Conditionals and Actuality*

(forthcoming in *Erkenntnis*)

Timothy Williamson

University of Oxford

Introduction

Indicative and subjunctive conditionals are known to interact differently with a natural rigidifying reading of the "actually" operator. The basic observation here goes back at least to Robert Stalnaker (1975), although Frank Jackson (1981: 129; 1987: 74-75) has used it in arguing against semantic accounts of the indicative conditional such as Stalnaker's that use a framework of possible worlds.[1] Here is a variant on the original examples:

(1i) If Jim is two metres tall, Jim is actually two metres tall.

(1s) If Jim had been two metres tall, Jim would have actually been two metres tall.

The indicative conditional (1i) is an obvious truism, verifiable on broadly logical grounds. By contrast, the subjunctive conditional (1s) is actually false on the relevant reading unless Jim is actually two metres tall, given that he could have been two metres tall. To put the point in terms of Lewis-Stalnaker semantics for counterfactuals, (1s) says that in the closest possible world(s) (if any) in which Jim is two metres tall, Jim is two metres tall back here in the actual world. Given that Jim could have been two metres tall, there are such worlds, so that truth-condition requires Jim to be two metres tall back here. The danger for a possible worlds account of *indicative* conditionals such as Stalnaker's is that it delivers the same result for them, falsely predicting that (1i) is false if Jim is not two metres tall. However, the aim of this paper is not to criticize Stalnaker but to explore the different ways in which conditionals can interact with an 'actually' operator, and their consequences. Throughout, we will be assuming a "rigidifying" reading of 'actually', on which its function is, in possible world terms, to return the world of evaluation to the world of utterance, or more generally the circumstance of evaluation to the context of utterance, although we shall see later that we can capture such a function without assuming as much as that about the semantic framework. The rigidifying reading of 'actually' is not the only one available in English, but it *is* available in English, and it is the reading relevant to present concerns.

A difference like that between (1i) and (1s) arises even for indicative and subjunctive conditionals about the future, which are generally hard to separate:

(2i) If Mary comes tomorrow, she will actually come tomorrow.

(2s) If Mary had come tomorrow, she would actually have come tomorrow.

As before, the indicative conditional (2i) is an obvious truism, verifiable on broadly logical grounds, whereas the subjunctive conditional (2s) is actually false on the relevant reading unless Mary does actually come tomorrow, given that she could have come tomorrow.

In the foregoing respect the indicative conditional (1s) works like the strict conditional (1m):

(1m) Necessarily, if Jim is two metres tall, then Jim is actually two metres tall.

Using **A** for "actually", we can formalize (1m) as:

(1m)' $\Box(\mathbf{j} \supset \mathbf{Aj})$

Given standard principles of modal logic with an "actually" operator, (1m)' will have the same truth-value in the actual world as (1m)":

(1m)" $\Diamond\mathbf{j} \supset \mathbf{j}$

For (1m)' entails $\Diamond\mathbf{j} \supset \Diamond\mathbf{Aj}$ and $\Diamond\mathbf{Aj}$ entails $\mathbf{Aj}$, which has the same actual world truth-value as $\mathbf{j}$; conversely, if (1m)" is true in the actual world so is either $\neg\Diamond\mathbf{j}$ or $\mathbf{j}$ and if the latter $\mathbf{Aj}$ too; but $\mathbf{Aj}$ entails $\Box\mathbf{Aj}$, which entails (1m)", as does $\neg\Diamond\mathbf{j}$.

Incidentally, (1m) is a problem for naive attempts to regard “actually” as merely a syntactic scope-indicating device that simply takes what it governs out of the scope of modal operators. In particular, (1m) is certainly not equivalent to:

(1m)* If Jim is necessarily two metres tall, then Jim is two metres tall.

For if Jim is not two metres tall then (1m) is false, whereas (1m)* is still true because necessity entails truth. More generally, (1m)” is not logically equivalent to any truth-function of □**j** and **j**, otherwise its truth-value would be settled by the falsity of **j** (which requires the falsity of □**j**), irrespective of the truth-value of ◊**j**. A much more sophisticated notion of semantic scope would be needed to make sense of such an idea.

The difference in the ways in which “actually” interacts with indicative and subjunctive conditionals permits non-trivial comparisons between conditionally supposed and actual states of affairs to be made in subjunctive but not in indicative conditionals.

(3i) If I kept quiet while hunting, I caught more than I actually did.

(3s) If I had kept quiet while hunting, I would have caught more than I actually did.

Here the subjunctive conditional (3s) records the sort of consideration that helps one to learn from experience, in particular from past mistakes. By contrast, the indicative conditional (3i) tells someone who knows full well that he did not keep quiet while hunting nothing new of interest. Thus the difference in interaction with "actually" is no mere curiosity: it is a sign of significant differences in the function of such conditionals.

In what follows, we investigate the interaction between "actually" and other operators in a formal setting. Our concern here is not primarily to investigate subtleties of the use of "actually" operators in natural languages. Rather, we will take for granted an operator in a formal language with the evidently intelligible rigidifying reading in order to investigate formally some ways in which it constrains the space of available semantic options for other constructions in the language, such as indicative and subjunctive conditionals.[3]

We will not assume a framework of possible worlds semantics to handle the formal analogue of "actually", **A**, on which a formula **Aα** is evaluated as true at an arbitrary world in a model if and only if **α** is evaluated as true at the actual world of the model (while other parameters of semantic evaluation remain fixed). Instead, we will provide a more general characterization of its logical role, relative to an arbitrary semantic notion of validity. This level of generality is not motivated by any objection to the possible worlds semantics for "actually" in itself but rather by the fact that many different sorts of semantic framework have been proposed for the treatment of conditionals (not all of them even truth-conditional). The aim is to remain neutral to the extent possible between those frameworks. For heuristic purposes, some points will be informally glossed in terms of possible worlds, but such glosses are inessential to the official arguments.

To anticipate: In the sorts of language with which we shall be concerned, the behaviour of the subjunctive conditional noted above is typical of a non-truth-functional conditional, whereas the corresponding behaviour of the indicative conditional seems to exclude any natural reading except that of a truth-functional conditional.[2]

Language

We consider a formal language L with a countable infinity of atomic sentence variables, the usual truth-functional operators, including the 0-place falsity constant **f** and the truth constant **t**, negation (¬), conjunction (**&**), disjunction (∨), the material conditional (⊃) the material biconditional (≡), the (atomic) one-place sentential operator **A** and an unspecified set of further (atomic) operators (specific consideration of non-truth-functional conditionals is postponed to the final section); **p**, **q**, **r**, … are atomic sentence variables of L and **α**, **β**, **γ**, … are any sentences of L. Below, 'tautology' means truth-functional tautology and 'variable' means sentence variable.

In what follows, we shall speak of sentential contexts, not just of sentential operators. For example, for any sentence **β** in L there is a sentential context **C** such that for any sentence **α**, **C(α)** = **¬α & β**, even if **C** does not correspond to an atomic operator of L. But we do not wish to describe *every* function from sentences of L to sentences of L as a sentential context. For instance, in L there is no sentential context **C** such that for every sentence **α**, **C(α)** is **t** if **α** is atomic and **C(α)** is **f** otherwise. Obviously, contexts in this intra-linguistic sense are not to be confused with contexts in the more usual extra-linguistic sense, in which a speaker, time and world are components of a context.

Although the notion of a sentential context seems intuitively clear, we pause to make it more precise.

A *substitution* is any function $\boldsymbol{\sigma}$ from L to L that commutes with all operators of L: that is, if $\mathbf{O}$ is an *n*-place operator then $\boldsymbol{\sigma}(\mathbf{O}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)) = \mathbf{O}(\boldsymbol{\sigma}(\boldsymbol{\alpha}_1), \ldots, \boldsymbol{\sigma}(\boldsymbol{\alpha}_n))$; in particular, $\boldsymbol{\sigma}(\mathbf{t}) = \mathbf{t}$, $\boldsymbol{\sigma}(\mathbf{f}) = \mathbf{f}$ and $\boldsymbol{\sigma}(\mathbf{A}\boldsymbol{\alpha}) = \mathbf{A}\boldsymbol{\sigma}(\boldsymbol{\alpha})$. Clearly, substitutions are closed under composition: if $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ are substitutions, there is a unique substitution $\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2$ such that $\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2(\boldsymbol{\alpha}) = \boldsymbol{\sigma}_1(\boldsymbol{\sigma}_2(\boldsymbol{\alpha}))$ for all $\boldsymbol{\alpha}$. Below, $\boldsymbol{\sigma}$ is any substitution.

A (sentential) *context* is an ordered pair $<\boldsymbol{\alpha}, \mathbf{p}>$. If $\mathbf{C} = <\boldsymbol{\alpha}, \mathbf{p}>$ then $\mathbf{C}(\boldsymbol{\beta}) = \boldsymbol{\sigma}_{\mathbf{p}/\boldsymbol{\beta}}(\boldsymbol{\alpha})$, where $\boldsymbol{\sigma}_{\mathbf{p}/\boldsymbol{\beta}}$ is the substitution such that $\boldsymbol{\sigma}_{\mathbf{p}/\boldsymbol{\beta}}(\mathbf{p}) = \boldsymbol{\beta}$ and $\boldsymbol{\sigma}_{\mathbf{p}/\boldsymbol{\beta}}(\mathbf{q}) = \mathbf{q}$ whenever $\mathbf{q} \neq \mathbf{p}$. Think of $<\boldsymbol{\alpha}, \mathbf{p}>$ as a metalinguistic analogue of $\boldsymbol{\lambda}\mathbf{p}.\boldsymbol{\alpha}$. For example, if $\mathbf{C} = <\mathbf{p}\ \&\ \neg\mathbf{q}, \mathbf{p}>$ and $\mathbf{D} = <\mathbf{p}\ \&\ \neg\mathbf{q}, \mathbf{q}>$ then $\mathbf{C}(\neg\mathbf{r}) = \neg\mathbf{r}\ \&\ \neg\mathbf{s}$ and $\mathbf{D}(\neg\mathbf{r}) = \mathbf{p}\ \&\ \neg\neg\mathbf{r}$. We say that contexts $\mathbf{C}$ and $\mathbf{D}$ are *equivalent* iff $\mathbf{C}(\boldsymbol{\alpha}) = \mathbf{D}(\boldsymbol{\alpha})$ for every sentence $\boldsymbol{\alpha}$.

We check the following fact:

0. 1. Composition of contexts. For any contexts $\mathbf{C}$ and $\mathbf{D}$, there is a context $\mathbf{CD}$ such that for all $\boldsymbol{\alpha}$, $\mathbf{CD}(\boldsymbol{\alpha}) = \mathbf{C}(\mathbf{D}(\boldsymbol{\alpha}))$.

Proof: Let $\mathbf{C} = <\boldsymbol{\gamma}, \mathbf{p}>$ and $\mathbf{D} = <\boldsymbol{\delta}, \mathbf{q}>$. Choose an $\mathbf{r}$ that occurs in neither $\boldsymbol{\gamma}$ nor $\boldsymbol{\delta}$. Set $\mathbf{E} = <\mathbf{C}(\mathbf{D}(\mathbf{r})), \mathbf{r}>$. Now for any formula $\boldsymbol{\alpha}$ the substitutions $\boldsymbol{\sigma}_{\mathbf{q}/\boldsymbol{\alpha}}$ and $\boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}\boldsymbol{\sigma}_{\mathbf{q}/\mathbf{r}}$ have the same effect on any sentence in which $\mathbf{r}$ does not occur, for $\boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}\boldsymbol{\sigma}_{\mathbf{q}/\mathbf{r}}(\mathbf{q}) = \boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}(\mathbf{r}) = \boldsymbol{\alpha} = \boldsymbol{\sigma}_{\mathbf{q}/\boldsymbol{\alpha}}(\mathbf{q})$ and $\boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}\boldsymbol{\sigma}_{\mathbf{q}/\mathbf{r}}(\mathbf{s}) = \boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}(\mathbf{s}) = \mathbf{s} = \boldsymbol{\sigma}_{\mathbf{q}/\boldsymbol{\alpha}}(\mathbf{s})$ whenever $\mathbf{s} \neq \mathbf{q}$, $\mathbf{s} \neq \mathbf{r}$. Since $\mathbf{r}$ does not occur in $\boldsymbol{\delta}$, $\boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}(\mathbf{D}(\mathbf{r})) = \boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}\boldsymbol{\sigma}_{\mathbf{q}/\mathbf{r}}(\boldsymbol{\delta}) = \boldsymbol{\sigma}_{\mathbf{q}/\boldsymbol{\alpha}}(\boldsymbol{\delta}) = \mathbf{D}(\boldsymbol{\alpha})$. Thus the substitutions $\boldsymbol{\sigma}_{\mathbf{p}/\mathbf{D}(\boldsymbol{\alpha})}$ and $\boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}\boldsymbol{\sigma}_{\mathbf{p}/\mathbf{D}(\mathbf{r})}$ have the same effect on any sentence in which $\mathbf{r}$ does not occur, for $\boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}\boldsymbol{\sigma}_{\mathbf{p}/\mathbf{D}(\mathbf{r})}(\mathbf{p}) = \boldsymbol{\sigma}_{\mathbf{r}/\boldsymbol{\alpha}}(\mathbf{D}(\mathbf{r})) =$

$\mathbf{D(\alpha)} = \boldsymbol{\sigma}_{\mathbf{p/D(\alpha)}}\mathbf{(p)}$ and $\boldsymbol{\sigma}_{\mathbf{r/\alpha}}\boldsymbol{\sigma}_{\mathbf{p/D(r)}}\mathbf{(s)} = \boldsymbol{\sigma}_{\mathbf{r/\alpha}}\mathbf{(s)} = \mathbf{s} = \boldsymbol{\sigma}_{\mathbf{p/D(\alpha)}}\mathbf{(s)}$ whenever $\mathbf{s} \neq \mathbf{p}$, $\mathbf{s} \neq \mathbf{r}$. Since $\mathbf{r}$ does not occur in $\boldsymbol{\gamma}$, $\mathbf{E(\alpha)} = \boldsymbol{\sigma}_{\mathbf{r/\alpha}}\mathbf{(C(D(r)))} = \boldsymbol{\sigma}_{\mathbf{r/\alpha}}\boldsymbol{\sigma}_{\mathbf{p/D(r)}}(\gamma) = \boldsymbol{\sigma}_{\mathbf{p/D(\alpha)}}(\gamma) = \mathbf{C(D(\alpha))}$. ■

Validity

We assume that some semantic notion of validity is given, and write $\models \boldsymbol{\alpha}$ to mean that $\boldsymbol{\alpha}$ is valid. For simplicity, we consider only single formula validity, although the discussion could certainly be extended to validity for arguments with one or more premises. We assume for purposes of inquiry that validity has the following properties:

MP If $\models \boldsymbol{\alpha} \supset \boldsymbol{\beta}$ and $\models \boldsymbol{\alpha}$ then $\models \boldsymbol{\beta}$

TAUTOLOGY If $\boldsymbol{\alpha} \equiv \boldsymbol{\beta}$ is a tautology then $\models \mathbf{C(\alpha)} \supset \mathbf{C(\beta)}$

SUBSTITUTION If $\models \boldsymbol{\alpha}$ then $\models \boldsymbol{\sigma}\mathbf{(\alpha)}$

ACTUALITY $\models \mathbf{C(A\alpha)} \equiv \mathbf{[[\alpha \supset C(t)] \ \& \ [\neg\alpha \supset C(f)]]}$

Comments: MP is modus ponens for the material conditional $\supset$; modus ponens for any other conditional the language may contain is another matter. TAUTOLOGY implies that the language does not contain operators so fine-grained that they are sensitive to differences between truth-functional equivalents; it ensures that the background logic for the usual connectives is classical. SUBSTITUTION is a "smoothness" condition; dropping it would make little difference to the main line of argument below, but SUBSTITUTION does allow us to talk of the validity of all instances of a schema by talking of the validity of a single instance of it in which distinct atomic sentences stand in for distinct schematic letters. ACTUALITY expresses the rigidifying effect of **A** in all contexts in L: actualized truths behave like tautologies and actualized falsehoods like contradictions. Implicit in ACTUALITY is the validity of the equivalence of truth and actual truth: in the terminology of Davies and Humberstone (1980), this corresponds to real world validity (truth at the actual world in each model) rather than the modal notion of general validity (truth at every world in each model). Indeed, since ACTUALITY applies to all sentential contexts in the language, **A** must rigidify all parameters in a circumstance of evaluation: the time as well as the world ("actually now") and even the assignment of values to individual variables in a quantified language. However, as already noted, for present purposes we are avoiding official commitment to such semantic glosses on **A** in terms of possible worlds.

We are not claiming that all languages must obey these four conditions. Rather, the claim is just that they are attractive enough for their consequences to be worth exploring. For example, if we add a standard "actually" operator **A**, a counterfactual conditional and operators for metaphysical possibility and necessity to a standard propositional calculus and consider real world validity on a possible worlds semantics, we obtain all four features.

For future use, we make explicit a few elementary consequences of these constraints. First, we check that validity is closed under classical truth-functional reasoning.

1.1. If $\boldsymbol{\alpha}$ is a tautology then $\models \boldsymbol{\alpha}$.

Proof: Since $\mathbf{t} \equiv \mathbf{t}$ is a tautology, $\models \mathbf{t} \supset \mathbf{t}$ by TAUTOLOGY. But if $\boldsymbol{\alpha}$ is a tautology then $[\mathbf{t} \supset \mathbf{t}] \equiv \boldsymbol{\alpha}$ is a tautology, so $\models [\mathbf{t} \supset \mathbf{t}] \supset \boldsymbol{\alpha}$ by TAUTOLOGY again, so $\models \boldsymbol{\alpha}$ by MP. ■

1.2. If $\models [\boldsymbol{\alpha}_1 \mathbin{\&} \ldots \mathbin{\&} \boldsymbol{\alpha}_n] \supset \boldsymbol{\beta}$ and $\models \boldsymbol{\alpha}_1, \ldots, \models \boldsymbol{\alpha}_n$ then $\models \boldsymbol{\beta}$.

Proof: By induction on n. Basis: For $n = 0$ we treat the conjunction as $\mathbf{t}$, so $\models \mathbf{t} \supset \boldsymbol{\beta}$, but $\models \mathbf{t}$ by 1.1, so $\models \boldsymbol{\beta}$ by MP. Induction step: Suppose that the result holds for n. Suppose also that $\models [\boldsymbol{\alpha}_1 \mathbin{\&} \ldots \mathbin{\&} \boldsymbol{\alpha}_{n+1}] \supset \boldsymbol{\beta}$ and $\models \boldsymbol{\alpha}_1, \ldots, \models \boldsymbol{\alpha}_{n+1}$. But $[[\boldsymbol{\alpha}_1 \mathbin{\&} \ldots \mathbin{\&} \boldsymbol{\alpha}_{n+1}] \supset \boldsymbol{\beta}] \supset [\boldsymbol{\alpha}_{n+1} \supset [[\boldsymbol{\alpha}_1 \mathbin{\&} \ldots \mathbin{\&} \boldsymbol{\alpha}_n] \supset \boldsymbol{\beta}]$ is a tautology. Hence by 1.1 and two steps of MP $\models [\boldsymbol{\alpha}_1 \mathbin{\&} \ldots \mathbin{\&} \boldsymbol{\alpha}_n] \supset \boldsymbol{\beta}$, so $\models \boldsymbol{\beta}$ by induction hypothesis. ■

1.3. If $[\boldsymbol{\alpha}_1 \mathbin{\&} \ldots \mathbin{\&} \boldsymbol{\alpha}_n] \supset \boldsymbol{\beta}$ is a tautology and $\models \boldsymbol{\alpha}_1, \ldots, \models \boldsymbol{\alpha}_n$ then $\models \boldsymbol{\beta}$.

Proof: By 1.1 and 1.2. ■

Most of the subsequent proofs in this paper use at least one of these three lemmas without comment.

We now make explicit that validity is real world validity:

1.4. $\models \mathbf{A}\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}$

Proof: Setting $\mathbf{C} = \langle\mathbf{p}, \mathbf{p}\rangle$ in ACTUALITY, $\models \mathbf{A\alpha} \equiv \mathbf{[[\alpha \supset t] \ \& \ [\neg\alpha \supset f]]}$. But $\mathbf{\alpha} \equiv \mathbf{[[\alpha \supset t] \ \& \ [\neg\alpha \supset f]]}$ is a tautology, so $\models \mathbf{A\alpha} \equiv \mathbf{\alpha}$. ■

Call a context $\mathbf{C}$ *congruential* iff whenever $\models \mathbf{\alpha} \equiv \mathbf{\beta}$, $\models \mathbf{C(\alpha)} \equiv \mathbf{C(\beta)}$ ('substitution of logical equivalents'). Call $\mathbf{C}$ *extensional* iff always $\models \mathbf{[\alpha \equiv \beta] \supset [C(\alpha) \equiv C(\beta)]}$. Of course, by MP all extensional contexts are congruential. Many languages contain operators that create congruential, non-extensional contexts: for instance, modal operators in simple languages of propositional modal logic. Extensionality is almost but not quite equivalent to truth-functionality (Humberstone 1986, 1997). All truth-functional contexts are extensional. An example of an extensional context that is not truth-functional is $\langle\mathbf{p \ \& \ q, p}\rangle$. Although $\models \mathbf{[\alpha \equiv \beta] \supset [[\alpha \ \& \ q] \equiv [\beta \ \& \ q]]}$, because it is a tautology, the truth-value of $\mathbf{p \ \& \ q}$ is not determined by the truth-value of $\mathbf{p}$ independently of that of $\mathbf{q}$. In practice, all the non-truth-functional operators commonly discussed in the literature are also non-extensional.

The following result shows a perhaps surprising effect of the presence of $\mathbf{A}$ in L:

1.5. All congruential contexts are extensional.

Proof: Suppose that $\mathbf{C}$ is congruential. Let $\mathbf{\alpha}$ and $\mathbf{\beta}$ be any sentences. By 1.4, $\models \mathbf{A\alpha} \equiv \mathbf{\alpha}$. Since $\mathbf{C}$ is congruential, $\models \mathbf{C(A\alpha)} \equiv \mathbf{C(\alpha)}$. So by ACTUALITY $\models \mathbf{C(\alpha) \equiv [[\alpha \supset C(t)] \ \& \ [\neg\alpha \supset C(f)]]}$. Similarly, $\models \mathbf{C(\beta) \equiv [[\beta \supset C(t)] \ \& \ [\neg\beta \supset C(f)]]}$. But $\mathbf{[\alpha \equiv \beta] \supset [[[\alpha \supset C(t)] \ \& \ [\neg\alpha \supset C(f)]] \equiv [[\beta \supset C(t)] \ \& \ [\neg\beta \supset C(f)]]]}$ is a tautology, so $\models \mathbf{[\alpha \equiv \beta] \supset [[[\alpha \supset C(t)] \ \& \ [\neg\alpha \supset C(f)]] \equiv [[\beta \supset C(t)] \ \& \ [\neg\beta \supset C(f)]]]}$. Hence $\models \mathbf{[\alpha \equiv \beta] \supset [C(\alpha) \equiv C(\beta)]}$. ■

The proof can be thought of as valid analogue of the notorious Frege-Church "slingshot" argument, for sentence rather than name position. Since we want to allow non-extensional contexts into L (such as the context of the antecedent or consequent of a subjunctive conditional), we must allow non-congruential contexts. Thus merely establishing that $\models \boldsymbol{\alpha} \equiv \boldsymbol{\beta}$ does not in general show that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are interchangeable with respect to validity. It would therefore also be nice to have some notion of semantic equivalence broader than truth-functional equivalence that *does* permit interchangeability of such equivalents without change of relevant semantic status. The next section shows how to define such a notion from present resources.

Validity*

We define a stronger notion of validity, written $\models^*$, by setting $\models^* \boldsymbol{\alpha}$ if and only if for every context $\mathbf{C}$, $\models \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\mathbf{t})$. Thus the valid* sentences are those semantically interchangeable with a tautology. We establish some properties of $\models^*$. First, we check that $\models^*$ is indeed at least as strong as $\models$, in the sense of imposing at least as high a standard for validity.

2.1. If $\models^* \boldsymbol{\alpha}$ then $\models \boldsymbol{\alpha}$.

Proof: Suppose that $\models^* \boldsymbol{\alpha}$. Setting $\mathbf{C} = \langle \mathbf{p}, \mathbf{p} \rangle$, $\models \boldsymbol{\alpha} \equiv \mathbf{t}$. Since $[\boldsymbol{\alpha} \equiv \mathbf{t}] \supset \boldsymbol{\alpha}$ is a tautology, $\models \boldsymbol{\alpha}$. ■

Now we check that classical truth-functional reasoning is valid in the stronger sense.

2.2. If $\boldsymbol{\alpha}$ is a tautology then $\models^* \boldsymbol{\alpha}$.

Proof: If $\boldsymbol{\alpha}$ is a tautology then $\mathbf{t} \equiv \boldsymbol{\alpha}$ and $\boldsymbol{\alpha} \equiv \mathbf{t}$ are tautologies, so by TAUTOLOGY $\models \mathbf{C(t)} \supset \mathbf{C}(\boldsymbol{\alpha})$ and $\models \mathbf{C}(\boldsymbol{\alpha}) \supset \mathbf{C(t)}$, so $\models \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C(t)}$. ■

2.3. If $\models^* \boldsymbol{\alpha} \supset \boldsymbol{\beta}$ and $\models^* \boldsymbol{\alpha}$ then $\models^* \boldsymbol{\beta}$.

Proof: Suppose that $\models^* \boldsymbol{\alpha} \supset \boldsymbol{\beta}$ and $\models^* \boldsymbol{\alpha}$. Fix a context $\mathbf{C}$. Let $\mathbf{D} = \langle \mathbf{q} \supset \boldsymbol{\beta}, \mathbf{q} \rangle$, where $\mathbf{q}$ does not occur in $\boldsymbol{\beta}$. Since $\models^* \boldsymbol{\alpha}$, $\models \mathbf{CD}(\boldsymbol{\alpha}) \equiv \mathbf{CD(t)}$. But $\mathbf{CD}(\boldsymbol{\alpha}) = \mathbf{C}(\boldsymbol{\alpha} \supset \boldsymbol{\beta})$ and $\mathbf{CD(t)} = \mathbf{C}(\mathbf{t} \supset \boldsymbol{\beta})$. Hence $\models \mathbf{C}(\boldsymbol{\alpha} \supset \boldsymbol{\beta}) \equiv \mathbf{C}(\mathbf{t} \supset \boldsymbol{\beta})$. But $[\mathbf{t} \supset \boldsymbol{\beta}] \equiv \boldsymbol{\beta}$ is a tautology, so by TAUTOLOGY $\models \mathbf{C}(\mathbf{t} \supset \boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta})$. Hence $\models \mathbf{C}(\boldsymbol{\alpha} \supset \boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta})$. Since $\models^* \boldsymbol{\alpha} \supset \boldsymbol{\beta}$, $\models \mathbf{C}(\boldsymbol{\alpha} \supset \boldsymbol{\beta}) \equiv \mathbf{C(t)}$. Thus $\models \mathbf{C}(\boldsymbol{\beta}) \equiv \mathbf{C(t)}$. Since $\mathbf{C}$ was an arbitrary context, $\models^* \boldsymbol{\beta}$. ■

2.4. If $\boldsymbol{\alpha} \equiv \boldsymbol{\beta}$ is a tautology then $\models^* \boldsymbol{\alpha}$ iff $\models^* \boldsymbol{\beta}$.

Proof: If $\boldsymbol{\alpha} \equiv \boldsymbol{\beta}$ is a tautology so are $\boldsymbol{\alpha} \supset \boldsymbol{\beta}$ and $\boldsymbol{\beta} \supset \boldsymbol{\alpha}$; now use 2.2 and 2.3. ■

Next, we check that semantic equivalence in the sense of $\models^*$ really does imply semantic interchangeability with respect to $\models$.

2.5. If $\models^* \boldsymbol{\alpha} \equiv \boldsymbol{\beta}$ then $\models \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})$.

Proof: Suppose that $\models^* \boldsymbol{\alpha} \equiv \boldsymbol{\beta}$. Let $\mathbf{D} = \langle \mathbf{q} \equiv \boldsymbol{\beta}, \mathbf{q} \rangle$ where $\mathbf{q}$ does not occur in $\boldsymbol{\beta}$. Since

$\models^* \boldsymbol{\alpha \equiv \beta}$, $\models \mathbf{CD}(\boldsymbol{\alpha \equiv \beta}) \equiv \mathbf{CD(t)}$. But $\mathbf{CD}(\boldsymbol{\alpha \equiv \beta}) = \mathbf{C}([\boldsymbol{\alpha \equiv \beta}] \equiv \boldsymbol{\beta})$ and $\mathbf{CD(t)} = \mathbf{C}(\mathbf{t} \equiv \boldsymbol{\beta})$. Thus $\models \mathbf{C}([\boldsymbol{\alpha \equiv \beta}] \equiv \boldsymbol{\beta}) \equiv \mathbf{C}(\mathbf{t} \equiv \boldsymbol{\beta})$. But $[[\boldsymbol{\alpha \equiv \beta}] \equiv \boldsymbol{\beta}] \equiv \boldsymbol{\alpha}$ and $[\mathbf{t} \equiv \boldsymbol{\beta}] \equiv \boldsymbol{\beta}$ are tautologies, so by TAUTOLOGY $\models \mathbf{C}([[\boldsymbol{\alpha \equiv \beta}] \equiv \boldsymbol{\beta}]) \equiv \mathbf{C}(\boldsymbol{\alpha})$ and $\models \mathbf{C}(\mathbf{t} \equiv \boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta})$. Hence $\models \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})$. ■

We can now show that all contexts in L have the analogue of congruentiality for $\models^*$ in place of $\models$:

2.6. If $\models^* \boldsymbol{\alpha \equiv \beta}$ then $\models^* \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})$.

Proof: Suppose that $\models^* \boldsymbol{\alpha \equiv \beta}$. Let $\mathbf{C} = \langle \gamma, \mathbf{p} \rangle$. Consider any context $\mathbf{D} = \langle \boldsymbol{\delta}, \mathbf{q} \rangle$. Let $\mathbf{E} = \langle \mathbf{q} \equiv \mathbf{C}(\boldsymbol{\beta})), \mathbf{q} \rangle$, where $\mathbf{q}$ does not occur in $\mathbf{C}(\boldsymbol{\beta})$. For any context $\mathbf{D}$, by 2.5 $\models \mathbf{DEC}(\boldsymbol{\alpha}) \equiv \mathbf{DEC}(\boldsymbol{\beta})$. But $\mathbf{DEC}(\boldsymbol{\alpha}) = \mathbf{D}(\mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta}))$ and $\mathbf{DEC}(\boldsymbol{\beta}) = \mathbf{D}(\mathbf{C}(\boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta}))$. Hence $\models \mathbf{D}(\mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})) \equiv \mathbf{D}(\mathbf{C}(\boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta}))$. But $[\mathbf{C}(\boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta})] \equiv \mathbf{t}$ is a tautology, so by TAUTOLOGY $\models \mathbf{D}(\mathbf{C}(\boldsymbol{\beta}) \equiv \mathbf{C}(\boldsymbol{\beta})) \equiv \mathbf{D(t)}$. Hence $\models \mathbf{D}(\mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})) \equiv \mathbf{D(t)}$. Since D was an arbitrary context, $\models^* \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})$. ■

We can also show that $\models^*$ is the weakest strengthening of $\models$ to have features 2.1, 2.4 and 2.6.

2.7. Suppose that $\models^\wedge$ obeys the following principles for all sentences $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$:

(a^) If $\models^\wedge \boldsymbol{\alpha}$ then $\models \boldsymbol{\alpha}$

(b^) If $\boldsymbol{\alpha \equiv \beta}$ is a tautology then $\models^\wedge \boldsymbol{\alpha}$ iff $\models^\wedge \boldsymbol{\beta}$

(c^) If $\models^\wedge \boldsymbol{\alpha \equiv \beta}$ then $\models^\wedge \mathbf{C}(\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\beta})$

Then for all sentences $\boldsymbol{\alpha}$, if $\models^\wedge \boldsymbol{\alpha}$ then $\models^* \boldsymbol{\alpha}$.

Proof: Suppose that $\models^{\wedge} \alpha$. Since $\alpha \equiv [\alpha \equiv \mathbf{t}]$ is a tautology, $\models^{\wedge} \alpha \equiv \mathbf{t}$ by (b^). Thus $\models^{\wedge} \mathbf{C}(\alpha) \equiv \mathbf{C}(\mathbf{t})$ by (c^), so $\models \mathbf{C}(\alpha) \equiv \mathbf{C}(\mathbf{t})$ by (a^). Since $\mathbf{C}$ was arbitrary, $\models^{*} \alpha$. ■

We should also check that the smoothness condition of closure under uniform substitutions is preserved under the transition from $\models$ to $\models^{*}$.

2.8. If $\models^{*} \alpha$ then $\models^{*} \sigma(\alpha)$

Proof: Fix α and σ. Suppose that $\models^{*} \alpha$. Let $\mathbf{C} = \langle\gamma, \mathbf{p}\rangle$. Let S be the set of variables in α, $\sigma(\alpha)$, γ or $\mathbf{p}$. Let φ be a substitution that maps distinct variables in S to distinct variables not in S and is constant on variables not in S. Let ψ be a substitution such that if $\mathbf{q}$ is in S then $\psi(\varphi(\mathbf{q})) = \mathbf{q}$ and ψ is constant outside S. Let $\mathbf{C}^{*} = \langle\varphi(\gamma), \varphi(\mathbf{p})\rangle$. Let σ^{*} be the substitution such that $\sigma^{*}(\mathbf{q}) = \sigma(\mathbf{q})$ if $\mathbf{q}$ is in S and $\sigma^{*}(\mathbf{q}) = \mathbf{q}$ otherwise. Now for any β, $\sigma^{*}(\sigma_{\varphi(p)/\beta}(\varphi(\mathbf{p}))) = \sigma^{*}(\beta) = \sigma_{\varphi(p)/\sigma^{*}(\beta)}(\varphi(\mathbf{p}))$ and for any other variable $\mathbf{q}$ not in S, $\sigma^{*}(\sigma_{\varphi(p)/\beta}(\mathbf{q})) = \sigma^{*}(\mathbf{q}) = \mathbf{q} = \sigma_{\varphi(p)/\sigma^{*}(\beta)}(\mathbf{q})$. Hence for any δ with no variables in S, $\sigma^{*}(\sigma_{\varphi(p)/\beta}(\delta)) = \sigma_{\varphi(p)/\sigma^{*}(\beta)}(\delta)$. Since $\varphi(\gamma)$ has no sentence variables in S, $\sigma^{*}(\mathbf{C}^{*}(\beta)) = \sigma^{*}(\sigma_{\varphi(p)/\beta}(\varphi(\gamma))) = \sigma_{\varphi(p)/\sigma^{*}(\beta)}(\varphi(\gamma)) = \mathbf{C}^{*}(\sigma^{*}(\beta))$. But $\models \mathbf{C}^{*}(\alpha) \equiv \mathbf{C}^{*}(\mathbf{t})$, so by SUBSTITUTION $\models \psi(\sigma^{*}(\mathbf{C}^{*}(\alpha))) \equiv \psi(\sigma^{*}(\mathbf{C}^{*}(\mathbf{t})))$, so $\models \psi(\mathbf{C}^{*}(\sigma^{*}(\alpha))) \equiv \psi(\mathbf{C}^{*}(\sigma^{*}(\mathbf{t})))$. Now $\sigma^{*}(\mathbf{t}) = \mathbf{t}$ and $\sigma^{*}(\alpha) = \sigma(\alpha)$ since every variable in α is in S. Thus $\models \psi(\mathbf{C}^{*}(\sigma(\alpha))) \equiv \psi(\mathbf{C}^{*}(\mathbf{t}))$. Let β have no variables outside S. Then $\psi(\sigma_{\varphi(p)/\beta}(\varphi(\mathbf{p})) = \psi(\beta) = \beta = \sigma_{p/\beta}(\mathbf{p})$ and if $\mathbf{q}$ is any other variable in S $\psi(\sigma_{\varphi(p)/\beta}(\varphi(\mathbf{q}))) = \psi(\varphi(\mathbf{q})) = \mathbf{q} = \sigma_{p/\beta}(\mathbf{q})$. Therefore, since γ has no variables outside S, $\psi(\sigma_{\varphi(p)/\beta}(\varphi(\gamma))) = \sigma_{p/\beta}(\gamma)$. But $\psi(\mathbf{C}^{*}(\beta)) = \psi(\sigma_{\varphi(p)/\beta}(\varphi(\gamma)))$ and $\mathbf{C}(\beta) = \sigma_{p/\beta}(\gamma)$, so $\psi(\mathbf{C}^{*}(\beta)) = \mathbf{C}(\beta)$. In particular, since $\sigma(\alpha)$ and $\mathbf{t}$ have no variables outside S, $\psi(\mathbf{C}^{*}(\sigma(\alpha))) = \mathbf{C}(\sigma(\alpha))$ and $\psi(\mathbf{C}^{*}(\mathbf{t})) = \mathbf{C}(\mathbf{t})$, Therefore

$\models \mathbf{C}(\boldsymbol{\sigma}(\boldsymbol{\alpha})) \equiv \mathbf{C}(\mathbf{t})$. But **C** was arbitrary, so $\models^* \boldsymbol{\sigma}(\boldsymbol{\alpha})$. ■

The price of allowing non-extensional contexts, such as are created by counterfactual conditionals and modal operators, is that we lose some real world valid features of the logic of the "actually" operator:

2.9. If L has non-extensional contexts, then not $\models^* \mathbf{Ap} \equiv \mathbf{p}$.

Proof: Suppose that $\models^* \mathbf{Ap} \equiv \mathbf{p}$. Then for all $\boldsymbol{\alpha}$ $\models^* \mathbf{A}\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}$ by 2.8, so far any context **C**, $\models^* \mathbf{C}(\mathbf{A}\boldsymbol{\alpha}) \equiv \mathbf{C}(\boldsymbol{\alpha})$. It follows as in the proof of 1.5 that **C** is extensional. ■

Nevertheless, we can still obtain all the standard generally valid features of the logic of "actually" for the strengthened notion of validity (and a fortiori for the original weaker notion). In possible worlds terms, all those formulas that are true at all worlds in every model, not just at the actual world of every model, are still valid in the new sense.

2.10. $\models^* \neg\mathbf{Ap} \equiv \mathbf{A}\neg\mathbf{p}$

Proof: Let **C** be any context. Let $\mathbf{D} = \langle \neg\mathbf{q} \equiv \mathbf{A}\neg\mathbf{p}, \mathbf{q} \rangle$. Applying ACTUALITY to the context **CD**:

(10a) $\models \mathbf{C}(\neg\mathbf{Ap} \equiv \mathbf{A}\neg\mathbf{p}) \equiv [[\mathbf{p} \supset \mathbf{C}(\neg\mathbf{t} \equiv \mathbf{A}\neg\mathbf{p})] \;\&\; [\neg\mathbf{p} \supset \mathbf{C}(\neg\mathbf{f} \equiv \mathbf{A}\neg\mathbf{p})]]$

By TAUTOLOGY from (10a):

(10b) $\models \mathbf{C}(\neg\mathbf{Ap} \equiv \mathbf{A}\neg\mathbf{p}) \equiv [[\mathbf{p} \supset \mathbf{C}(\neg\mathbf{A}\neg\mathbf{p})] \;\&\; [\neg\mathbf{p} \supset \mathbf{C}(\mathbf{A}\neg\mathbf{p})]]$

Let $\mathbf{E} = \langle \neg\mathbf{q}, \mathbf{q} \rangle$. Applying ACTUALITY to $\neg\boldsymbol{\alpha}$ for **C** and **CE** respectively:

(10c) $\models \mathbf{C}(\mathbf{A}\neg\mathbf{p}) \equiv [[\neg\mathbf{p} \supset \mathbf{C}(\mathbf{t})] \;\&\; [\neg\neg\mathbf{p} \supset \mathbf{C}(\mathbf{f})]]$

(10d) $\models C(\neg A\neg p) \equiv [[\neg p \supset C(\neg t)] \ \&\ [\neg\neg p \supset C(\neg f)]]$

By TAUTOLOGY (10d) simplifies to:

(10e) $\models C(\neg A\neg p) \equiv [[\neg p \supset C(f)] \ \&\ [p \supset C(t)]]$

From (10b), (10c) and (10e):

(10f) $\models C(\neg Ap \equiv A\neg p) \equiv$

$$[[p \supset [[\neg p \supset C(f)] \ \&\ [p \supset C(t)]]] \ \&\ [\neg p \supset [[\neg p \supset C(t)] \ \&\ [p \supset C(f)]]]]$$

The right-hand side of (10f) is tautologously equivalent to $C(t)$. Hence (10f) gives:

(10g) $\models C(\neg Ap \equiv A\neg p) \equiv C(t)$

■

2.11. $\models^* A[p \supset q] \supset [Ap \supset Aq]$

Proof: Let C be any context. Let $D = \langle r \supset [Ap \supset Aq], r\rangle$. Applying ACTUALITY to $p \supset q$ for CD:

(11a) $\models C(A[p \supset q] \supset [Ap \supset Aq]) \equiv$

$$[[[p \supset q] \supset C(t \supset [Ap \supset Aq])] \ \&\ [\neg[p \supset q] \supset C(f \supset [Ap \supset Aq])]]$$

TAUTOLOGY from (11a) gives:

(11b) $\models C(A[p \supset q] \supset [Ap \supset Aq]) \equiv [[[p \supset q] \supset C(Ap \supset Aq)] \ \&\ [\neg[p \supset q] \supset C(t)]]$

Let $E = \langle r, r \supset Aq\rangle$. Applying ACTUALITY to CE:

(11c) $\models C(Ap \supset Aq) \equiv [[p \supset C(t \supset Aq)] \ \&\ [\neg p \supset C(f \supset Aq)]]$

By TAUTOLOGY from (11c):

(11d) $\models C(Ap \supset Aq) \equiv [[p \supset C(Aq)] \ \&\ [\neg p \supset C(t)]]$

By ACTUALITY again:

(11e) $\models C(Aq) \equiv [[q \supset C(t)] \ \&\ [\neg q \supset C(f)]]$

From (11d) and (11e):

(11f) $\models$ **C(Ap ⊃ Aq) ≡ [[p ⊃ [[q ⊃ C(t)] & [¬q ⊃ C(f)]]] & [¬p ⊃ C(t)]]**

Simplifying the right-hand side of (11f) by TAUTOLOGY:

(11g) $\models$ **C(Ap ⊃ Aq) ≡ [[p ⊃ q] ⊃ C(t)] & [¬[p ⊃ q] ⊃ C(f)]]**

From (11b) and (11g):

(11h) $\models$ **C(A[p ⊃ q] ⊃ [Ap ⊃ Aq]) ≡**

[[[p ⊃ q] ⊃ [[p ⊃ q] ⊃ C(t)] & [¬[p ⊃ q] ⊃ C(f)]]] & [¬[p ⊃ q] ⊃ C(t)]]

But the right-hand side of (11k) is tautologously equivalent to **C(t)**, so:

(11i) $\models$ **C(A[p ⊃ q] ⊃ [Ap ⊃ Aq]) ≡ C(t)**

■

2.12. $\models$* **AAp ≡ Ap**

Proof: Let **C** be any context. Let **D** = <**q ≡ Ap**, **q**>. Applying ACTUALITY to **Ap** for **CD**:

(12a) $\models$ **C(AAp ≡ Ap) ≡ [[Ap ⊃ C(t ≡ Ap)] & [¬Ap ⊃ C(f ≡ Ap)]]**

By (1.4) this simplifies to:

(12b) $\models$ **C(AAp ≡ Ap) ≡ [[p ⊃ C(t ≡ Ap)] & [¬p ⊃ C(f ≡ Ap)]]**

By TAUTOLOGY from (12b):

(12c) $\models$ **C(AAp ≡ Ap) ≡ [[p ⊃ C(Ap)] & [¬p ⊃ C(¬Ap)]]**

From 2.6 and 2.10:

(12d) $\models$ **C(¬Ap) ≡ C(A¬p)**

Thus from (12c) and (12d):

(12e) $\models$ **C(AAp ≡ Ap) ≡ [[p ⊃ C(Ap)] & [¬p ⊃ C(A¬p)]]**

Applying ACTUALITY to **p** and **¬p**:

(12f) $\models \mathbf{C(Ap) \equiv [[p \supset C(t)] \; \& \; [\neg p \supset C(f)]]}$

(12g) $\models \mathbf{C(A\neg p) \equiv [[\neg p \supset C(t)] \; \& \; [\neg\neg p \supset C(f)]]}$

From (12e)-(12g):

(12h) $\models \mathbf{C(AAp \equiv Ap) \equiv}$

$$\mathbf{[[p \supset [[p \supset C(t)] \; \& \; [\neg p \supset C(f)]]] \; \& \; [\neg p \supset [[\neg p \supset C(t)] \; \& \; [p \supset C(f)]]]]}$$

But the right-hand side of (12h) is tautologously equivalent to **C(t)**, so from (12h):

(12i) $\models \mathbf{C(AAp \equiv Ap) \equiv C(t)}$

■

2.13. If $\models^* \boldsymbol{\alpha}$ then $\models^* \mathbf{A}\boldsymbol{\alpha}$

Proof: Suppose that $\models^* \boldsymbol{\alpha}$. Then by (2.1) $\models \boldsymbol{\alpha}$. By ACTUALITY,

$\models \mathbf{C(A\alpha) \equiv [[\alpha \supset C(t)] \; \& \; [\neg\alpha \supset C(f)]]}$, so $\models \mathbf{C(A\alpha) \equiv C(t)}$. ■

<u>Adding a conditional</u>

Let us now assume that, in addition to the truth-functional conditional ⊃ , L contains a conditional →. Beyond the general assumptions already made, we make only these three specific to →:

CONSEQUENCE If $\models \boldsymbol{\alpha \supset \beta}$ then $\models \boldsymbol{\alpha \rightarrow \beta}$

DISTRIBUTION $\models$ **[p → [q & r]] ≡ [[p → q] & [p → r]]**

REDUCTIO $\models$ **[p → f] ⊃ ¬p**

These seem rather mild assumptions. CONSEQUENCE does not say that the material conditional entails the corresponding → conditional; it merely says that when the material conditional is *valid*, so that we can exclude the case of **α** without **β**, the corresponding → conditional is valid too. Almost all standard conditionals support DISTRIBUTION. REDUCTIO permits us to exclude whatever implies a contradiction in the sense of →. All three principles are derivable in the Lewis-Stalnaker systems for counterfactual conditionals with respect to their original languages. We could also consider stronger versions of the three constraints with $\models$* in place of $\models$, and use them to prove correspondingly strengthened versions of the results below, but for present purposes that is unnecessary. For these three constraints, combined with the four earlier ones, already suffice to show that → is extensionally equivalent to the material conditional.

We first note two elementary lemmas and then prove the main result, 3.3.

3.1. If **β** ⊃ **γ** is a tautology then $\models$ **[α → β]** ⊃ **[α → γ]**.

Proof: Suppose that **β** ⊃ **γ** is a tautology. Then **β** ≡ **[β & γ]** is a tautology. Choose **p** not to appear in **α**. Let **C** = <**α → p**, **p**>. By TAUTOLOGY, $\models$ **C(β)** ≡ **C(β & γ)**, in other words $\models$ **[α → β]** ≡ **[α → [β & γ]]**. Thus by DISTRIBUTION $\models$ **[α → β]** ⊃ **[α → γ]**. ■

3.2. If $[\beta \,\&\, \gamma] \supset \delta$ is a tautology then $\models [[\alpha \rightarrow \beta] \,\&\, [\alpha \rightarrow \gamma]] \supset [\alpha \rightarrow \delta]$.

Proof: By 3.1 and DISTRIBUTION. ■

3.3. $\models [p \rightarrow q] \equiv [p \supset q]$.

Proof: By 1.4, $\models p \supset [Aq \equiv q]$, so by CONSEQUENCE:

(4a) $\models p \rightarrow [Aq \equiv q]$

But since $[q \,\&\, [Aq \equiv q]] \supset Aq$ and $[Aq \,\&\, [Aq \equiv q]] \supset q$ are tautologies, by 3.2:

$\models [[p \rightarrow q] \,\&\, [p \rightarrow [Aq \equiv q]]] \supset [p \rightarrow Aq]$ and

$\models [[p \rightarrow Aq] \,\&\, [p \rightarrow [Aq \equiv q]]] \supset [p \rightarrow q]$, so by (4a):

(4b) $\models [p \rightarrow q] \equiv [p \rightarrow Aq]$

Let $C = \langle p \rightarrow r, r \rangle$. Applying ACTUALITY to C with q for p:

(4c) $\models [p \rightarrow Aq] \equiv [[q \supset [p \rightarrow t]] \,\&\, [\neg q \supset [q \rightarrow f]]]$

Since $p \supset t$ is a tautology, by TAUTOLOGY and CONSEQUENCE $\models p \rightarrow t$. Hence from (4c):

(4d) $\models [p \rightarrow Aq] \equiv [\neg q \supset [p \rightarrow f]]$

From (4b) and (4d):

(4e) $\models [p \rightarrow q] \equiv [\neg q \supset [p \rightarrow f]]$

By truth-functional reasoning, the right-to-left direction of (4e) yields:

(4f) $\models q \supset [p \rightarrow q]$

From the left-to-right direction of (4e) and REDUCTIO $\models [p \rightarrow q] \supset [\neg q \supset \neg p]$, so:

(4g) $\models [p \rightarrow q] \supset [p \supset q]$

By CONSEQUENCE, $\models p \rightarrow p$, so putting p for q in (4e) yields:

(4h) $\models \neg p \supset [p \rightarrow f]]$

But **f** $\supset$ **q** is a tautology, so by 3.1 $\models$ **[p → f]** $\supset$ **[p → q]**, so from (4h):

(4i) $\models$ **¬p** $\supset$ **[p → q]**

Collecting (4f), (4g) and (4i) together:

(4j) $\models$ **[p → q]** $\equiv$ **[p** $\supset$ **q]**

■

What is the informal significance of 3.3, on either a subjunctive or an indicative reading of →?

On a subjunctive reading of → and rigidifying reading of **A**, the problem is clear. DISTRIBUTION and REDUCTIO are unproblematic, but CONSEQUENCE cannot hold with full generality on the operative reading of $\models$ as real world validity (1.4). For **p** $\supset$ **Ap** is real world valid, whereas **p → Ap** fails on a subjunctive reading when **p** is contingently false, as was noted above in respect of (1s) and (2s). Indeed, ACTUALITY by itself tells us that $\models$ **¬q** $\supset$ **[[p → Aq]** $\equiv$ **[p → f]]**: subjunctively implying a rigidified falsehood is equivalent to subjunctively implying a contradiction. On this reading, the failure of CONSEQUENCE is just like the failure of the rule of necessitation in modal logic for real world validity: although **p** $\supset$ **Ap** is real world valid, □**(p** $\supset$ **Ap)** is not. Like necessitation, consequence should be restricted to general validity: if $\models$* **α** $\supset$ **β** then $\models$* **α → β**.

Let us switch to an indicative reading of →. If it is read truth-functionally, no problem arises. CONSEQUENCE becomes trivial and DISTRIBUTION and REDUCTIO are obvious. But what happens if we try to read → as a non-truth-functional conditional, supposed to fit the indicative use of "if" in English? CONSEQUENCE still looks compelling. If we apply it to $\mathbf{p \supset Ap}$, the case that made trouble for the subjunctive conditional, we still conclude that $\mathbf{p \rightarrow Ap}$ is real world valid, but we have already noted that indicative conditionals of that form, such as (1i) and (2i), seem to be logically trivial.[4] Indeed, one might expect real world validity to be the standard of validity best adapted to an indicative conditional.

Caution is required, for 3.3 does not imply that $\mathbf{p \rightarrow q}$ embeds in the same way as $\mathbf{p \supset q}$ in non-truth-functional contexts. For by 1.4 $\models \mathbf{[p \rightarrow q] \equiv A[p \supset q]}$ too, but $\mathbf{p \supset q}$ and $\mathbf{A[p \supset q]}$ embed differently in modal contexts (only the latter entails its own necessitation). However, many of the standard objections to the truth-functional reading of "if" extend to 3.3, because they concern cases that do not involve problematic embeddings. For instance, we often seem to be much more confident of the negation of $\mathbf{p}$ than of an indicative conditional $\mathbf{p \rightarrow q}$; that is hard to reconcile with 3.3. Thus 3.3 shows the language L to constitute an extremely hostile environment for a non-extensional indicative conditional.

What is the source of the hostility? Without ACTUALITY, nothing like 3.3 is derivable, since all the other principles hold on the interpretation of $\rightarrow$ as the strict conditional in a modal language with $\models$ as general validity. Indeed, even this variant of ACTUALITY is generally valid:

ACTUALITY* $\qquad \models \mathbf{C(A\alpha) \equiv [[A\alpha \supset C(t)] \ \& \ [\neg A\alpha \supset C(f)]]}$

Given the other initial constraints, ACTUALITY is equivalent to the conjunction of 1.4 with ACTUALITY*, for ACTUALITY and ACTUALITY* are trivially equivalent in the presence of 1.4 and the other constraints. ACTUALITY* depends on the rigidifying effect of **A** but not on the privileged role with respect to validity of the point of evaluation to which it rigidifies.

The treatment of $\models$ as real world validity, associated with 1.4, is not wholly uncontentious, since some philosophers regard general validity as a better candidate than real world validity for the analysis of the informal notion of validity; they want validity to have a modal dimension. However, this concern is orthogonal to the issue about non-extensional indicative conditionals, since even if general validity is the preferred candidate, real world validity is still a perfectly intelligible notion, whether or not we dignify it with the title "validity". Moreover, simple real world valid sentences such as (1i) and (2i) still obviously hold. Thus a preference for general validity would not block the proof of 3.3 for the technical notion $\models$; we could still conclude on broadly logical grounds that $\boldsymbol{\alpha} \rightarrow \boldsymbol{\beta}$ is extensionally equivalent to $\boldsymbol{\alpha} \supset \boldsymbol{\beta}$. Since defenders of a non-extensional indicative reading of → want to avoid that conclusion, the heart of the issue must concern the treatment of embedded occurrences of sentences of the form **Aα**, as in ACTUALITY*, irrespective of the choice between real world and general validity.

Accounts of indicative "if" as a non-extensional conditional typically assign it some kind of doxastic or epistemic meaning. We should therefore consider instances of ACTUALITY* involving doxastic or epistemic contexts. For instance, let **Bel** be a belief operator and **p** say that it is pouring. An instance of ACTUALITY* is:

(!) $\models \mathbf{Bel(Ap)} \equiv \mathbf{[[Ap \supset Bel(t)] \; \& \; [\neg Ap \supset Bel(f)]]}$

According to (!), if it is actually pouring, then it is believed that it is actually pouring if and only if a tautology is believed; if it is not actually pouring, then it is believed that it is actually pouring if and only if a contradiction is believed. If the truth of **Bel(α)** depends on the agent's reaction to the sentence **α**, provided that the agent understands **α**, then (!) is hopeless, even as evaluated with respect to the actual world. If it is pouring but I do not believe that it is pouring, I can assent to a tautology without having any inclination to assent to the sentence "It is actually pouring"; thus the right-hand side of (!) may hold while the left-hand side fails. If it is not pouring but I falsely believe that is pouring, I can assent to the sentence "It is actually pouring" without having any inclination to assent to a contradiction; thus the left-hand side of (!) may hold while the left-hand side fails. These problems for (!) depend simply on error and ignorance about the weather, not about logic; they are quite compatible with logical omniscience. Thus a defender of a non-extensional indicative conditional might take it to create a doxastic or epistemic context of such a kind for which ACTUALITY* would fail.[5]

However, standard frameworks for formal semantics enable one to introduce a rigidifying operator such as **A** by force of stipulation; thus one cannot simply assert that ACTUALITY* must fail. Instead, one should use ACTUALITY* to understand the effect of such an operator. For instance, one uses the right-hand side of (!) to understand its left-hand side, rather than assuming that one understands its left-hand side independently. If **p** is true, then **Ap** expresses a tautologous proposition, and what it takes to believe **Ap** is to believe a tautology; if **p** is false, then **Ap** expresses a contradictory proposition and what it takes to believe **Ap** is to believe a contradiction. The operator **A** is indexical because what proposition **Ap** expresses depends on the world in which one is speaking, even if what proposition **p** expresses does not so depend. On this understanding, the problem is after all not with the treatment of **A**, but with the possibility of its unforeseen repercussions for sentences in which **A** interacts with →. In particular, what looks like a harmless principle about the conditional might turn out to be implausible when properly understood. For a toy example, give **α → β** the crude doxastic reading **Bel(α ⊃ β)**, where the believer is assumed to be rational. In particular, **p → Ap** is interpreted as **Bel(p ⊃ Ap)**. If **p** is actually false, that requires **Bel(p ⊃ f)** (in effect, **Bel(¬p)**) by ACTUALITY*, which may fail on account of the rational agent's ignorance of the weather. On this toy doxastic reading, **p → Ap** would fail, and with it CONSEQUENCE for real world validity.[6]

The problem is that it is very hard, perhaps impossible, to hear ordinary indicative conditionals of the form **p → Ap** as failing in any such way. No ignorance of the weather threatens the validity of (1i) and (2i), however much one focuses on a rigidifying reading of "actually". Thus it is very doubtful that indicative conditionals are doxastic or epistemic in the sort of way required to block the derivation of 3.3, in particular to invalidate CONSEQUENCE.

In the present state of understanding, it would be premature to draw any conclusion too confidently for the prospects for a non-extensional reading of indicative "if". Nevertheless, the possibility of a rigidifying operator poses a threat to any such reading that must be taken seriously.

Notes

* Thanks to participants in the 2008 First Formal Epistemology Festival in Konstanz for stimulating comments on an earlier version of this material, and to Franz Huber for written comments.

1 Stalnaker uses an example from Anderson (1951) to make the point (Stalnaker 1999: 71 in the reprinted version).

2 The present paper extends the investigation of such issues in Williamson 2006. There are also relevant remarks at Williamson 2007: 137-141, 144-145, 152-153, 295-296.

3 One can observe that sentences such as (1i) and (2i) are truisms prior to taking any theoretical stance on whether indicative conditionals are truth-functional.

4 Of course, we do not expect $\models^{*} \mathbf{p} \rightarrow \mathbf{Ap}$, since the conditional may have a true antecedent and a false consequent with respect to a counterfactual circumstance of evaluation.

5 Such an account of propositional attitudes would be incompatible with Stalnaker's; his is coarse-grained and gives much less weight to the agent's reaction to a

sentence that expresses the proposition at issue, even when the agent understands the sentence.

6 The agent's assumed rationality allows us to ignore objections to TAUTOLOGY that concern agents who are not omniscient about truth-functional logic.

Bibliography

Anderson, A.R. 1951. "A note on subjunctive and counterfactual conditionals". *Analysis* 12: 35-38.

Davies, M., and Humberstone, L. 1980. "Two notions of necessity". *Philosophical Studies* 38: 1-30.

Humberstone, L. 1986. "Extensionality in sentence position". *Journal of Philosophical Logic* 15: 27-54.

Humberstone, L. 1997. "Singulary extensional connectives: a closer look". *Journal of Philosophical Logic* 26: 341-356.

Jackson, F. 1981. "Conditionals and possibilia". *Proceedings of the Aristotelian Society* 81: 125-137.

Jackson, F. 1987. *Conditionals*. Oxford: Blackwell.

Stalnaker, R.C. 1975. "Indicative conditionals". *Philosophia* 5.

Stalnaker, R.C. 1999. *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford: Oxford University Press.

Williamson, T. 2006. "Indicative versus subjunctive conditionals, congruential versus non-hyperintensional contexts". *Philosophical Issues* 16: 310-333.

Williamson, T. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.