible that once the tree is built (by expansion of the wffs in the agent's theory), that we impose a reflexive and transitive relation over these justifications. We will show that the expanded tree is the same for any order in which the wffs ∈ O (the agent's base set) are considered. The intuition here is that since this compilation might be done in any order (and the construction is completely general in this respect), the agent is not committed to any particular order, rather to the totality of different ways in which this expansion might be done. The intention here is to represent what it is for an agent to have a theory consisting of a set of sentences, all justified by each other.

Wffs are defined recursively in the usual way but we include (or will do eventually) other shapes of formulae including those of the following form:

X:Y/Z

where X,Y and Z are sets of wffs defined as above (though possibly subject to restrictions). We also include a distinguished wff T which is a tautology.

Consider by way of example the following default theory:

W = {a,b v c, d -> e} D = {a:M(b & d)/d}

We translate it (informally at the moment) into the following single theory of the world:

(1) {a,b v c, d -> e, a:b & d/a -> d}

Such sets will be finite and will be subject to some kind of numbering, the need for which will become apparent.

Note that we remove the M. Consistency as we will see is treated in a slightly unexpected way. Notice also the use of the material conditional (or this current system's prrof-theoretic analogue) in the translation of a default rule.

The best way is to watch it work. Here is the tree of the theory (1) above.

1 <j = <T>, d = <a, b v c, d -> e, a:b & d/a -> d>>
2 <<T ∪ a>, <a>>
3.1 <<T ∪ a ∪ b> <b>>
3.2 <<T ∪ a ∪ c> <c>>
4.1.1 <<T ∪ a ∪ b ∪ ¬d> <¬d>>
4.1.2 <<T ∪ a ∪ b ∪ e> <e>
4.2.1 <<T ∪ a ∪ c ∪ ¬d> <¬d>>
4.2.2 <<T ∪ a ∪ c ∪ e> <e>>

(notice that a is justified back to level 1, and remembering the semantic tableau, true in all models (though we won't talk about models for some time yet. Consequently the default is applied to all branches) in general, a default is applicable only on those branches in which the prerequisite is present.

5.1.1.1 <<T ∪ a ∪ b ∪ ¬d ∪ (b & d)> <a -> d>>.

Note that the justification of the default rule goes into the justification set, and the conclusion into the derivation set. etc.

The idea is that this process runs, creating a cumulative justification. Semi-normal defaults are now handled in a natural way. There will come a time when the justification part of the line of the tree is <u>stable</u> . Just as temporal logic may say that a sentence is true at or after a certain time point, there is an idea of a derivation being true at or after such time as the justification is. A default extension is allowed in some way under a trust (by default) that the justification will indeed stabilise at some (hitherto unreached) later point on the agent's cumulative cognitive history.

Notice that on their own, translated default rules ensure that extensions (of the tree) are closed under their conclusions. The translated rule applies to any branch where the prerequisite holds.

We could, if we wished, employ disjunctions of translated default rules i.e. ((a:b & c/a->d) v (a:e & f/a->f)). The v operator behaves as normal, so the default rules now fit on distinct branches, where if they were not disjoined, they would fit above and below each other. It will turn out that the disjunction idea captures Lucaszewicz and the conjunction, Reiter.

I want to show that under certain exact stability conditions, any sequence in such a tree will correspond to a model. On account of the cumulative way in which the structure is built it will be provable that any subset of such a sequence will also have a model. These reults will be established by means of a downward inductive generating system over the tree-structure which will be built. Details later.

Here's what I have so far.

Let X be a set of formulae (justification set). Let A,B ∈ O, where O is a subset of the wffs ∈ L (agents theory, or a subset of that). We may represent proof rules as follows.

1.

```
        (X),(A & B)
               |
       (X ∪ A),(A)
   (X ∪ A ∪ B),(B)
(not entirely happy with this)
```

2.

```
            (X),(A v B)
               __|__
              |     |
(X ∪ A),(A)     (X ∪ B),(B)
```

3.

```
   (X),(A:B & C/A ->C)
            |
 (X ∪ B ∪ C),(A->C)
```

etc.

Define the operational complexity of a wff (a function from L to N where $N \in \{1,2,3,...\}$) written c(wff) = n as follows;

1. c(A) = 1 for A atomic;

2. c(A v B) = (1 + c(A) + c(B)) etc.;

3. c(A:B & C/C) = (1 + c(C)).

Define the result of the application of a wff to any non-empty set of series of pairs S (function from L * S to S') written res(wff,S) = S' as follows;

1. For c(A) = 1 (i.e. for A atomic) , res(A,S) = S' = $\{<X_0,Y_0>,...,<X_i,Y_i>,<X_{i+1},Y_{i+1}>: <X_0,Y_0>,...,<X_i,Y_i> \in S\}$, where $X_{i+1} = X_i \cup A$, and $Y_{i+1} = A$

2. For c(A) = 2;

Case i); A = P v Q (other wffs analagous) , res(A,S) = $\{ <X_0,Y_0>,...,<X_i,Y_i>,<X_{i+1},Y_{i+1}> \cup <X_0,Y_0),...,<X_i,Y_i>,<X'_{i+1},Y'_{i+1}>: <X_0,Y_0>,...,<X_{i,i}Y> \in S\}$, where $X_{i+1} = X_i \cup P$, $Y_{i+1} = P$, $X'_{i+1} = X_i \cup Q$, and $Y'_{i+1} = Q$.

3. For c(A) = n + 1 ($n \geq 2$); Let A have a composition $P \circ Q$ where $\circ$ is the functor of A. Let S_r = res(P,S). Then res(A,S) = res(Q,S_r).

Define a series of sets of series of pairs $S_0 \subseteq S_1 \subseteq S_2$,... as follows;

Let $o_1, o_2, ..., o_n \in O$

S_0 = $\{<T,O>\}$ where T is the distinguished wff being always true, and O is the subset of L being the agent's theory of the world.

S_{i+1} = res(o_i,S_i)

then $$S = \bigcup_{i=0}^{n} S_i$$

Back on an intuitive level, the idea will be to define exact conditions under which a series of pairs $\in$ S corresponds to a model. We can then show that any subset of that series also has a model.

For those series which are stable, an assignment function g over the justification and a formula to T is introduced.

(It might turn out better to make the derivation set cumulative as the justification set is, and let g be a mapping over the justification and accumulated formulae to a model).

Representing other formalisms in such a framework looks interesting. Some analogue of circumscription might easily go into the justification set like this:

Bird(x):fly(x) & tweety(x)/fly(x).

Details not yet conceived. Nor yet how exactly to tie in other formalisms.

I am by no means entirely happy with the formulation to date - I wish to give formal stability conditions once I am more certain that the tree I envisage is being built accurately.

It is important to achieve the following results.

1. Show that the structure exactly captures Reiter' notion of a default proof.

2. That every extension in this system
   a. Corresponds exactly to one of Reiter's
   b. Is a maximal extension

3. That every extension in Reiter's system corresponds to an extension in this system.

4. Exact conditions regarding when a default has an extension. It looks hopeful that we may be able to achieve this with more clarity and detail than Reiter,s formulation.

Looking ahead

Remarkable expressive power is to be had if the agent's set O is taken to be a subset of the set of wffs $\in$ event-based temporal logic where Time-points are reflexive and transitive.

In particular, it will be possible to describe the temporal truth-conditions of one sentence under a justification expressing the time-conditions of another. The hope is that this will make certain types of reasoning about change quite natural.