**Философия. Текст 1.**

**Natural Deduction for Modal Logic**

**with a Backtracking Operator**

**Jonathan Payne**

Received: 14 October 2013 / Accepted: 6 May 2014

© Springer Science+Business Media Dordrecht 2014

**Abstract** Harold Hodes in [1] introduces an extension of first-order modal logic featuring

a *backtracking* operator, and provides a possible worlds semantics, according

to which the operator is a kind of device for ‘world travel’; he does not provide a proof

theory. In this paper, I provide a natural deduction system for modal logic featuring

this operator, and argue that the system can be motivated in terms of a reading of the

backtracking operator whereby it serves to indicate modal scope. I prove soundness

and completeness theorems with respect to Hodes’ semantics, as well as semantics

with fewer restrictions on the accessibility relation (Hodes restricts his attention

to S5).

**Keywords** Modal logic · Backtracking · Natural deduction

**1 Introduction**

Harold Hodes in [1] introduces an extension of first-order modal logic featuring a

*backtracking* operator ‘↓’. The purpose of this operator is similar to that of an actuality

operator. But, instead of exempting what follows from the scope of *all* enclosing

modal operators, it exempts it only from the *innermost* modal operator. Or, in terms

of the possible worlds semantics, instead of causing a formula to be evaluated at some

specified ‘actual’ world, the backtracking operator causes a formula to be evaluated

at the ‘last visited’ world, so to speak.

J.Hodes gives a semantics (which I give an overview of in Section 2), but does not

supply a proof theory.1 This semantics reflects a reading of the operator as one which

allows more flexible ‘travel’ through worlds in evaluating the truth value of a formula.

If the semantics of \_ and ♦ are thought of as instructions to travel to an accessible

world in which the truth value of a formula is evaluated, Hodes’ semantics for ↓ gives

instructions to travel back to the world most recently passed through, to evaluate the

truth value of a formula there. A useful comparison is to the actuality operator. The

usual semantics for an actuality operator gives directions to evaluate the truth value

of a formula at a privileged possible world in the model—the *actual world*.

**Философия. Текст 2.**

 But, as mentioned, there is another—more syntactic—reading of the operator, as

a scope indicator. That is, the operator simply indicates that what follows it is to be

exempt from the scope of the innermost modal operator. Again, compare with the

actuality operator; this may be taken to be an indicator that what follows is to be

evaluated as exempt from the scope of *all* enclosing modal operators. Thus, the role

of ↓ is akin to that of parentheses, yet allowing for more nuanced scope distinctions.

The aim of this paper is to provide a proof system for modal logics featuring the

operator, which I will claim does justice to this reading; the proof theory gives rules

for ‘looking inside’ the scope of a modal operator, and then, when the ↓ operator is

encountered, pulling the appended formula out of that scope.

As well as explaining this alternative reading of the operator, such a proof theory

may be desirable from the point of view of certain philosophical or other uses of

modal logic. For example, extensions of modal logic may be desired to gain expressive

power without committing oneself to quantifying over possible worlds (or their

analogues) or to the members of domains of possible worlds. In some cases, modal

logic may be introduced specifically for the purposes of avoiding quantification over

some entities or other. A proof theory would allow somebody not to rely on the

semantics to give sense to claims involving ↓.2

 Nonetheless, even if one is not persuaded of the need for a proof theory for such

purposes, the fact that the operator appears to admit of the syntactic reading should

be motivation enough to develop a proof theory which represents such a reading.

Before presenting the proof system, in Section 2 I will give an overview of the

semantics which Hodes provides for the operator, albeit with a few minor differences.

In Section 3 I present a natural deduction system for the operator for propositional

modal logic featuring the operator, which makes uses of labelling each line of a proof.

**Философия. Текст 3.**

**2 Semantics**

Let *L*↓ be a typical language for propositional modal logic; it consists of countably

many propositional variables *p,q, r, . . .*, connectives ∧ and ¬ and a necessity

operator \_. In addition, it shall have an additional sentential operator ↓, called the

*backtracking operator*. The intended effect of the backtracking operator will be to

exempt what follows it from the scope of the innermost modal operator from which

it is not already exempt (so, for example, *p*, \_↓*p* and \_\_↓↓*p* should all be counted

as equivalent). Other sentential connectives ∨*,*→and a possibility operator ♦ can be

defined in the usual way.

The semantics presented here is essentially that of [1], with the main differences

being: (a) Hodes’ semantics is for quantified modal logic, whereas I will

only describe the case for propositional modal logic (I discuss extending to quantified

modal logic briefly in Section 7). (b) Hodes’ logic is an extension of S5, so

that the accessibility relation is an equivalence, whereas the only restriction on the

equivalence relation here is that it is serial. (c) Hodes only defines satisfaction for

a certain class of formulas, whereas the semantics presented here places no such

restriction.

A model is a triple *M* = \_*W,R, a*\_, where *W* is a set (of possible worlds), *R* ⊆

*W* Ч*W* is the accessibility relation, and *a* is an assignment function which assigns to

each propositional variable *p* at a world *w* ∈ *W* a truth value *a(w, p)* ∈ {*T,F*}.

Only one restraint will be placed on the accessibility relation for now, and that is

that it is *serial*. So, for any *w* ∈ *W* there is a *w*

 ∈ *W* such that *wRw*

.

Then, a satisfaction relation is defined, not for each world, but for each finite

sequence of worlds of the appropriate type. So, we first make the following

definition:

**Definition 1** Given a model*M*, a *world sequence* is a member of the following set:

WS*M*

= {\_*w*1*, . . . , wk*\_ : *k* ≥ 1*,* ∀*i* ≤ *k,wi* ∈ *W* and ∀*i < k,wi Rwi*+1}

As a result of seriality, for every world sequence there will be a world sequence

extending it (and so there are world sequences of arbitrary length). The sequence of

worlds at which a formula is evaluated may be thought of as a kind of memory, which

keeps track of which worlds have been travelled through.

Some terminology for members of WS*M* will be useful. I shall write **w** for an

arbitrary member of WS*M*. Where **w** = \_*w*1*, . . . , wk*\_, then:

**Философия. Текст 4.**

Before proving that the inference rules given here are both sound and complete with

respect to the semantics, I would first like to say more about the motivation behind

the proof theory. There are two aims which I have. The first concerns a worry which

may be had if the proof theory is wanted in order to avoid reliance on the possible

worlds semantics for more than pragmatic reasons. Itmight be worried that, due to the

presence of labels—which it is tempting to take as referring to worlds or sequences

of worlds—the proof theory does not succeed in avoiding reliance on the possible

worlds semantics. Secondly, I claimed before that the natural deduction system can

be seen as explaining the reading of the ↓ operator as exempting formulas from the

scope of other operators. Here I will argue for that claim.

One way in which we may try to motivate the proof system is by reference to the

semantics. On this view, a labelled formula is a formula of a kind of extended language,

and the labels are something like variables referring to sequences of worlds.

Then, a formula (of the extended language) *φ*; **s** makes the claim that *φ* is true at the

sequence of worlds **s**. The inference rules then aim to capture certain valid inferences

in this language. A soundness theorem will then be an essential part of the motivation

of the proof theory, in that it will show that the inference rules are indeed valid

inference rules—that is, they are truth preserving in the sense of the semantics. (This

will also require an extension of the semantics given in Section 2 so that labelled

formulas are given satisfaction conditions.)

But given one motivation for developing a proof theory, this will not do. There

are potential philosophical applications of such a proof theory whose aim is to

Natural Deduction for Modal Logic with a Backtracking Operator

avoid reference to, and quantification over, worlds. That is, they wish to relegate

the semantics to a purely secondary, pragmatic role. But if the proof theory is ultimately

motivated by the semantics, then this can not be the case. In any case, such a

motivation would fail to shed light on the scope exemption reading of the ↓ operator.

**Философия. Текст 5.**

So, instead, we wish to have a motivation for the proof theory which derives from

this scope exemption reading of the operator. Then, a soundness theorem will not play

a role in motivating the proof theory in light of the semantics, but will rather play the

role—along with a completeness theorem—of motivating the pragmatic value of the

semantics. For, given a soundness theorem, one will be able to use the semantics for

useful ends, in proving that a certain formula is not derivable from other formulas,

and so on.

How might such a motivation look? Firstly, labels must not be thought of as

referring to worlds. Indeed, it should be borne in mind that they are not part of the language

at all. They are merely part of the proof theory, and can be explained as a kind

of bookkeeping device, not dissimilar to the use of line numbers, the lists of undischarged

assumptions which are common in many ways of laying out formal proofs,

or even to the various horizontal and vertical lines which appear in many ways of laying

out proofs. It is perhaps better to think, not of labelled *formulas*, but of labelled

*lines* (it just happens that it is simpler for metatheoretical purposes to treat labels as

attaching to formulas).

If labels are not part of the language, then there can be no danger that they refer to

anything in the semantics (just as line numbers and the like do not). Indeed, labelled

formulas are not the kind of thing that can be asserted, or the kind of thing that have

truth-conditions or satisfaction-conditions. Since labelled formulas are not the kind

of thing that can be true or false, and the inference rules are relations between labelled

formulas, it follows that the inference rules can not be motivated in terms of validity

(i.e. necessary truth preservation). A different motivation is thus required.

The main rules which need motivating are the \_ and ↓ rules. These can be motivated,

not in terms of validity, but rather as rules for temporarily ignoring, and then

reinstating, modal operators, whilst the labels serve as a reminder as to when a modal

operator is being ignored. So, rather than serving as a memory of worlds, **s** serves as

a memory of modal scope.