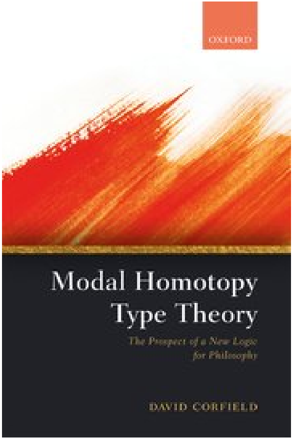
Graded modalities and dependent type theory

David Corfield

University of Kent

June 2022

New tools for philosophy



One central thesis of the book

Wherever philosophers deploy first-order (modal) logic – whether in philosophy of language, metaphysics or epistemology, they would gain enormously by switching to some form of modal type theory.

Graded modalities provide a further opportunity to demonstrate this thesis.

Since computer science has been at the forefront of the practical use of type theory, we should expect to learn from there.

Introduction

Philosophers' modal logic

- The examination of various modalities:
  - *alethic* (it is necessarily/possibly the case that),
  - *epistemic* (it is known that),
  - *doxastic* (it is believed that),
  - *deontological* (it ought to be that),
  - *temporal* (it has been the case that),
  - ...
- We might consider the differences, if any, between *physical*, *metaphysical* and *logical* necessity and possibility.
- Technically, still largely using modal propositional logics (K, S4, S5, etc.) and modal first-order logic, then Kripke models for semantics.

Computer scientists' modal logic

- New modalities to represent security levels, resource usage, and generally, *effects* and *coeffects*.
- Re-use of philosophers' modalities for different purposes, e.g.,:
  - Model-checking (temporal)
  - Multi-agent systems (epistemic)
- Technically, formalisms include: sub-structural logics, (idempotent) monads, coalgebra, labelled transition systems, bisimulations, ...

Note that these formalisms fit closely with the type-theoretic/ category-theoretic outlook.

Graded modalities

These first appear in philosophical works in the 1970s (Goble, Fine, Lewis).

Topics of my fellow speakers:

- In philosophy/linguistics: *it is quite possible*, *it is more permissible*, *it is strongly believed*, ...
- In computing, e.g., for a finer treatment of resource use: use once, use twice, etc. Technical tools include *graded modalities*.

Advantages of dependent type theory over FOL

Advantages of dependent type theory over FOL

First-order logic is just a delimited form of dependent type theory.

DTT

Advantages of Dependent type theory over FOL

First-order logic is just a delimited form of dependent type theory.

FOL

FOL perspectives

The adoption of FOL encourages certain philosophical perspectives:

- An epistemic agent's state is given in terms of the collection of propositions they hold.
- There's a domain of everything which is to be treated as a uniform collection.
- Entities and properties are treated very distinctly.

Judgement in DTT

Fundamentally, judgement is of an element in a type, $\vdash a : A$.

Since a proposition is a kind of type, namely, a type with at most one element, this includes judging the truth of a proposition.

It also includes the judgement of something being a type.

The epistemic agent knows about type formation and the construction and use of elements of types.

For every proposition we need an account of how it has been formed. Let's try a couple of examples.

Have you been drinking this morning?

This question arises from the type ‘You have been drinking this morning’.

Let’s take it to be the case that this proposition is true if any alcoholic drink has passed your lips this morning.

This proposition arises in turn from the type of (alcoholic) drinks you’ve had this morning.

G & T at 11am, Whisky at 10am: Drinks you’ve had this morning.

Both your G & T at 11am and your Whisky at 10am give rise to the truth of the proposition.

They corresponding to equal elements, |G & T at 11am| and |Whisky at 10am|, of the proposition above.

We might name them ‘The fact that you have been drinking this morning’, a term which can be used in a proposition ‘...means that you’ve broken your promise to me’.

Four entities

- Drinks you've had this morning: Type
- You've had a drink this morning: Proposition
- the G & T at 11am: Drinks you've had this morning
- $|$G & T at 11am$|$ = the fact that you've been drinking this morning.

They don't belong to some common universe of discourse to be quantified over, or to be set as potentially equal.

(Of course, we also need to see how the type 'Drinks you've had this morning' is formed. A type of events that are accomplishments.)

Anaphoric pronouns

Whenever someone's child dishonours them, they punish them for it.

- $x : Person, y : Child(x), z : Dishonour(x, y) \vdash Punish(x, y, z) : Type$
- $x : Person, y : Child(x), z : Dishonour(x, y) \vdash p : Punish(x, y, z)$

(Then dependent product)

Whenever an author creates a character in one of their books, there's something of their life represented there in the character's story.

- $x : Author, y : Book(x), z : Character(x, y) \vdash Autobiographical(x, y, z) : Type$
- $x : Author, y : Book(x), z : Character(x, y) \vdash a : Autobiographical(x, y, z)$

Judgement structure

Iterated dependency structure of any length:

- $x : A, y : B(x), z : C(x, y) \vdash d : D(x, y, z)$

We may consider first-order logic (FOL) as a form of dependent type-theory, but a very restricted one.

Typed FOL has one layer of non-dependent types (sets), all other types are dependent propositions – predicates, relations.

- $x : A, y : B, z : C, P(x), Q(y, z), R \vdash S(x, y, z)$

Note that we typically don't mention elements of propositions, but take the presence of these propositions to indicate that they hold.

Now in *untyped* (unityped) FOL, we only allow one non-dependent type, and then omit to mention it.

- $x : A, y : B, z : C, P(x), Q(y, z), R \vdash S(x, y, z)$
- $x, y, z : A, P(x), Q(y, z), R \vdash S(x, y, z)$
- $P(x), Q(y, z), R \vdash S(x, y, z)$

The domain remains implicit, and all variables range over it.

The form of a logic gives rise to a metaphysical perspective.

- Untyped FOL suggests that objects belong to a domain of 'everything', and that then there are properties corresponding to subclasses of the domain.
- There's a clear distinction between object and property.

DTT has a type of entities with elements (A, a) for $a : A$. But such elements (in HoTT) may have complicated identity structures. We no longer have an untyped proposition, $a = b$.

Properties are dependent propositions, so dependent types, depending on a specific type. There's a type of properties (A, P), where $x : A \vdash P(x) : \mathit{Proposition}$.

So then $(A, P) : \mathit{Property}$ and $((A, P), \mathit{Property}) : \mathit{Entity}$.

Type formation

After deciding on the judgement structure, we now need to provide ways to form types.

Type formation is specified by rules saying when we can introduce a type, how to construct terms of that type and how to use or "eliminate" terms of that type.

- $\mathbf{0}; \mathbf{1}; A \times B; A + B; [A, B]; Id_A(a, b); \ldots$

manifest themselves under the FOL-restriction as

- False; True; conjunction; disjunction; functions and implication; equality;...

We also need to extend the product and function types to dependent types.

Consider the role of conjunction:

- $A, B \vdash C$, therefore $A \& B \vdash C$.

Similarly, there is the type formation known as dependent sum (or dependent pair)

- $y : Author, z : Book(y)$
- $(y, z) : \sum_{x:Author} Book(x)$

Pairing with a predicate gives 'An author who is ...'

- $y : Author, z : British(y)$,
- $(y, z) : \sum_{x:Author} British(x)$

- $A \vdash B$, therefore $\vdash A \rightarrow B$.

Then dependent product (or dependent function)

- $x : Author \vdash f(x) : Book(x)$
- $\vdash f : \prod_{x:Author} Book(x)$ (only such an f if all authors are British).

In the case of a proposition:

- $x : Author \vdash f(x) : British(x)$
- $\vdash f : \prod_{x:Author} British(x)$

Quantifiers

- Dependent sum/pair; dependent product/function

when delimited, appear in FOL as

- Existential quantifier; universal quantifier

FOL perspectives

The adoption of FOL encourages certain philosophical perspectives:

- An epistemic agent's state is given in terms of the collection of propositions they hold.
- There's a domain of everything which is to be treated as a uniform collection.
- Objects and properties are treated very distinctly (quantification extended to properties in second-order logic).

By contrast, with DTT:

- The agent knows about the formation of types, and the construction and use of their elements.
- Each entity is typed. There is no uniform collection of everything.
- Objects and properties are elements of their respective types.

Modal logic

What then of modal logic with expressions: $\Diamond P$, $\exists x \Diamond Q(x)$, $\Diamond \exists x Q(x)$?

We should expect modal operators in DTT to apply to types in general.

(Graded) Modality as variation over types

The Kant-Sellars thesis about modality

Robert Brandom writes in *From empiricism to expressivism*:

> *...in being able to use nonmodal, empirical descriptive vocabulary, one already know how to do everything one needs to know how to do in order to deploy modal vocabulary, which according can be understood as making explicit structural features that are always already implicit in what one does in describing. (Brandom 2015, p. 143)*

The uses of nonmodal vocabulary being made explicit by modal language include those where one describes how the state of something would be under certain kinds of **variation of its current situation**.

(Graded) Modality as variation over types

Whenever there's a function between types $f : A \to B$, it generates a triple of functions going between A-dependent things and B-dependent things.

(Dependent sum, dependent product and base change within a context, or interpreted in a slice category over B, if that is of any help.)

Three functions: 1

Take a mapping

$$Owner : Dog \to Person,$$

then any property of people can be transported to a property of dogs, e.g.,

$$Being\ French \mapsto Being\ owned\ by\ a\ French\ person.$$

We shouldn't expect every property of dogs will occur in this fashion.

In other words, we can't necessarily invert this mapping to send, say, 'Pug' to a property of People.

Three functions: 2

We can try...

$$Pug \mapsto Owning\ some\ pug \mapsto ???$$

Three functions

But then

$$Pug \mapsto Owning\ some\ pug \mapsto Owned\ by\ someone\ who\ owns\ a\ pug.$$

However, people may own more than one breed of dog.

Three functions: 3

How about

$$Pug \mapsto Owning\ only\ pugs \mapsto ???$$

Three functions

But this leads to

$$Pug \mapsto Owning\ only\ pugs \mapsto Owned\ by\ someone\ owning\ only\ pugs$$

But again, not all pugs are owned by single breed owners.

A glimpse of modality

Note the modal flavour of these properties:

- *Pug*
- *Owned by someone who owns a pug* = *possibly pug*
- *Owned by someone owning only pugs* = *necessarily pug*

Under variation over co-owned dogs, the pug property is allowed/ensured.

Modal logic

What if we take a map $Worlds \to \mathbf{1}$?

Consider (world-dependent) propositions, or subsets of worlds,

$$w : W \vdash P(w) : Prop$$

Things then work out best if we compose dependent sum (function 2) or dependent product (function 3) followed by base change (function 1).

This yields *possibly* P and *necessarily* P as propositions dependent on the type *Worlds*, and as such comparable to P.

$$w : W \vdash \Diamond P(w) : Prop \qquad\qquad w : W \vdash \Box P(w) : Prop$$

Accessible worlds

More generally, we might consider an equivalence relation generated by a function: $f : W \to V$. Then

- Necessarily $P(w)$ holds at a world if P holds at all related worlds.
- Possibly $P(w)$ holds at a world if it holds at some related world.

Note that we don't need to restrict to propositions.

First glimpse of grading

Now we might consider a sequence of such V:

$$V_1 \twoheadrightarrow V_2 \twoheadrightarrow \cdots \twoheadrightarrow V_n,$$

with compatible maps $f_i : W \to V_i$. So a sequence of progressively coarser relations.

E.g., people classed as to home town, home province, home country.

Then we have a sequence of operators, $\Box_i$ and $\Diamond_i$, so that, e.g., for $i < j$, (the truth of) $\Box_j P$ implies $\Box_i P$ and $\Diamond_i P$ implies $\Diamond_j P$.

In history

> Since the notions of necessity and contingency assume *sets* of more or less similar events, their application is *inherently sensitive to the descriptions we use in referring to events*. To assess degree of necessity, we need to know whether the same type of event would have occurred given a certain change or intervention. Our assessment hinges, therefore, on modes of sorting and individuation, on what we consider a type, or the same type. "The war was necessary" means "a similar kind of war would have occurred in any event," hence a judgment about the degree of necessity that should be ascribed to an event will depend on how broadly or narrowly we construe the type in question. Knowing early twentieth-century European history, we may believe a war would have started sooner or later. But if the historian uses finer distinctions – a war in 1914, a war triggered by an assassination, she might lower the level of necessity she ascribes to the war, attributing increased significance to the assassination in Sarajevo. (Ben-Menahem, Y., p 124)

'Historical Necessity and Contingency', in A. Tucker (ed.) *A Companion to the Philosophy of History and Historiography*, Blackwell, 2009.

Grading

Indices of the modal operators belong to some structured set.

E.g., in the case of a sequence of n equivalence relations, for $i < j$, $\Diamond_i \Diamond_j P$ is equivalent to $\Diamond_j P$.

Obligatoriness

An idea: degrees of obligatoriness could be captured by the size of the domain of variation under which an action is required.

- So long as Y you must do X.
- It is unconditional that you do X.

Variation via spans

We can think of the modalities resulting from a map $f : W \to V$ as a passage through $W \to V \leftarrow W$, and this in turn through the associated span $W \leftarrow R \to W$, representing the equivalence relation.

This affords the possibility of generalization to any relation (not necessarily an *equivalence* relation), such as the relation of time instants, $R(t_1, t_2)$ iff $t_2 - t_1 \geq 0$.

But then we might consider $R_d(t_1, t_2)$ iff $t_2 - t_1 \geq d$, allowing us to express, e.g., that something happened at least d days ago, a graded temporal operator. (There is some R_d interval ending now which began with...)

There would be an additive structure on the grading.

Intermodalities

We may generalise further to any span $A \leftarrow C \rightarrow B$. This gives rise to the 'intermodalities' of Fong, Myers, Spivak in *Behavioral Mereology: A Modal Logic for Passing Constraints* (2021):

$$A \leftarrow C \rightarrow B$$

Here we have *intermodalities* relating properties of A to properties of B. Think of A and B as the types of behaviours of subparts of a system, where we are interested in cases where the first part being in a certain kind of state *allows* or *ensures* that the second part is in a specific kind of state.

Then we might have a variety of spans, corresponding to the degree of tightness of the constraint between the parts, or to a way to parameterise the dynamics if seen as temporal parts.

Specially structured type of worlds

Worlds with metric

We might also consider a type of worlds with some extra structure.

- A designated type of worlds, W, and a (possibly asymmetric) distance function $g : W \times W \to \mathbb{R}^{\geq 0}$.
- A designated type of worlds, W, and a (possibly asymmetric) goodness function $g : W \times W \to D$.

Something along these lines might work for Lewis on counterfactuals and obligations.

Let's consider a possible route to develop a natural metric on worlds.

Worlds as contexts

Consider what might be the beginning of a story, or a play:

> *A man walks into a bar. He's whistling a tune. A woman sits at a table in the bar. She's nursing a drink. On hearing the tune, she jumps up, knocking over the drink. She hurls the glass at him. "Is that any way to greet your husband?", he says.*

For Ranta (1994, *Type-theoretic grammar*), this kind of narrative should be treated as the extension of one long *context*, with its dependency structure, which begins as follows:

$$x_1 : Man, x_2 : Bar, x_3 : WalksInto(x_1, x_2), x_4 : Tune, x_5 : Whistle(x_1, x_4), x_6 : Woman, x_7 : Table, x_8 : Locate(x_7, x_2), x_9 : SitsAt(x_6, x_7), x_{10} : Drink, x_{11} : Nurse(x_6, x_{10}), x_{12} : Hear(x_6, x_5), \ldots$$

Contexts

In general, a context in type theory takes the form

$$\Gamma \;=\; x_0 : A_0,\; x_1 : A_1(x_0),\; x_2 : A_2(x_0, x_1),\; \ldots, x_n : A_n(x_0, \ldots, x_{n-1}),$$

where the A_i are types which may be legitimately formed. As we add an item to a context, there may be dependence on any of the previous variables.

A context need not take full advantage of this array of dependencies. For instance, in the case above, *Whistle* only depends upon x_1 and x_4 and not upon x_2 or x_3. But it certainly cannot depend on a variable ahead of it.

Contexts

Any such context corresponds itself to an object, the iterated dependent sum of the context. Let W represent the iterated sum, and W_i the stages of the construction of W, then the maps we considered earlier

$$W\text{-}dependent \overset{\longrightarrow}{\underset{\longrightarrow}{\longleftarrow}} \mathbf{1}\text{-}dependent$$

now factor through the successive stages of the construction of the context:

$$W\text{-}dependent \overset{\longrightarrow}{\underset{\longrightarrow}{\longleftarrow}} \cdots \overset{\longrightarrow}{\underset{\longrightarrow}{\longleftarrow}} W_2\text{-}dependent \overset{\longrightarrow}{\underset{\longrightarrow}{\longleftarrow}} W_1\text{-}dependent \overset{\longrightarrow}{\underset{\longrightarrow}{\longleftarrow}} \mathbf{1}\text{-}dependent$$

$$\Gamma \;=\; x_0 : A_0,\; x_1 : A_1(x_0),\; x_2 : A_2(x_0, x_1),\; \ldots x_n : A_n(x_0, \ldots, x_{n-1}),$$

We could construct modal operators, etc., relative to an initial segment of the context.

Counterfactuals could work by stripping back a context until the counterfactual antecedent can hold, an idea first developed by Aarne Ranta (1991, *Constructing possible worlds*, Theoria 57 (1-2): 77-99).

Or perhaps one might use the tree-like dependency structure (Cf. first-order logic's and propositional logic's weak dependency structures)

There is such a vast store of shared knowledge that it seems that any story can go almost anywhere its author wishes. Continuing our Western,

- An elephant escapes from the box car in which it is travels with the circus and tramples all in the saloon underfoot.
- A tornado rips through the town and takes the couple somewhere over the rainbow.

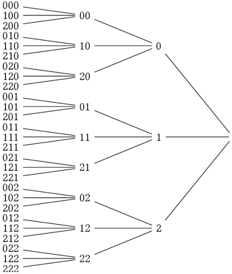
Instead, one might imagine a more controlled setting of what can occur next, where paths fan out according to circumscribed choices, as we find with computations paths in computer science or, in a more extreme form, with the collection of real numbers for the intuitionist.

Like the real numbers formed from all possible infinite decimal expansions, here we can conceive of the specification of a collection of worlds which are all possible 'complete' extensions of a context Γ.

> *Worlds appear as total infinite extensions of finitely representable approximations of them. Moreover, all we can say about a world is on the basis of some finite approximation of it, and hence at the same time about indefinitely many worlds extending that approximation. (Ranta 1991, p. 79)*

p-adic tree

The analogy with expansions of real numbers introduces the possibility of a metric on our type of worlds. Or maybe better yet, $p-$adic numbers rather than the reals.



Distances are calculated from the position of the leftmost common node. (Some form of valuation on a tree for deontological modalities?)

To conclude

I have sketched a few possibilities we might explore:

- Variation over types; over functions; over spans (relations and more general)
- Worlds as contexts, and their tree-structure.

Other leads:

- Formal treatment of graded modal dependent type theory
- Probability theory meets type theory/category theory.