

Logic Informed

Justin Bledin*

November 9, 2014

Abstract: Do logically valid arguments necessarily preserve truth? Certain inferences involving informational modal operators and indicative conditionals suggest that truth preservation and good deductive argument come apart. Given this split, I recommend an alternative to the standard truth preservation view of logic on which validity and good deductive argument coincide: logic is a descriptive science that is fundamentally concerned not with the preservation of truth, but with the preservation of structural features of information. Along the way, I defend *modus ponens* for the indicative against an attack by Kolodny and MacFarlane [2010], and I present a new proof system, **Info**, appropriate to this informational view.

1 Rival Conceptions of Logic

What is logic? Two answers by Frege dominate contemporary discussion.

First: Logic is a descriptive science, a body of truths about *truth*. Frege [1918] opens his essay *Thoughts* with this characteristic passage:¹

*Published in *Mind* 123(490): 277–316, 2014. This research is the core contribution of my same-titled dissertation written at the University of California, Berkeley. My biggest debt of gratitude is to my team of advisors: John MacFarlane, Niko Kolodny, and Seth Yalcin. They inspired and sharpened the ideas in this paper and are, in many respects, model philosophers. For extended discussions and comments on earlier drafts, let me also thank Adam Bradley, Michael Caie, Fabrizio Cariani, Kenny Easwaran, Melissa Fusco, Ethan Jerzak, Jeffrey Kaplan, Wesley Holliday, Matthew Ritchie, Justin Vlasits, the editor and reviewers for *Mind*, and audiences at UC Berkeley, Johns Hopkins University, and University College London.

¹See also Frege [1879-1891], p. 3, and [1897], p. 128.

Just as ‘beautiful’ points the way for aesthetics and ‘good’ for ethics, so do words like ‘true’ for logic. All sciences have truth as their goal; but logic is also concerned with it in a quite different way: logic has much the same relation to truth as physics has to weight or heat. To discover truths is the task of all sciences; it falls to logic to discern the laws of truth...Here of course it is not a matter of what happens but of what is. (p. 351)

In keeping with this characterization, most logicians and philosophers today define logical validity as follows: where $\varphi_1, \dots, \varphi_n, \psi$ are sentences in a formal or informal language, the argument from $\varphi_1, \dots, \varphi_n$ to ψ is logically valid if and only if it is impossible for each of $\varphi_1, \dots, \varphi_n$ to be true and for ψ to be false by virtue of their logical form.²

Second: Logic is a normative science, a body of rules that govern our thinking. Frege [1893] offers this other characterization in his *Grundgesetze der Arithmetik*:³

It will be granted by all at the onset that the laws of logic ought to be guiding principles for thought in the attainment of truth, yet this is only too easily forgotten, and here what is fatal is the double meaning of the word ‘law’. In one sense a law asserts what is; in the other it prescribes what ought to be. Only in the latter sense can the laws of logic be called ‘laws of thought’: so far as they stipulate the way in which one ought to think. Any law asserting what is, can be conceived as prescribing that one ought to think in conformity with it, and is thus in that sense a law of thought. This holds for laws of geometry and physics no less than for laws of logic. The latter have a special title to the name ‘laws of thought’ only if we mean to assert that they are the most general laws, which prescribe universally the way in which one ought to think if one is to think at all. (p. 12)

²Likewise, φ is a logical truth iff it is impossible for φ to be false by virtue of its logical form, $\varphi_1, \dots, \varphi_n$ are logically consistent iff it is possible for $\varphi_1, \dots, \varphi_n$ to be jointly true by virtue of their logical form, and so on.

The rider ‘by virtue of their logical form’ in the definition of logical validity is meant to preclude arguments like this from counting as valid:

(P1) Alfred is a bachelor.

(C) Alfred is an unmarried man.

But drawing a principled divide between the logical and non-logical expressions of a language is a difficult open problem in the philosophy of logic. See MacFarlane [2009] for a good critical survey of various approaches to this demarcation problem.

³See also Frege [1893], pp. 14, 15, and [1897], pp. 128, 146, 149.

Whereas Euclid's principles are in one sense norms for thinking about the geometry of space, and similarly the laws of physics are norms for thinking about the physical world, the laws of logic are constitutive norms for thought *as such*.⁴ That is, something not subject to the laws of logic just isn't thinking. Though some philosophers will surely balk at this constitutivity claim, most still think that logic is normative for thought in some form or another.⁵

For Frege, of course, these two conceptions of logic are compatible. The above quote from *Thoughts* continues,

From the laws of truth there follow prescriptions about asserting, thinking, judging, inferring. (p. 351)

Thus, on what I think is the most natural reading of Frege, logical laws have descriptive content about truth. They say nothing about how we ought to reason or what we ought to believe. However, the laws of logic still imply prescriptions for thought and related intentional activity. In this regard too, Frege's view remains popular. In fact, something like his picture of logic has a good claim to being the 'standard view' on the contemporary philosophical scene. Logically valid arguments, it is said, necessarily preserve truth. But fans of truth preservation needn't deny that logic has normative import, nor that the normative component of logic can play an important role in demarcating it from closely related disciplines like geometry and psychology. They must only recognize that logical laws are not themselves norms for thought. Bridge principles are needed at the logic-epistemology interface.⁶

But can we hold onto this standard conception of logic? Is logic a descriptive science of truth that gives rise to norms for thought? After

⁴cf. MacFarlane [2002].

⁵For example, Barwise and Etchemendy [1999] write on the first page of their popular logic textbook:

All rational inquiry depends on logic, on the ability of people to reason correctly most of the time, and, when they fail to reason correctly, on the ability of others to point out the gaps in their reasoning.

They continue a few pages later:

Rational inquiry, in our sense, is not limited to academic disciplines, and so neither are the principles of logic. If your beliefs about a close friend logically imply that he would never spread rumors behind your back, but you find out that he has, then your beliefs need revision. Logical consequence is central, not only to the sciences, but to virtually every aspect of everyday life.

⁶That such bridge principles are required is one of the key lessons of Gilbert Harman's [1986] *Change in View*. Here are a couple of candidate synchronic principles that bind at a moment in time t :

clarifying what I take to be the standard truth preservation view in §2, I consider some inferences in §3 that militate against it. Given what one of our best formal semantic theories says about the semantic values of informational modals and the indicative conditional, some good deductive arguments involving these *informational constants*—that is, deductive arguments that we can appropriately make in both categorical and hypothetical deliberative contexts by virtue of their logical form—are arguably not truth preserving. We seem, as I discuss in §4, to face a difficult choice: we can either maintain that these good deductive arguments are valid, and abandon the view that even *material* truth preservation is a necessary condition for validity; or we can maintain that logic is about truth preservation, and undermine the connection between validity and good deductive argument.⁷

In fact, I argue that we can take the first fork with few tears. Inspired by Yalcin [2007] in particular, I propose in §5 that logic is not the science of which arguments necessarily preserve truth, but rather a descriptive science whose proper subject is *information*.⁸ Facts about validity, on

Consistency: Where $\varphi_1, \dots, \varphi_n$ are logically inconsistent, you are rationally required at t (either not to believe that φ_1 is true, ..., or not to believe that φ_n is true).

Closure: Where $\varphi_1, \dots, \varphi_n$ logically imply ψ , you are rationally required at t (either not to believe that φ_1 is true, ..., not to believe that φ_n is true, or to believe that ψ is true).

Harman himself argues that such bridge principles linking logic and “reasoned change in view” are all problematic, so he negatively concludes that logic is not normative for thought in any special sense. However, see MacFarlane [*ms.*], and Field [2009a] and [*ms.*] for carefully formulated closure principles that avoid Harman’s objections.

⁷To my mind, ‘good deductive argument’ is best understood as an *evaluative* notion applicable in the third-person standpoint of appraisal, not a *normative* notion applicable in the first-person standpoint of deliberation and the related second-person standpoint of advice. It is one thing to know which inferences we can appropriately make in our reasoning. It is another thing to know what we ought to believe at a time or how we ought to revise our beliefs over time. However, I think good deductive arguments nevertheless play a strong normative role in our linguistic and epistemic practices, so the tension between truth preservation and good argument is also one between Frege’s two conceptions of logic as a descriptive science of truth and a normative science.

⁸Let me stress, though, that while Yalcin defends an informational consequence relation over natural language, he does not explicitly endorse an informational view of *logical* consequence. Any conceptual errors associated with the shift from natural language semantics to logic are my own.

I am certainly not the first to propose an intimate connection between logic and information. At Stanford in the 1980s and 1990s, Jon Barwise and his colleagues developed the information-based *situation semantics*. For some relevant work in this tradition, see Barwise [1989], Barwise and Etchemendy [1990], and Devlin [1991]. At the Institute for Logic, Language, and Computation in Amsterdam, Veltman

this informational view, tell us about the structure of the bodies of information that we generate, encounter, absorb, and exchange as we interact with one another and learn about our world. Roughly put, the argument from $\varphi_1, \dots, \varphi_n$ to ψ is logically valid if and only if information with the structural features (to be made precise in due course) corresponding to each of $\varphi_1, \dots, \varphi_n$ also has the structural feature corresponding to ψ by virtue of the logical form of these sentences.

This informational concept of validity lines up with the standard concept defined in terms of truth preservation over simple languages without informational modals and the indicative conditional. But the informational concept also dovetails with good deductive argument over richer languages that include these informational constants—or so I will argue. In §6 and §7, I consider cases of defective argumentation that purportedly show that *modus ponens* for the indicative—a valid form of argument on the informational view—is unreliable in hypothetical contexts. I argue that these cases show no such thing; rather, they show that methods of argumentation involving hypothetical reasoning like *reductio ad absurdum* and constructive dilemma do not always fit the classical model when performed in a language with informational constants. I conclude in §8 in favor of something close to Frege’s picture: logic is a descriptive science that informs what we ought to do and believe. It is just that logical facts describe not relations of truth, but instead structural properties of information.

2 Truth Preservation

I use the definite description ‘the truth preservation view’ to denote a cluster of widespread intuitions about the *informal* concept of logical validity.⁹ The most basic intuition, of course, is that a logically valid

[1996], van Benthem [2011], and other logicians and philosophers have also proposed dynamic information-based conceptions of logic. Charting this ‘informational turn’ in logic is a historical project that I plan to pursue in the future. For the moment, let me just say that the informational view presented in this paper is more closely tied to work in the Amsterdam school.

⁹Tarski [1936] opens his seminal essay on logical consequence with this remark on the subject of mathematical logic:

The concept of logical consequence is one whose introduction into the field of strict formal investigation was not a matter of arbitrary decision on the part of this or that investigator; in defining this concept, efforts were made to adhere to the common usage of the language of everyday life. (p. 409)

Let me be clear that what I have in mind here is not the “common usage” of people on the street (I am doubtful that this “common usage” even exists), but the informal

argument with true premises has a true conclusion. But two further intuitions sharpen this core condition. The first is that validity involves a modal element: it is *impossible* for each of the premises of a logically valid argument to be true and for the conclusion to be false. The second is that a logically valid argument preserves truth *by virtue of the logical form* of the sentences in the argument, and not due to the meaning of any non-logical symbols. These intuitions all come together in the definition from §1: the argument from $\varphi_1, \dots, \varphi_n$ to ψ is logically valid if and only if it is impossible for each of $\varphi_1, \dots, \varphi_n$ to be true and for ψ to be false by virtue of their logical form.¹⁰

In mathematical logic, this informal notion is typically analyzed in terms of *truth in a model*.¹¹ Open up just about any logic textbook and you will see something like this: the argument from $\varphi_1, \dots, \varphi_n$ to ψ is logically valid if and only if there is no model \mathcal{M} for formal language \mathcal{L} such that the translations of $\varphi_1, \dots, \varphi_n$ into \mathcal{L} are all true in \mathcal{M} but the translation of ψ into \mathcal{L} is false in \mathcal{M} , where \mathcal{L} purportedly makes the logical form of $\varphi_1, \dots, \varphi_n, \psi$ explicit, and a model \mathcal{M} for \mathcal{L} is, roughly, something that provides enough information to determine the extensions of all well formed sentences $S_{\mathcal{L}}$ of this formal language.¹² In sentential logic, a model is (basically) just a reference row of a truth table. In classical first-order logic, a model is a non-empty domain of individuals and an interpretation function that maps constants in \mathcal{L} to individuals in the domain and maps predicates in \mathcal{L} to sets of individuals in the domain. In intuitionistic logic, a model is a Kripke tree with valuations at each node. And so forth. By varying our formal languages and models, we have generated a large family of formal characterizations of logical validity.

But while logicians and philosophers agree on which arguments count as logically valid on this or that formal characterization, they disagree

concept of logical validity that has been influential in philosophy and mathematics since Aristotle.

¹⁰Even among logicians and philosophers who accept this definition, there is room for disagreement. Is the modality alethic, metaphysical, or epistemic? Are higher order quantifiers, say, part of logical form? ‘The truth preservation view’ is best regarded as an umbrella term that covers the many ways to make this definition precise.

Relevance logicians like Anderson, Belnap, and Dunn [1992] also insist that the premises of a logically valid argument must be relevant to its conclusion. However, they argue that this requirement of relevance is not a separate virtue but is actually required to ensure a kind of truth preservation. See Lewis [1982] for good discussion.

¹¹Besides this model-theoretic semantic approach, logicians also study syntactic characterizations of validity in proof systems.

¹²For a cleaner exposition, I will often use $\varphi_1, \dots, \varphi_n, \psi$ to designate both sentences in English and their translations into a formal language.

on which of the formal notions explicate the *genuine* informal notion of logical validity. That is, they disagree on which of the formal notions extensionally coincide with our pre-theoretic, intuitive notion of validity. Having an extensionally adequate formal explication of validity would certainly be useful. By investigating it, we could learn things about our target, the informal notion.

For example, assume that \mathcal{L} is the basic sentential modal language. Then a mathematical explication of logical validity is typically going to look something like this:

A model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, @, \mathcal{V} \rangle$ for \mathcal{L} consists of a set of possible worlds \mathcal{W} where one of the worlds $@ \in \mathcal{W}$ is designated the actual one, a binary accessibility relation $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$ between worlds, and an interpretation function $\mathcal{V} : At_{\mathcal{L}} \times \mathcal{W} \mapsto \{T, F\}$ that maps each sentence letter $p \in At_{\mathcal{L}}$ and world $w \in \mathcal{W}$ to a truth value in $\{T, F\}$. A recursive specification of truth uses \mathcal{R} and \mathcal{V} to establish the complete interpretation function $\llbracket \cdot \rrbracket_{\mathcal{M}} : S_{\mathcal{L}} \times \mathcal{W} \mapsto \{T, F\}$ for \mathcal{L} that maps each sentence $\varphi \in S_{\mathcal{L}}$ and world $w \in \mathcal{W}$ to a truth value.¹³ Sentence φ is true at w in \mathcal{M} if and only if $\llbracket \varphi \rrbracket_{\mathcal{M}}^w = T$. This sentence is true in \mathcal{M} if and only if $\llbracket \varphi \rrbracket_{\mathcal{M}}^@ = T$. Logical validity is defined in terms of truth in \mathcal{M} :

Def 1. The argument from $\varphi_1, \dots, \varphi_n$ to ψ is valid, $\{\varphi_1, \dots, \varphi_n\} \models \psi$, just in case there is no model \mathcal{M} where $\llbracket \varphi_1 \rrbracket_{\mathcal{M}}^@ = \dots = \llbracket \varphi_n \rrbracket_{\mathcal{M}}^@ = T$ and $\llbracket \psi \rrbracket_{\mathcal{M}}^@ = F$.¹⁴

It remains to add at least the identity relation ‘=’ and the standard first-order quantifiers ‘ \forall ’ and ‘ \exists ’ to \mathcal{L} . But the hope, again, is that the formal relation ‘ \models ’ extensionally captures the informal notion of logical validity over the fragment of our language without these symbols—that is, $\{\varphi_1, \dots, \varphi_n\} \models \psi$ just in case it is impossible for each of $\varphi_1, \dots, \varphi_n$ to be true and for ψ to be false by virtue of their logical form (here I use ‘true’ and ‘false’ in an ordinary, pre-theoretic sense, and not in the technical sense relativized to a model). If this equivalence holds, then investigating the laws governing ‘ \models ’ can teach us about the target intuitive notion and, given the purported tight link between logic and epistemology, about what we ought to believe.

¹³See any modal logic textbook for details. See also n. 21 below.

¹⁴If $[\varphi]_{\mathcal{M}} =_{def} \{w \in \mathcal{W} : \llbracket \varphi \rrbracket_{\mathcal{M}}^w = T\}$ is the *proposition* expressed by sentence φ in \mathcal{M} , then $\{\varphi_1, \dots, \varphi_n\} \models \psi$ just in case there is no \mathcal{M} where $[\varphi_1]_{\mathcal{M}} \cap \dots \cap [\varphi_n]_{\mathcal{M}} \not\subseteq [\psi]_{\mathcal{M}}$. The relation ‘ \models ’ between a set of sentences and a sentence in $S_{\mathcal{L}}$ corresponds to a set-theoretic relation between the propositions expressed by these sentences.

3 Against Truth Preservation

However, there are deductive arguments involving informational modal operators and indicative conditionals that put pressure on this standard view of logic. These arguments suggest that truth preservation and good deductive argument come apart.

For example, consider this variant of Vann McGee’s [1985] famous ‘counterexample’ to *modus ponens* based on the 1980 U.S. Presidential election:

(P1) If a married woman committed the murder, then if Mrs. Peacock didn’t do it, it was Mrs. White.

(P2) A married woman committed the murder.

(C) If Mrs. Peacock didn’t do it, it was Mrs. White.

Pace McGee, I think that we can appropriately make this argument in both categorical and hypothetical deliberative contexts.¹⁵ Publicly, if someone asserts or supposes that if a married woman did it then if it wasn’t Mrs. Peacock it was Mrs. White, and also asserts or supposes that a married woman did it, then we can infer on this basis that if it wasn’t Mrs. Peacock it was Mrs. White.¹⁶ Privately, if you activate your beliefs or simply suppose in an episode of internal theoretical deliberation that the conditional and its antecedent both hold, then you can infer that the consequent holds.¹⁷

Nevertheless, the inference from (P1) and (P2) to (C) arguably fails to preserve truth. Let me elaborate.¹⁸

Assume that in addition to sentence letters A, B, C, \dots, \perp , the logical connectives $\neg, \vee, \wedge, \supset, \equiv$, and parentheses $()$, the symbols of \mathcal{L} also include the indicative conditional \Rightarrow and *informational* necessity \square and possibility \diamond operators (these modals are typically called *epistemic*

¹⁵I defend this claim in §6 and §7.

¹⁶I use ‘infer’ in a thin sense. Inference consists of recognizing what follows; it needn’t culminate in belief.

¹⁷This is *not* to say that in categorical deliberative contexts involving assertion and belief activation you should come to believe that if it wasn’t Mrs. Peacock it was Mrs. White. Perhaps you believe that (P1) holds or that (P2) holds in the face of strong evidence to the contrary. Still, the *modus ponens* inference can shed important light on the normativity of your situation—for instance, that you ought either not to believe both that a married woman did it *and* if a married woman did it then if it wasn’t Mrs. Peacock it was Mrs. White, or to believe that if it wasn’t Mrs. Peacock then it was Mrs. White.

¹⁸Going forward, I will be loose about use and mention, omitting Quinean quasi-quotes and the like.

modals since sentences involving them are often evaluated relative to what someone knows or believes).¹⁹

To give a semantics for \mathcal{L} , I follow Yalcin [2007] and [2011], and Kolodny and MacFarlane [2010] in adding another argument place to the interpretation function $\llbracket \cdot \rrbracket_{\mathcal{M}} : S_{\mathcal{L}} \times \mathcal{W} \mapsto \{T, F\}$ from §2. Sentences in $S_{\mathcal{L}}$ are now evaluated for truth relative both to a world $w \in \mathcal{W}$ and to an *information state* $i \in 2^{\mathcal{W}}$ (a set of worlds in \mathcal{W}). A model \mathcal{M} for \mathcal{L} will still consist partly of a function $\mathcal{V} : At_{\mathcal{L}} \times \mathcal{W} \mapsto \{T, F\}$ that interprets each sentence letter $p \in At_{\mathcal{L}}$, but a recursive specification of truth at an *index* $\langle w, i \rangle$ now lifts this function to the full interpretation function $\llbracket \cdot \rrbracket_{\mathcal{M}} : S_{\mathcal{L}} \times \mathcal{W} \times 2^{\mathcal{W}} \mapsto \{T, F\}$ mapping $\varphi \in S_{\mathcal{L}}$, $w \in \mathcal{W}$, and $i \in 2^{\mathcal{W}}$ to a truth value.²⁰

The clauses for sentence letters, \perp , and the logical connectives are straightforward:

$$\begin{aligned} \llbracket p \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \mathcal{V}(p, w) = T \\ \llbracket \perp \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } 0 = 1 \\ \llbracket \neg\varphi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = F \\ \llbracket \varphi \vee \psi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = T \text{ or } \llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = T \\ \llbracket \varphi \wedge \psi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = T \text{ and } \llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = T \\ \llbracket \varphi \supset \psi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = F \text{ or } \llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = T \\ \llbracket \varphi \equiv \psi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = \llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} \end{aligned}$$

However, the semantic clauses for the informational modal operators and indicative conditional are more interesting since the information state parameter i in the index comes into action:

$$\begin{aligned} \llbracket \Box\varphi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \forall v \in i (\llbracket \varphi \rrbracket_{\mathcal{M}}^{v,i} = T) \\ \llbracket \Diamond\varphi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \exists v \in i (\llbracket \varphi \rrbracket_{\mathcal{M}}^{v,i} = T)^{21} \\ \llbracket \varphi \Rightarrow \psi \rrbracket_{\mathcal{M}}^{w,i} = T & \text{ iff } \forall v \in i + \varphi (\llbracket \psi \rrbracket_{\mathcal{M}}^{v,i+\varphi} = T) \end{aligned}$$

¹⁹The catalogue of logical constants is controversial. But I assume in what follows, along with many logicians and philosophers, that the informational modal operators and indicative conditional make the list.

²⁰Lewis [1980], drawing on Kaplan [1989], argues that a sentence φ in a natural language must be evaluated relative both to the *context* c in which φ is used, and to an *index*, an n -tuple of features of context. But given the formal language \mathcal{L} under consideration, context dependence and much index dependence can be ignored; the index $\langle w, i \rangle$ will do.

²¹The main difference between this compositional semantics and the standard textbook semantics for \Box and \Diamond involving an accessibility relation $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$ is effectively that the information state parameter i in the index $\langle w, i \rangle$ is independent of

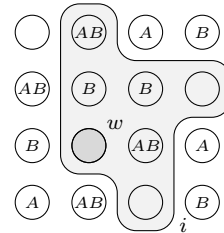


Fig 1. An index $\langle w, i \rangle$

where $i + \varphi$ appearing in the clause for the indicative is the largest subset $i' \subseteq i$ such that $\forall w \in i' (\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i'} = T)$.²² Intuitively: ‘Colonel Mustard must have done it’ is true at index $\langle w, i \rangle$ just in case all of the worlds in i are worlds in which Colonel Mustard did it. ‘Colonel Mustard might have done it’ is true at index $\langle w, i \rangle$ just in case Colonel Mustard did it at some of these worlds. And ‘If Colonel Mustard did it then he used the candlestick’ is evaluated for truth at index $\langle w, i \rangle$ by first ‘adding’ the information that Colonel Mustard did it to the existing stock of information i and then ‘checking’ whether Colonel Mustard used the candlestick at all worlds in this updated information state. Given the index $\langle w, i \rangle$ in Fig 1, for example, $\llbracket \Box A \rrbracket_{\mathcal{M}}^{w,i} = F$, $\llbracket \Diamond A \rrbracket_{\mathcal{M}}^{w,i} = T$, and $\llbracket A \Rightarrow B \rrbracket_{\mathcal{M}}^{w,i} = T$.

What about the formal logical consequence relation? This is where things get more interesting—there are multiple *prima facie* attractive

the world parameter w . On the textbook semantics, the clauses for the informational modals can be stated as follows:

$$\begin{aligned} \llbracket \Box \varphi \rrbracket_{\mathcal{M}}^w = T & \quad \text{iff} \quad \forall v \in \{v : wRv\} (\llbracket \varphi \rrbracket_{\mathcal{M}}^v = T) \\ \llbracket \Diamond \varphi \rrbracket_{\mathcal{M}}^w = T & \quad \text{iff} \quad \exists v \in \{v : wRv\} (\llbracket \varphi \rrbracket_{\mathcal{M}}^v = T) \end{aligned}$$

²²A few asides on the semantics of the indicative conditional:

—Yalcin [2007] requires that $i + \varphi \neq \emptyset$ but I follow Kolodny and MacFarlane [2010] and relax this restriction.

—If the antecedent φ of an indicative conditional can include informational modals, then $i + \varphi$ is not well-defined since there needn’t be a unique *maximal* φ -subset $i' \subseteq i$ such that $\forall w \in i' (\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i'} = T)$ and $\exists w \in i'' (\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i''} = F)$ for each i'' such that $i' \subset i'' \subseteq i$. For example, when $i = \{w_1, w_2\}$, w_1 is an $A \wedge \neg B$ -world, and w_2 is a $\neg A \wedge B$ -world, both $\{w_1\}$ and $\{w_2\}$ are candidates for $i + \Box A \vee \Box B$. To handle this non-uniqueness, Kolodny and MacFarlane [2010] present the following alternative clause for the indicative:

$$\llbracket \varphi \Rightarrow \psi \rrbracket_{\mathcal{M}}^{w,i} = T \quad \text{iff} \quad \forall i' \in i_{\varphi} (\forall v \in i' (\llbracket \psi \rrbracket_{\mathcal{M}}^{v,i'} = T))$$

where i_{φ} is the set of maximal φ -subsets of i . But for ease of exposition, I will just work with the simpler clause in the main text since I consider only conditionals with non-modal antecedents.

—Like Yalcin [2007], Kolodny and MacFarlane [2010], and Gillies [2010], but unlike Lewis [1975] and Kratzer [1986], I am effectively treating ‘if’ as an *operator* rather than a *restrictor* (note that the clause for $\varphi \Rightarrow \Box \psi$ involves double quantification whereas a restrictor-theoretic clause would involve only single quantification). Doing so allows me to avoid the need for covert informational necessity modals in the semantic clause for bare indicatives like ‘If Colonel Mustard did it then he used the candlestick’ that lack an explicit modal in their consequent (cf. Kratzer [1986]). However, I will not defend the operator approach here since I do not think that the difference between operators and restrictors matters much for present purposes given the simple conditionals under consideration. (For some sense of what a defense would require, see the forceful objections to Gillies’ operator-based semantics in Khoo [2011].)

options for \models in this setting.

Running with the truth preservation view of logic, a natural first suggestion is that logically valid arguments preserve truth at *all* indices in all models:

Def 2. The argument from $\varphi_1, \dots, \varphi_n$ to ψ is valid, $\{\varphi_1, \dots, \varphi_n\} \models_0 \psi$, just in case there is no model \mathcal{M} such that for some world $w \in \mathcal{W}$ and information state $i \in 2^{\mathcal{W}}$, $\llbracket \varphi_1 \rrbracket_{\mathcal{M}}^{w,i} = \dots = \llbracket \varphi_n \rrbracket_{\mathcal{M}}^{w,i} = T$ and $\llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = F$.

However, Def 2 seems too demanding. This is a good argument:

(P1) Mrs. Peacock must have done it.

(C) Mrs. Peacock did it.

But $\{\Box A\} \not\models_0 A$, since when w_1 is an A -world but w_2 is not, truth is not preserved at the index $\langle w_2, \{w_1\} \rangle$. Similarly, this is good:

(P1) Mrs. Peacock did it.

(C) Mrs. Peacock might have done it.

But $\{A\} \not\models_0 \Diamond A$, since truth is not preserved at the index $\langle w_1, \{w_2\} \rangle$.

One diagnosis of this badness is that we are considering far too many indices, many of which— $\langle w_2, \{w_1\} \rangle$, $\langle w_1, \{w_2\} \rangle$, for example—cannot correspond to an actual or possible agent’s situation and knowledge state.²³ This suggests the following fix:

Def 3. The argument from $\varphi_1, \dots, \varphi_n$ to ψ is valid, $\{\varphi_1, \dots, \varphi_n\} \models_{Tr} \psi$, just in case there is no model \mathcal{M} such that for some information state $i \in 2^{\mathcal{W}}$ and world $w \in i$, $\llbracket \varphi_1 \rrbracket_{\mathcal{M}}^{w,i} = \dots = \llbracket \varphi_n \rrbracket_{\mathcal{M}}^{w,i} = T$ and $\llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = F$.

The relation \models_{Tr} preserves truth at all and only proper indices $\langle w, i \rangle$ where $w \in i$ in all models.²⁴ So $\{\Box \varphi\} \models_{Tr} \varphi$ and $\{\varphi\} \models_{Tr} \Diamond \varphi$ for all $\varphi \in S_{\mathcal{L}}$, and \models_{Tr} supports our judgments that these implications are good.

However, \models_{Tr} also invalidates some good deductive arguments. For example, we can appropriately make this argument in both categorical and hypothetical contexts:

(P1) Professor Plum didn’t do it.

²³cf. Kolodny and MacFarlane [2010], drawing on Kaplan [1989]. If a context c supplies an information state i_c —the knowledge state of some contextually salient agent—that always includes the world w_c of this context—given the *factivity* of knowledge—then one might insist that Def 2 excessively requires the preservation of truth, not only at *proper* indices corresponding to some possible c , but at indices not corresponding to any context.

²⁴Since we often investigate what is so according to nonfactual information that rules out the actual state of the world, this restriction to indices $\langle w, i \rangle$ where $w \in i$ strikes me as a bad move. However, this is not a point that I want to press here.

(C) It is not the case that Professor Plum might have done it.²⁵

But $\{\neg A\} \not\models_{Tr} \neg \Diamond A$, since when w_1 is an A -world but w_2 is not, truth is not preserved at the proper index $\langle w_2, \{w_1, w_2\} \rangle$ (see Fig 2).

Further, this argument is good:

(P1) Either Mrs. White did it or Miss Scarlett did it.

(P2) Miss Scarlett didn't do it.

(C) Mrs. White must have done it.²⁶

But $\{A \vee B, \neg B\} \not\models_{Tr} \Box A$, since when w_1 is an $A \wedge \neg B$ -world and w_2 is a $\neg A$ -world, truth is not preserved at the proper index $\langle w_1, \{w_1, w_2\} \rangle$ (see Fig 3).

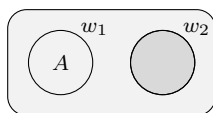


Fig 2. The index $\langle w_2, \{w_1, w_2\} \rangle$

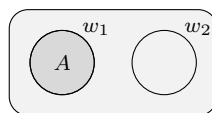


Fig 3. The index $\langle w_1, \{w_1, w_2\} \rangle$

Lastly, the *modus ponens* argument that opened this section is good:

(P1) If a married woman committed the murder, then if Mrs. Peacock didn't do it, it was Mrs. White.

(P2) A married woman committed the murder.

(C) If Mrs. Peacock didn't do it, it was Mrs. White.

But $\{A \Rightarrow (\neg B \Rightarrow C), A\} \not\models_{Tr} \neg B \Rightarrow C$, since when w_1 is an $A \wedge B \wedge \neg C$ -world, w_2 is an $A \wedge \neg B \wedge C$ -world, and w_3 is a $\neg A \wedge \neg B \wedge \neg C$ -world, this argument does not preserve truth at the proper index $\langle w_1, \{w_1, w_2, w_3\} \rangle$ (see Fig 4).

²⁵Yalcin [2007] calls $\{\neg\varphi\} \models \neg\Diamond\varphi$ “Łukasiewicz’s Principle” since Łukasiewicz [1930] appears to endorse it.

²⁶One might object that this argument only seems good because of the modal element in the standard characterization of logical validity—whenever the premises are true, the conclusion *must* be true—and the validity of disjunctive syllogism. But the following good argument is also invalidated by \models_{Tr} :

(P1) The murder occurred in the library and either Mrs. White did it or Miss Scarlett did it.

(P2) Miss Scarlett didn't do it.

(C) The murder occurred in the library and Mrs. White must have done it.

It is harder to hear the embedded ‘must’ in (C) as the modality of logical validity.

The upshot: If validity is understood in terms of necessary truth preservation and the formal relation \models_{Tr} explicates this informal target notion, then these arguments reveal that validity and good deductive argument do not line up.

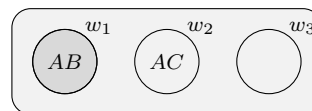


Fig 4. The index $\langle w_1, \{w_1, w_2, w_3\} \rangle$

4 Responses

How to respond to this mismatch? One might, of course, just dismiss my positive evaluations of the previous three arguments as misguided. However, my intuition that these are good arguments is widely shared.²⁷

Alternatively, one might agree that these are good arguments and maintain the truth preservation view of logic, but concede that \models_{Tr} is not up to the job. However, it is unclear what a better explication of logical validity, at least as traditionally understood in terms of necessary truth preservation at a context, could look like that validates all of these arguments.²⁸

Changing tack, one might accept the separation of truth preservation and good deductive argument and come out on one or the other side of the fence. This response might already have come to mind:

“Why insist that all good deductive arguments are logically valid? Robert Stalnaker [1975] famously distinguishes between the semantic concept of *entailment* and the pragmatic concept of *reasonable inference*. The *orto-if* inference from ‘Either the butler did it or the gardener did it’ to ‘If the butler didn’t do it then the gardener did’, says Stalnaker, is logically invalid since the former sentence does not even semantically entail the latter, but it is still a reasonable inference. Similarly, one might say that the inference from ‘The butler didn’t do it’ to ‘It is not the case that the butler might have done it’ is logically invalid but still a good

²⁷Again, it is not universally shared. Kolodny and MacFarlane [2010], for instance, argue that *modus ponens* can lead us astray in hypothetical contexts. I respond to their argument in §6.

²⁸See Bledin [2013], Ch. 3, for a more comprehensive discussion of the threat that informational constants pose to the truth preservation view. There are more sophisticated versions of this view that validate all of the arguments in §3. However, there is a deeper worry with the truth preservation view of logic in general. The worry is that sentences involving informational modals and the indicative conditional are not, properly speaking, even true or false at a context (cf. Adams [1975], Edgington [1986], and Yalcin [2011]). On this *nonfactualist* position, the truth preservation view seems thoroughly ill-suited to explain the goodness of deductive arguments involving these informational constants.

inference.”²⁹

Of course, we might just stipulate ‘logically valid’ is to mean something like ‘necessarily truth preserving by virtue of logical form’ and accept that many good deductive arguments fall outside its extension. However, this move is unattractive. Though the idea that logically valid arguments necessarily preserve truth by virtue of logical form is well entrenched in the philosophical tradition, so too is the idea that good deductive arguments are logically valid. We maintain one of these ideas only by seriously undermining another.

This response might have also come to mind:

“Why not surrender the idea that logically valid arguments preserve truth? Perhaps we can understand logic in some other way such that validity and good deductive argument coincide.”

Hartry Field [2006], [2008], [2009a], [2009b], [*ms.*], for one, has argued for this approach in recent work.³⁰ He boldly suggests that we should

²⁹I bring up Stalnaker’s distinction between semantic entailment and reasonable inference to stimulate this response, but I should emphasize that, as I understand these concepts, good deductive inference and reasonable inference do not coincide. Inspired by Stalnaker’s original definition, let us call an inference from $\varphi_1, \dots, \varphi_n$ to ψ a *reasonable inference* iff there is no model \mathcal{M} such that for some $i \in 2^{\mathcal{W}}$ one can appropriately assert φ_1 against background information i, \dots , and also appropriately assert φ_n against $i + \varphi_1 + \dots + \varphi_{n-1}$, but $\exists w \in i + \varphi_1 + \dots + \varphi_n (\llbracket \psi \rrbracket_{\mathcal{M}}^{w, i + \varphi_1 + \dots + \varphi_n} = F)$. The good inferences in §3 are all reasonable given the semantics under consideration. However, while the *or-to-if* inference from $A \vee \Diamond A$ to $\neg A \Rightarrow \Diamond A$ is reasonable given the Gricean assumption that one can appropriately assert $\varphi \vee \psi$ against background information i only if $\exists w \in i (\llbracket \varphi \wedge \neg \psi \rrbracket_{\mathcal{M}}^{w, i} = T)$ and $\exists w \in i (\llbracket \neg \varphi \wedge \psi \rrbracket_{\mathcal{M}}^{w, i} = T)$ (since it is *never* appropriate to assert $A \vee \Diamond A$), this inference is not good; one does poorly to make it in categorical and hypothetical contexts.

³⁰Field thinks that we should abandon the truth preservation view not because of the inferences discussed in §3, but because of various inferences involving a general untyped truth predicate $Tr(x)$. If we look at our best formal truth theories developed since the 1970s to handle the Liar paradox and related semantic paradoxes, Field argues, these theories formulated in a language with $Tr(x)$ include axioms that they do not regard as true, or rules of inference that they do not regard as unrestrictedly truth preserving. Worse, adding to truth theory T either the sentence saying that all of T ’s axioms are true or the sentence saying that all of T ’s rules of inference preserve truth results in inconsistency. But all of us accept, or should accept, one or the other of these theories, taking its axioms/rules to govern our inferential practices. So if we want to hold onto the idea that logic—where ‘logic’ is broad enough to include the logic of $Tr(x)$ —lines up with good deductive inference such that these axioms/rules are logical truths/logically valid, then we are not in a position to consistently accept that logic is about unrestricted truth preservation.

In this essay, however, I am less concerned with Field’s truth-theoretic argument than with his normative characterization of logic. There are reasons aside from the semantic paradoxes to think that truth preservation and good deductive argument

regard the normative dimension of logic as fundamental:

If logic is not the science of what [forms of inference] necessarily preserve truth, it is hard to see what the subject of logic could possibly be, if it isn't somehow connected to norms of thought. (Field [2009a], p. 263)

This is not to say that logical validity should be *defined* in terms of its normativity for thought. Field finds this tack repugnant, and argues that it is best not to define logical validity at all but to treat it as a primitive notion and illuminate its conceptual role. Field [*ms.*] suggests the following role for validity:

To regard an inference or argument as valid is (in large part anyway) to accept a constraint on belief: one that prohibits fully believing its premises without fully believing its conclusion. (The prohibition should be ‘due to logical form’: for any other argument of that form, the constraint should also prohibit fully believing the premises without fully believing the conclusion.) (p. 11)³¹

Moreover, there is nothing more to understanding logical validity than understanding what it is to regard an argument as valid. If you and I disagree over whether *modus ponens* for the indicative is valid, then one of us accepts a constraint on belief that the other rejects—that one ought either not to fully believe in φ , not to fully believe in $\varphi \Rightarrow \psi$, or to fully believe in ψ . But there needn't be any ultimate metaphysical basis for accepting one view over the other. In the course of our dispute, of course, you or I might appeal to certain objective facts to back up a position. You might argue for the belief constraint, say, by explaining that if φ and $\varphi \Rightarrow \psi$ are both true, then ψ is true as well. But the arguments in §3 show that this kind of justification is not always available. When it is not, our differing judgments about validity will simply reflect the different norms that we take to govern our epistemic practices.³²

However, this Fieldian position also involves radically abandoning orthodoxy. The arguments in §3 invite us to revisit what we talk about when we talk about logic. But, once we do so, it is not obvious why we should conclude as Field [2009a] does that “logic is essentially normative”

come apart.

³¹This proposal is refined in Field [2009a], §I, and [*ms.*], §2. But I will just focus on the cruder proposal stated in the main text.

³²Field [*ms.*], §3, motivates this “projectivism” about logical validity by analogy to the concept of chance.

(p. 268). Not only must we give up the standard truth preservation view of logic, we must give up the more basic idea that logic is a descriptive science. This seems drastic. Is there a better option?

5 A Better Option

There is an alternative conception of logic that Field overlooks. Like the truth preservation view, this neglected alternative takes logic to be a descriptive science—indeed, it can help us understand the widespread attraction of truth preservation. But on this alternative, validity also coincides with good deductive argument.

The new characterization of logic is motivated by the semantics in §3. Note that the compositional semantics in that section suggests another formal consequence relation:³³

Def 4. The argument from $\varphi_1, \dots, \varphi_n$ to ψ is valid, $\{\varphi_1, \dots, \varphi_n\} \models_I \psi$, just in case there is no model \mathcal{M} such that for some information state $i \in 2^{\mathcal{W}}$, $\forall w \in i (\llbracket \varphi_1 \rrbracket_{\mathcal{M}}^{w,i} = \dots = \llbracket \varphi_n \rrbracket_{\mathcal{M}}^{w,i} = T)$ and $\neg \forall w \in i (\llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = T)$.

Unlike the relations in Def 2 and Def 3, the ‘informational consequence’ relation \models_I does not preserve truth at an index $\langle w, i \rangle$ in all models, whether or not $w \in i$.³⁴ Instead, facts about \models_I correspond to facts about information states. \models_I holds between a set of sentences and a single sentence in $S_{\mathcal{L}}$ just in case all information states in all models

³³In fact, the semantics suggests even more options. For example, Willer [2012] proposes this “dynamic logical consequence” relation:

$\langle \varphi_1, \dots, \varphi_n \rangle \models_D \psi$ just in case there is no model \mathcal{M} such that for some $i \in 2^{\mathcal{W}}$ and world $w \in i$, $\llbracket \varphi_1 \rrbracket_{\mathcal{M}}^{w,i} = \dots = \llbracket \varphi_n \rrbracket_{\mathcal{M}}^{w, i \boxplus \dots \boxplus \varphi_{n-1}} = T$ and $\llbracket \psi \rrbracket_{\mathcal{M}}^{w, (i \boxplus \dots \boxplus \varphi_{n-1}) \boxplus \varphi_n} = F$

where $i \boxplus \varphi =_{def} i \cap \{w : \llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = T\}$. The relation \models_D validates the arguments from §3: $\langle \Box A \rangle \models_D A$, $\langle A \rangle \models_D \Diamond A$, $\langle \neg A \rangle \models_D \neg \Diamond A$, $\langle A \vee B, \neg B \rangle \models_D \Box A$, and $\langle A \Rightarrow (\neg B \Rightarrow C), A \rangle \models_D \neg B \Rightarrow C$. But this relation also has some strange, non-standard structural properties. First, \models_D is sensitive to the order of premises: $\langle \neg A, \Diamond A \rangle \models_D A$ but $\langle \Diamond A, \neg A \rangle \not\models_D A$ (hence my use of ordered sequences). Also, \models_D is not right monotonic: $\langle \Diamond A \rangle \models_D \Diamond A$ but $\langle \Diamond A, \neg A \rangle \not\models_D \Diamond A$. So I focus on a more conservative relation. (Willer actually regards the non-monotonicity of \models_D as a virtue since it allows us to “disentangle *modus ponens* from *modus tollens*” and take the former to be valid and the latter to be invalid. But this is a poor motivation for \models_D since non-monotonicity is not really required to disentangle the two *modi*. More on this in n. 37 below.)

³⁴This is Yalcin’s [2007] terminology. Veltman [1996] proposes a similar formal consequence relation \models_3 defined over his dynamic ‘update semantics’. Also, Kolodny and MacFarlane [2010] consider a variant of \models_I . When $\{\varphi_1, \dots, \varphi_n\} \models_I \psi$ but $\{\varphi_1, \dots, \varphi_n\} \not\models_{Tr} \psi$, they say that the argument from $\varphi_1, \dots, \varphi_n$ to ψ is “quasi-valid” but not “valid”.

have a particular kind of *structure*.

Let me explain. Every sentence $\varphi \in S_{\mathcal{L}}$ corresponds to a potential feature $\triangleright \varphi$ of information states: $i \triangleright \varphi$ just in case $\forall w \in i(\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = T)$. For instance, $i \triangleright A$ just in case $\forall w \in i(\llbracket A \rrbracket_{\mathcal{M}}^{w,i} = T)$; that is, $i \triangleright A$ just in case each world in i is an A -world. Also, $i \triangleright \Diamond B$ just in case $\forall w \in i(\llbracket \Diamond B \rrbracket_{\mathcal{M}}^{w,i} = T)$; that is, $i \triangleright \Diamond B$ just in case some world in i is a B -world or i is empty. And so on. Def 4 can thus be restated as follows:

$\{\varphi_1, \dots, \varphi_n\} \models_I \psi$ just in case there is no model \mathcal{M} such that for some $i \in 2^{\mathcal{W}}$, $i \triangleright \varphi_1, \dots, i \triangleright \varphi_n$, and $i \not\triangleright \psi$.

Intuitively, ψ is an informational consequence of $\{\varphi_1, \dots, \varphi_n\}$ just in case any information state in any model with the structure determined by each of $\varphi_1, \dots, \varphi_n$ (consisting entirely of A -worlds, containing at least one B -world, and so forth) has the structure determined by ψ . For convenience, let us say that when $i \triangleright \varphi$, φ is ‘incorporated’ in i .³⁵ We can then think of informational consequence as preserving not truth at an index, but incorporation in all information states.³⁶

Now, \models_I to its credit respects more of our implication judgments than \models_{Tr} . It is easy to verify that \models_I validates the good deductive arguments in §3 that give \models_{Tr} a rough time: $\{\neg A\} \models_I \neg \Diamond A$, $\{A \vee B, \neg B\} \models_I \Box A$, and $\{A \Rightarrow (\neg B \Rightarrow C), A\} \models_I \neg B \Rightarrow C$.³⁷ Certainly, the entailment

³⁵Yalcin [2007] says that when $i \triangleright \varphi$, φ is “accepted” in i . But I worry that ‘acceptance’ talk can suggest—at least to readers of Stalnaker [1984]—that $i \in 2^{\mathcal{W}}$ must explicate the informational content of someone’s propositional attitudes, and not information that would exist even if we did not.

³⁶Admittedly, Def 4 still involves the technical notion of truth at an index and might be thought to motivate a kind of *global* truth preservation view on which logical validity necessarily preserves truth across clusters of contexts (see Bledin [2013], Ch. 3, for more discussion). But as already mentioned in n. 28 above, nonfactualism about the informational fragment of our language spells trouble for all versions of the truth preservation view.

³⁷ $i \triangleright \neg \varphi$ iff $\forall w \in i(\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = F)$. $i \triangleright \neg \Diamond \varphi$ iff $\forall w \in i(\llbracket \Diamond \varphi \rrbracket_{\mathcal{M}}^{w,i} = F)$. Since $i \triangleright \neg \varphi$ implies $i \triangleright \neg \Diamond \varphi$, $\{\neg \varphi\} \models_I \neg \Diamond \varphi$.

$i \triangleright \varphi \vee \psi$ iff $\forall w \in i(\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = T$ or $\llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = T)$. $i \triangleright \neg \psi$ iff $\forall w \in i(\llbracket \psi \rrbracket_{\mathcal{M}}^{w,i} = F)$. $i \triangleright \Box \varphi$ iff $\forall w \in i(\llbracket \varphi \rrbracket_{\mathcal{M}}^{w,i} = T)$. Since $i \triangleright \varphi \vee \psi$ and $i \triangleright \neg \psi$ together imply $i \triangleright \Box \varphi$, $\{\varphi \vee \psi, \neg \psi\} \models_I \Box \varphi$.

$i \triangleright \varphi \Rightarrow \psi$ iff $\forall w \in i + \varphi(\llbracket \psi \rrbracket_{\mathcal{M}}^{w,i+\varphi} = T)$ iff $i + \varphi \triangleright \psi$. $i \triangleright \varphi$ iff $i = i + \varphi$. Since $i \triangleright \varphi \Rightarrow \psi$ and $i \triangleright \varphi$ together imply $i \triangleright \psi$, $\{\varphi \Rightarrow \psi, \varphi\} \models_I \psi$.

It is commonly thought that *modus ponens* and exportation trade off against each other (cf. McGee [1985]) so it is perhaps not too surprising that exportation for the indicative is invalidated by \models_I : $\{(\varphi \wedge \psi) \Rightarrow \xi\} \not\models_I \varphi \Rightarrow (\psi \Rightarrow \xi)$ for some $\varphi, \psi, \xi \in S_{\mathcal{L}}$. For instance, $\{((A \vee \Diamond A) \wedge \neg A) \Rightarrow \Diamond A\} \not\models_I (A \vee \Diamond A) \Rightarrow (\neg A \Rightarrow \Diamond A)$. I suspect that it is a bit more surprising that *modus tollens* for the indicative is invalidated by \models_I : $\{\varphi \Rightarrow \psi, \neg \psi\} \not\models_I \neg \varphi$ for some $\varphi, \psi \in S_{\mathcal{L}}$. Consider this argument based on Lewis Carroll’s [1894] barbershop paradox (cf. MacFarlane and Kolodny [2010], p. 140):

$\{A\} \models_I \Box A$ is slightly odd, but I do not think that it threatens the link between incorporation preservation and good deductive argument. I have found that many who feel that the argument from ‘Colonel Mustard did it’ to ‘Colonel Mustard must have done it’ is odd also feel that the argument from ‘Either Colonel Mustard did it or Professor Plum did it’ and ‘Professor Plum didn’t do it’ to ‘Colonel Mustard must have done it’ is fine. So I suspect that the perceived oddness of the argument from A to $\Box A$ stems from its redundancy, not from its unreliability.³⁸

But the relation \models_I does not preserve truth at an index and so is not a possible explication of an informal consequence relation that preserves truth at a context in the ordinary sense. Insofar as we want logical validity to line up with good deductive argument, then, informational consequence pushes us to reconsider the informal concept of validity itself. Why is the argument valid from ‘Professor Plum didn’t do it’ to ‘It is not the case that Professor Plum might have done it’? Informational consequence suggests an attractive alternative to truth preservation: the argument is logically valid because information has a particular kind of structure. Information incorporating that Professor Plum didn’t do it rules out the possibility that he did it, and is therefore also information incorporating that it is not the case that Professor Plum might have done it.³⁹ Crucially, the suggestion is *not* that this argument is valid

(P1) If Colonel Mustard is not in the study, then if Professor Plum is not in the study, Reverend Green is.

(P2) It is not the case that if Professor Plum is not in the study, then Reverend Green is.

(C) It is not the case that Colonel Mustard is not in the study.

$\{\neg A \Rightarrow (\neg B \Rightarrow C), \neg(\neg B \Rightarrow C)\} \not\models_I \neg\neg A$, since when w_1 is an $A \wedge \neg B \wedge \neg C$ -world and w_2 is a $\neg A \wedge \neg B \wedge C$ -world, $\{w_1, w_2\} \triangleright (\neg A \Rightarrow (\neg B \Rightarrow C)) \wedge \neg(\neg B \Rightarrow C)$ but $\{w_1, w_2\} \not\triangleright \neg\neg A$. However, \models_I delivers the right verdict here since this argument is no good; we do badly to make it in categorical and hypothetical contexts. See Yalcin [2012] for some more bad *modus tollens* arguments that do not preserve incorporation.

³⁸I consider a more sophisticated worry with this argument in n. 49.

³⁹I am using ‘information’ here in the rough intuitive sense of that which eliminates certain possibilities while leaving others open, and not in the technical sense of a set of possible worlds $i \in 2^{\mathcal{W}}$. As we should distinguish the technical notion of truth at an index in our formal semantic theories from the ordinary pre-theoretic notion of truth, we should also distinguish the technical notion of incorporation relating a sentence $\varphi \in S_{\mathcal{L}}$ and a formal mathematical object $i \in 2^{\mathcal{W}}$ from the ordinary pre-theoretic sense in which a body of information incorporates, say, that Colonel Mustard might have done it. That is, we should distinguish the theoretical fact that the sentence ‘Colonel Mustard might have done it’ is incorporated in information state i from the pre-theoretic fact that a particular body of information is information according to which Colonel Mustard might have done it. You might think of the informal notion of incorporation as the core concern of logic understood informationally: logically

because it preserves truth. It is not this: ‘Professor Plum didn’t do it’ logically implies ‘It is not the case that Professor Plum might have done it’ because it is impossible for the former sentence to be true and for the latter to be false by virtue of their logical form. Instead, the proposal is this: ‘Professor Plum didn’t do it’ logically implies ‘It is not the case that Professor Plum might have done it’ because any body of information (the content of an eyewitness’s utterances, the evening news, and so forth) according to which Professor Plum didn’t do it is therefore, by virtue of logical form, also information according to which it is not the case that he might have done it.⁴⁰

Note that we are, in a sense, turning the study of logic on its head. The core concern of mathematical logic is commonly taken to be the formulation of a formal concept of logical validity that extensionally coincides with the informal concept of validity cashed out in terms of truth preservation. But now it is the formal relation \models_I that spurs us to reconsider the informal concept itself. Underlying the distinction between \models_{Tr} and \models_I is, I am suggesting, a deeper distinction between two different ways of thinking about the target informal notion of logical validity: there is the truth preservation view, and also what I will call the ‘informational view’ on which logic is fundamentally concerned with the structure of information.

Fortunately, the informational view allows us to make sense of most of the things that have been said about logic over the years. We can still regard logic as a descriptive science. We can even make sense of the considerable popularity of the truth preservation view itself. Restricted to sentences in $S_{\mathcal{L}}$ sans informational modal operators and the indicative conditional, $\{\varphi_1, \dots, \varphi_n\} \models_{Tr} \psi$ if and only if $\{\varphi_1, \dots, \varphi_n\} \models_I \psi$.⁴¹ So it is little wonder on the informational view that truth preservation has been so influential.

Further, we can maintain the idea that logical validity and good deductive argument coincide. I will conclude that a deductive argument counts as logically valid on the informational view if and only if we can appropriately make it in both categorical and hypothetical deliberative

valid arguments preserve incorporation in bodies of information, logical truths are incorporated in all bodies of information, logically consistent sentences are jointly incorporated in some body of information, and so on.

⁴⁰This suggestion also applies to arguments validated by \models_{Tr} . ‘Mrs. Peacock did it’ logically implies ‘Mrs. Peacock might have done it’ because information according to which Mrs. Peacock did it is therefore, by virtue of logical form, also information according to which she might have done it.

⁴¹The left-to-right direction clearly holds for the full language \mathcal{L} . The right-to-left direction holds for sentences without informational constants since $\llbracket \varphi \rrbracket^{w,i} = T$ iff $\{w\} \triangleright \varphi$ for such sentences.

contexts by virtue of logical form. But before that, let me respond to some serious threats to this equivalence.

6 *Reductio*

I have claimed that *modus ponens* for the indicative conditional—a valid form of inference on the informational view—is a good form of deductive argument. That is, we do well to employ it in our inferential practices. My claim requires defense.

Kolodny and MacFarlane [2010] argue that *modus ponens* is reliable inside categorical deliberative contexts but it can lead you astray inside hypothetical contexts triggered by supposition. Suppose you have been in your office for hours with the blinds down and have not heard the weather forecast. Given your evidence, you reflect that the streets might not be wet, but that if it is raining the streets must be wet. You then enter a hypothetical context by supposing that it is raining. Applying *modus ponens* in this hypothetical context, you infer that the streets must therefore be wet. Recognizing that this conflicts with the first premise, you then conclude by *reductio* that it is not raining. Something has gone horribly wrong.

1	The streets might not be wet	Premise		
2	If it is raining, the streets must be wet	Premise		
3	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-left: 1px solid black; padding-left: 10px; vertical-align: top;">It is raining</td> <td style="padding-left: 20px; vertical-align: top;">Supposition</td> </tr> </table>	It is raining	Supposition	
It is raining	Supposition			
4	The streets must be wet	From 2,3		
5	It is not the case that the streets must be wet	From 1		
6	⊥	From 4,5		
7	It is not raining	From 3,4-6		

Kolodny and MacFarlane pinpoint the use of *modus ponens* at step 4 as the root of the trouble. If they are right, then this argument form tells against the informational view of logic. Whereas before it seemed that the truth preservation view undergenerates by failing to count good deductive inferences as valid, it now seems that the informational view overgenerates by counting bad deductive inferences as valid.

However, is *modus ponens* really unreliable inside the hypothetical context? It seems to me that Kolodny and MacFarlane misdiagnose the problem, which arises only at step 5 or step 7, depending on how your

supposition works at step 3. Consider the *informational background* of your deliberation. Sitting in the office, your information leaves open the possibility that the streets are not wet, but rules out that it is raining and the streets are not wet. If your supposition consists in tentatively *adding* the information that it is raining to your basic information, then your nondegenerate information in the induced hypothetical context rules out the possibility that the streets are not wet, so you should not reflect at step 5 that it is not the case that the streets must be wet.⁴²

On the other hand, your supposition might trigger a hypothetical context in which your salient body of information rules out that it is not raining but also still has the structural features corresponding to both premises.⁴³ In this case, you do well to recognize the contradiction. But while your information in the hypothetical context is degenerate—it can be explicated by the empty set \emptyset —you *cannot* conclude from this that it is not raining. It follows only that your actual information does not rule out that it is not raining. That is, you can conclude that it *might* not be raining. Going forward, I assume that your supposition works in this second way, so the problem with the *reductio* is just that your conclusion is overly strong.

In a close variation of this example that also involves the use of *modus*

⁴²The root of Kolodny and MacFarlane’s misdiagnosis, on this proposal, is their assimilation of good deductive argumentation to reasoning according to a natural deduction system that allows unrestricted use of premises within a subproof. Though he does not discuss this particular example, Willer [2012] would also pin the blame on step 5. He argues that we must restrict what can be used inside hypothetical contexts when working with informational modals and the indicative because logical consequence is non-monotonic (recall his dynamic relation \models_D from n. 33). Note that $\langle \diamond \neg W, R \Rightarrow \Box W \rangle \models_D \neg \Box W$, but $\langle \diamond \neg W, R \Rightarrow \Box W, R \rangle \not\models_D \neg \Box W$. Thus, while you can infer in the main categorical context of your deliberation that it is not the case that the streets must be wet, you cannot infer this in the hypothetical context triggered by the supposition that it is raining.

However, unlike Willer, I am suggesting here that a subproof restriction is violated in the *reductio* proof only because of how your supposition affects the informational background of your theoretical deliberation, not because of the non-monotonicity of logical consequence. Suppose that sitting in your office, you reflect that the streets might not be wet, that if it is raining the streets must be wet, *and* that it is raining. What can you infer on this basis? According to Willer, you cannot infer that it is not the case that the streets must be wet. However, it seems to me that you can infer *anything you please*. The situation is like one where you begin reasoning from the contradictory basis that it is raining and that it is not raining.

⁴³In Bledin [2013], Appendix A, I distinguish this *lossless* kind of supposition from the earlier *lossy* kind. Lossy supposition triggers hypothetical contexts in which the premises of an argument can fail to hold. Lossless supposition always triggers hypothetical contexts in which one’s information incorporates everything that was incorporated before.

ponens in a hypothetical context, you *can* conclude that it is not raining. Suppose instead that the blinds in your office are raised high enough to see the streets but not the sky. Given your evidence, you reflect that the streets are not wet, but that if it is raining the streets must be wet. In fact, your evidence rules out that it is raining, but you are slow to realize this so you suppose that it is raining. Applying *modus ponens* in this hypothetical context, you infer that the streets must be wet, and therefore that they are wet. Recognizing that this conflicts with the first premise, you conclude by *reductio* that it is not raining.

1	The streets are not wet	Premise
2	If it is raining, the streets must be wet	Premise
3	It is raining	Supposition
4	The streets must be wet	From 2,3
5	The streets are wet	From 4
6	The streets are not wet	From 1
7	⊥	From 5,6
8	It is not raining	From 3,4-7

The only significant difference between this episode of reasoning and the previous one is that the first premise is now non-modal.⁴⁴ But your argumentation, though protracted, is impeccable. This strongly suggests that the problem with the previous example was not the use of *modus ponens*. To be fair, Kolodny and MacFarlane claim only that *modus ponens* will sometimes lead you astray in hypothetical contexts, not that it always will. However, this does not satisfactorily explain why you can argue to the stronger conclusion that it is not raining in this second example, but not in the first example.

Here is an explanation that leaves *modus ponens* untouched. If you were to gain new evidence in this second example that ruled out some possibilities left open by your initial evidence, then you could still reflect that the streets are not wet, but that if it is raining the streets must be wet. Unlike the modal premise in the first example, both premises in this second example continue to hold when the informational background of your deliberation contracts—and in particular, when it contracts to any

⁴⁴The second example also involves the inference from ‘The streets must be wet’ to ‘The streets are wet’ but this inference is uncontroversially valid.

single possibility where it is raining. The conflict between these two premises and your supposition that it is raining thus reveals not just that your original information does not rule out that it is not raining, but also that this information rules out the possibility that it is raining. On the basis of this conflict, you can felicitously conclude that it is not raining.⁴⁵

The real lesson from the above examples is not that *modus ponens* for the indicative is unreliable in hypothetical deliberative contexts, but instead that we must exercise caution when using *reductio ad absurdum* in languages with informational constants. Let us call a sentence $\varphi \in S_{\mathcal{L}}$ *persistent* (cf. Veltman [1996]) when there is no model \mathcal{M} such that for some $i, i' \in 2^{\mathcal{W}}$ where $i' \subset i$, $i \triangleright \varphi$ and $i' \not\triangleright \varphi$. The sentences $\neg W$, $R \Rightarrow \Box W$, and R are all persistent, but $\Diamond \neg W$ is non-persistent since when $\mathcal{V}(W, w_1) = T$ and $\mathcal{V}(W, w_2) = F$, $\{w_1, w_2\} \triangleright \Diamond \neg W$ and $\{w_1\} \not\triangleright \Diamond \neg W$. We should then distinguish between these different good forms of indirect proof:⁴⁶

1-*reductio*: If \perp follows from A and the premise set Γ , then $\Diamond \neg A$ follows from Γ .

2-*reductio*: If \perp follows from A and only *persistent* members of Γ , then $\neg A$ follows from Γ .

2-*reductio* is applied in the second example. But the *reductio* in the first example fits neither of these good forms.⁴⁷

7 Constructive Dilemma

However, my defense of *modus ponens* is incomplete. One might still worry that this argument form can lead a reasoner astray in a different environment: the hypothetical contexts of constructive dilemma.⁴⁸

⁴⁵Some formalism can help sharpen this point. Let i explicate your information. It rules out the possibility that the streets are wet, and rules out the possibility that it is raining and the streets are not wet, so $i \triangleright \neg W$ and $i \triangleright R \Rightarrow \Box W$. Consider some arbitrary $w \in i$. Since $\{w\} \triangleright \neg W$ and $\{w\} \triangleright R \Rightarrow \Box W$, $\mathcal{V}(R, w) = T$ only if $\{w\} \triangleright \perp$. Hence $\mathcal{V}(R, w) = F$, and since w was arbitrary, $i \triangleright \neg R$.

⁴⁶There are other good forms of indirect proof besides. Here, I present only the simplest ones where you suppose atomic A .

⁴⁷Kolodny and MacFarlane also present “a more powerful variant” of McGee’s famous ‘counterexample’ to *modus ponens* mentioned at the start of §3, but it involves the infelicitous use of *reductio* with a non-persistent premise of the form $\neg(\varphi \Rightarrow \psi)$. Again, *modus ponens* is not to blame.

⁴⁸The classical constructive dilemma is the following argument form: if χ follows from φ and Γ , and ξ follows from ψ and Γ , then $\chi \vee \xi$ follows from $\varphi \vee \psi$ and Γ .

Suppose you know that either John or Niko is in his office. You also know that if John is in his office then it must be Monday, and that if Niko is in his office then it must be Friday, but you do not know which day it is. After reflecting that either John or Niko is in his office, you suppose that John is in his office. Applying *modus ponens* in this hypothetical context, you infer that it must be Monday. You then suppose that Niko is in his office and infer that it must be Friday. By constructive dilemma, you conclude that either it must be Monday or it must be Friday. But given your evidence, it might not be Monday and it might not be Friday. Again, something has gone very wrong.⁴⁹

1	John is in or Niko is in	Premise
2	If John is in, it must be Monday	Premise
3	If Niko is in, it must be Friday	Premise
4	John is in	Supposition
5	It must be Monday	From 2,4
6	Niko is in	Supposition
7	It must be Friday	From 3,6
8	It must be Monday or it must be Friday	From 1,4-5,6-7

It might be tempting here to think that the use of *modus ponens* at steps 5 and 7 causes the trouble. But again, I think this misdiagnoses the problem, which arises only at the final step. And again, I think this misdiagnosis stems from a failure to appreciate the informational background of your deliberation. In the hypothetical context triggered by your supposition that John is in his office, your information leaves open the possibility only that John is in his office on a Monday, so is therefore information according to which either it must be Monday or it must be Friday. Likewise, in the hypothetical context triggered by your supposition that Niko is in his office, your information leaves open

⁴⁹On the basis of your evidence, you *can* conclude that it must be either Monday or Friday. But note the difference between this necessary disjunction and the disjunction of necessities in the main text.

A similar example might lead one to worry that the argument from φ to $\Box\varphi$ —another form of argument that is validated by \models_I but not by \models_{Tr} —is unreliable in hypothetical contexts. Using this argument form and constructive dilemma, you can reason from the premise that either John or Niko is in his office to the conclusion that either John must be in his office or Niko must be in his office. However, I do not think that the argument from φ to $\Box\varphi$ is the problem. My diagnosis of this example and the one in the main text is the same.

the possibility only that Niko is in his office on a Friday, so is therefore information according to which either it must be Monday or it must be Friday. But it does not follow from this that your original information is information according to which it must be Monday or it must be Friday. Information that leaves open the possibility only that either John or Niko is in his office needn't be information that either leaves open the possibility only that John is in his office or leaves open the possibility only that Niko is in his office.

Suppose instead that you know that either John must be in his office or Niko must be in his office. You also know that if John is in his office then it is Monday, and that if Niko is in his office then it is Friday. In this case, you *can* felicitously conclude that either it must be Monday or it must be Friday:

1	John must be in or Niko must be in	Premise		
2	If John is in, it is Monday	Premise		
3	If Niko is in, it is Friday	Premise		
4	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-bottom: 1px solid black; padding-right: 5px;">John is in</td> <td style="padding-left: 5px;">Supposition</td> </tr> </table>	John is in	Supposition	
John is in	Supposition			
5	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-bottom: 1px solid black; padding-right: 5px;">It is Monday</td> <td style="padding-left: 5px;">From 2,4</td> </tr> </table>	It is Monday	From 2,4	
It is Monday	From 2,4			
6	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-bottom: 1px solid black; padding-right: 5px;">Niko is in</td> <td style="padding-left: 5px;">Supposition</td> </tr> </table>	Niko is in	Supposition	
Niko is in	Supposition			
7	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border-bottom: 1px solid black; padding-right: 5px;">It is Friday</td> <td style="padding-left: 5px;">From 3,6</td> </tr> </table>	It is Friday	From 3,6	
It is Friday	From 3,6			
8	It must be Monday or it must be Friday	From 1,4-5,6-7		

If your information in the first hypothetical context has a structural feature—ruling out that it is not Monday—and your information in the second hypothetical context has another structural feature—ruling out that it is not Friday—then your original information has at least one of these two features.

Importantly, there are also cases where you can felicitously use the premise that either John or Niko is in his office as input for constructive dilemma:

1	John is in or Niko is in	Premise
2	If John is in, it's Monday	Premise
3	If Niko is in, it is Friday	Premise
4	John is in	Supposition
5	It is Monday	From 2,4
6	Niko is in	Supposition
7	It is Friday	From 3,6
8	It is Monday or Friday	From 1,4-5,6-7

This non-modal variation of the first example is fine. Why? In this case, you conclude, not that either it must be Monday or it must be Friday, but instead that either it is Monday or Friday—and this weaker conclusion *is* established by your hypothetical reasoning. From the fact that your information in the first hypothetical context leaves open the possibility only that it is Monday, and your information in the second hypothetical context leaves open the possibility only that it is Friday, it does not follow that your original information has one or the other of these structural features. It does follow that every possibility left open by your original information is a possibility where either it is Monday or Friday.⁵⁰

The lesson from these examples, like before, is not that *modus ponens* for the indicative is unreliable, but instead that we should distinguish between these two good forms of constructive dilemma:⁵¹

1-constructive dilemma: If C follows from A and Γ , and D follows from B and Γ , then $\Box C \vee \Box D$ follows from $\Box A \vee \Box B$ and Γ .

2-constructive dilemma: If C follows from A and only *persistent* members of Γ , and D follows from B and only *persistent* members of Γ , then $C \vee D$ follows from $A \vee B$ and Γ .⁵²

⁵⁰Let i explicate your information. It rules out the possibility that neither John nor Niko is in his office, rules out the possibility that John is in his office and it is not Monday, and rules out the possibility that Niko is in his office and it is not Friday, so $i \triangleright J \vee N$, $i \triangleright J \Rightarrow M$, and $i \triangleright N \Rightarrow F$. Consider some arbitrary $w \in i$. Since $i \triangleright J \vee N$, $\mathcal{V}(J, w) = T$ or $\mathcal{V}(N, w) = T$. Since $i \triangleright J \Rightarrow M$, $\mathcal{V}(J, w) = T$ only if $\mathcal{V}(M, w) = T$. Since $i \triangleright N \Rightarrow F$, $\mathcal{V}(N, w) = T$ only if $\mathcal{V}(F, w) = T$. Hence either $\mathcal{V}(M, w) = T$ or $\mathcal{V}(F, w) = T$, and since w was arbitrary, $i \triangleright M \vee F$.

⁵¹I present only the simplest forms of constructive dilemma where you infer C from supposition A and infer D from supposition B .

⁵²Without the restriction to persistent members, 2-constructive dilemma would

1-constructive dilemma and 2-constructive dilemma are applied in the second and third example respectively. But the constructive dilemma in the first example fits neither of these good forms.

8 Conclusion

I have advertised throughout this essay that logical validity, understood along informational lines, extensionally coincides with good deductive argument. But my case so far has been piecemeal. In §5, I claimed that a handful of good deductive arguments involving informational modals and the indicative conditional are incorporation preserving. In §6 and §7, I defended the reliability of one of these arguments—*modus ponens* for the indicative—in hypothetical contexts. In these later sections, though, I also discussed the *informational background* of deliberation, and so gestured at a general explanation for this extensional equivalence.

Deductive argumentation, on the informal, pre-theoretic picture I have had in mind, is an information-driven enterprise in which an agent investigates what is so according to a salient body of information that incorporates the premises of an argument. In many contexts, this body of information is the informational content of the agent’s beliefs—the agent is trying to determine how things are in the actual world. But it needn’t be. An agent might be investigating what is so according to the clues in the famous zebra puzzle, or a politician’s stump speech, or the testimony of an untrustworthy eyewitness to the murder, and so on. If this testimony incorporates that Colonel Mustard did it and that if Colonel Mustard did it then he used the candlestick, what else does this information incorporate?

One answer, of course, is that Colonel Mustard used the candlestick. Another is that either Colonel Mustard used the candlestick or Reverend Green used the revolver. In general, the inferences that the agent can appropriately make on the basis of this testimony are precisely those that preserve incorporation. My punch line is already, I hope, obvious: on the natural, intuitive picture of deductive argumentation just sketched, the *good* deductive arguments are precisely those that count as valid on the informational view. The fact that $\{\neg A\} \models_I \neg \Diamond A$, $\{A \vee B, \neg B\} \models_I \Box A$, and $\{A \Rightarrow (\neg B \Rightarrow C), A\} \models_I \neg B \Rightarrow C$ is not a happy coincidence. A clear picture of what is going on in deductive argumentation reveals that the informational concept of logical validity coincides with good

lead to bad results. From the premise that either John or Niko is in his office, and the premise that John might not be in his office, you could conclude that Niko is in his office.

deductive argument.

I have also stressed how this informational picture of inquiry leads to a better understanding of hypothetical reasoning. Suppose that you want to establish that Colonel Mustard did it. On the standard truth-centric picture where inquirers seek to determine what is true based on an initial body of truths, you might assume that Colonel Mustard did not do it and show that this together with your premises leads to contradiction. By establishing that your premises and assumption cannot be jointly true, the old story goes, you have established that Colonel Mustard did it on the basis of your premises. However, when we inquire in a language with informational constants, this picture is too simple. As discussed in §6, if one of the premises used in the indirect proof is non-persistent, then you can conclude only that Colonel Mustard might have done it. To establish the stronger conclusion that Colonel Mustard did it by *reductio ad absurdum*, you must not appeal to any persistent sentences inside the hypothetical context.

More generally, the informational picture explains the greater variety of good argument forms involving hypothetical reasoning in a language with informational constants.⁵³ It is hard to see how we could make sense of this complexity without moving to this deeper picture where inquirers seek to determine what is so according to information that meets the structural constraints imposed by the premises of an argument. Of course, there is still plenty of work left to do. For one thing, logicians have developed elegant proof systems appropriate to the standard truth preservation view that accurately codify good deductive argumentation

⁵³In addition to *modus ponens* for the indicative conditional, we can felicitously use this method of conditional proof:

conditional proof: If $\Box B$ follows from A and only *persistent* members of Γ , then $A \Rightarrow \Box B$ follows from Γ .

When you suppose that Professor Plum did it, you enter a hypothetical context where your information rules out that he did not do it. If you then establish in this context that the murder must have occurred in the library—using only persistent premises—you also establish that according to your original information if Professor Plum did it then the murder must have occurred in the library.

Compare the material conditional. Since $\{\varphi \supset \psi, \varphi\} \models_{Tr} \psi$, $\{\varphi \supset \psi, \varphi\} \models_I \psi$. But we do not want the above form of conditional proof for the material conditional on the informational view. When you suppose that Professor Plum did it, he must have done it according to your information in the hypothetical context triggered by your supposition. But your original information might leave open the possibilities both that Professor Plum did it and that he did not do it, so needn't be information according to which either Professor Plum did not do it or he must have done it. Someone who understands logic along informational lines might draw this moral: if we want both *modus ponens* and conditional proof for our conditional over languages with \Box , then we are after \Rightarrow not \supset .

in simple languages. It would be nice to have a formal proof system that models, *inter alia*, the different forms of indirect proof and constructive dilemma discussed in §6 and §7. But I think that progress can be made. For a large fragment of the language \mathcal{L} from §3, I present such a system in the appendix.

A The Natural Deduction System Info

In this appendix, I present a Fitch-style natural deduction proof system, **Info**, appropriate to the informational view of logic. In A.1, I define the languages relevant to this system. In A.2, I catalogue its rules. In A.3, I prove soundness with respect to informational consequence \models_I . In A.4, I prove completeness.

A.1 Languages

There will be a few languages in play. The most rudimentary is the language of sentential logic \mathcal{L}_0 :

Syntax of \mathcal{L}_0 . The symbols of \mathcal{L}_0 are $A, B, C, \dots, \perp, \neg, \vee, \wedge$, and parentheses.⁵⁴ The atomic sentences in $S_{\mathcal{L}_0}$ are A, B, C, \dots , and \perp . If $P, Q \in S_{\mathcal{L}_0}$, then $\neg P, (P \vee Q), (P \wedge Q) \in S_{\mathcal{L}_0}$. Nothing else is in $S_{\mathcal{L}_0}$.

Semantics of \mathcal{L}_0 . See §3.

The next language \mathcal{L}_1 is a restricted fragment of the language from §3 sans complex embeddings of \Box, \Diamond , and \Rightarrow , and certain hybrid sentences that combine subsentences involving informational constants with those not involving these constants:⁵⁵

Syntax of \mathcal{L}_1 . The symbols of \mathcal{L}_1 are those of \mathcal{L}_0 plus \Box, \Diamond , and \Rightarrow . If $P, Q \in S_{\mathcal{L}_0}$, then $P, \Box P, \Diamond P, (P \Rightarrow Q) \in S_{\mathcal{L}_1}$. If $\varphi, \psi \in S_{\mathcal{L}_1} \setminus S_{\mathcal{L}_0}$, then $\neg\varphi, (\varphi \vee \psi), (\varphi \wedge \psi) \in S_{\mathcal{L}_1}$. Nothing else is in $S_{\mathcal{L}_1}$.

Semantics of \mathcal{L}_1 . See §3.

The final language \mathcal{L}_2 is the *language of incorporation*:

Syntax of \mathcal{L}_2 . The symbols of \mathcal{L}_2 are those of \mathcal{L}_0 plus $\mathbf{i}, +$, and the relation \triangleright . The atomic sentences in $S_{\mathcal{L}_2}$ are all of the form $\mathbf{i} + P \triangleright Q$

⁵⁴ \supset and \equiv can be defined in the usual fashion.

⁵⁵For example, $\Diamond\Diamond A, A \Rightarrow \Box B, A \Rightarrow (B \Rightarrow C)$, and $\Diamond A \vee B$ are not in $S_{\mathcal{L}_1}$. Working with the simpler language \mathcal{L}_1 allows for a simpler proof system, and makes it easier to prove completeness in A.4. However, simplicity comes at the cost of limited applicability. I plan to extend **Info** in future research.

where $P, Q \in S_{\mathcal{L}_0}$ (I omit the $+P$ when P is logically equivalent to $\neg\perp$). If $\varphi, \psi \in S_{\mathcal{L}_2}$, then $\neg\varphi, (\varphi \vee \psi), (\varphi \wedge \psi) \in S_{\mathcal{L}_2}$. Nothing else is in $S_{\mathcal{L}_2}$.

Semantics of \mathcal{L}_2 . $\llbracket i + P \triangleright Q \rrbracket^{w,i} = T$ iff $\forall v \in i + P(\llbracket Q \rrbracket^{v,i+P} = T)$ iff $i + P \triangleright Q$. For the rest, see §3.

Note that $\llbracket i + P \triangleright Q \rrbracket^{w,i} = \llbracket P \Rightarrow Q \rrbracket^{w,i}$, so \mathcal{L}_2 is not really needed in addition to \mathcal{L}_1 . But the language of incorporation will be put to good use. With it, the informational background of argumentation in \mathcal{L}_1 will be made explicit.

Consider $|\cdot| : S_{\mathcal{L}_1} \mapsto S_{\mathcal{L}_2}$ where $P, Q \in S_{\mathcal{L}_0}$ and $\varphi, \psi \in S_{\mathcal{L}_1} \setminus S_{\mathcal{L}_0}$:

$$\begin{aligned} |P| &= i \triangleright P \\ |\Box P| &= i \triangleright P \\ |\Diamond P| &= \neg i \triangleright \neg P \\ |P \Rightarrow Q| &= i + P \triangleright Q \\ |\neg\varphi| &= \neg|\varphi| \\ |\varphi \vee \psi| &= |\varphi| \vee |\psi| \\ |\varphi \wedge \psi| &= |\varphi| \wedge |\psi| \end{aligned}$$

A proof by induction on the complexity of sentences in $S_{\mathcal{L}_1}$ establishes that $i \triangleright \varphi$ if and only if $\forall w \in i(\llbracket |\varphi| \rrbracket^{w,i} = T)$.⁵⁶ The argumentation $|\Pi|$ in \mathcal{L}_2 obtained by substituting $|\varphi|$ for φ at each line of argumentation Π in \mathcal{L}_1 then keeps track of incorporation facts as Π proceeds. For example:⁵⁷

1	$A \Rightarrow B$	Premise	1	$i + A \triangleright B$
2	$\Box\neg B \wedge \Box\neg C$	Premise	2	$i \triangleright \neg B \wedge i \triangleright \neg C$
3	\overline{A}	Supposition	3	$\overline{i \triangleright A}$
4	B	From 1,3	4	$\overline{i \triangleright B}$
5	$\Box\neg B$	From 2	5	$i \triangleright \neg B$
6	\perp	From 4,5	6	$i \triangleright \perp$
7	$\neg A$	From 3-6	7	$i \triangleright \neg A$
8	$\neg\Diamond A$	From 7	8	$\neg\neg i \triangleright \neg A$

⁵⁶The proof is left to the punctilious reader. Note that when $i \neq \emptyset$, $i \triangleright \varphi$ iff $\forall w(\llbracket |\varphi| \rrbracket^{w,i} = T)$. However, this equivalence fails in the degenerate case where $i = \emptyset$: $\emptyset \triangleright \Diamond A$, but $\llbracket \Diamond A \rrbracket^{w,\emptyset} = \llbracket \neg i \triangleright \neg A \rrbracket^{w,\emptyset} = F$ for any w , since $\emptyset \triangleright \neg A$.

⁵⁷We should really distinguish between lossy and lossless supposition in our proof systems (see n. 43). But for ease of exposition, I stick to the standard Fitch-style proof structure here.

The goodness of Π will turn on whether $|\Pi|$ is or can be expanded into a proof in the system **Info**. If $|\Pi|$ or its expansion is a proof in **Info**, then we might confer proofhood on Π as an inherited honorific status.

A.2 Rules of Info

These proof rules fix the *inner logic* of \triangleright , where $P, Q, R, S \in S_{\mathcal{L}_0}$ and $|\varphi| \vdash_s^* |\psi|$ designates that $|\psi|$ has been proven in subproof s initiated with $|\varphi|$ using only certain kinds of sentences:

(\star) Convert $|\chi| \in S_{\mathcal{L}_2}$ into *disjunctive normal form*: a sentence of the form $\chi_1 \vee \dots \vee \chi_n$ where each χ_i is a conjunction of literals—atomic sentences $\mathbf{i} + P \triangleright Q$ or their negations $\neg \mathbf{i} + P \triangleright Q$. If any literal is of the form $\neg \mathbf{i} + P \triangleright Q$ and $|\chi|$ is appealed to inside of s , then $|\varphi| \vdash_s^* |\psi|$.

$$\begin{array}{l}
\neg\mathbf{Intro}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright Q \vdash_s^* \mathbf{i} + P \triangleright \perp}{\mathbf{i} + P \triangleright \neg Q} \\
\neg\mathbf{Elim}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright \neg\neg Q \quad \neg \mathbf{i} + P \triangleright \neg\neg Q}{\mathbf{i} + P \triangleright Q \quad \neg \mathbf{i} + P \triangleright Q} \\
\vee\mathbf{Intro}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright Q}{\mathbf{i} + P \triangleright Q \vee R} \\
\vee\mathbf{Elim}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright Q \vee R \quad \mathbf{i} + P \triangleright Q \vdash_{s_1}^* \mathbf{i} + P \triangleright S \quad \mathbf{i} + P \triangleright R \vdash_{s_2}^* \mathbf{i} + P \triangleright S}{\mathbf{i} + P \triangleright S} \\
\wedge\mathbf{Intro}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright Q \quad \mathbf{i} + P \triangleright R}{\mathbf{i} + P \triangleright Q \wedge R} \\
\wedge\mathbf{Elim}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright Q \wedge R \quad \mathbf{i} + P \triangleright Q \wedge R}{\mathbf{i} + P \triangleright Q \quad \mathbf{i} + P \triangleright R} \\
\perp\mathbf{Intro}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright Q \quad \mathbf{i} + P \triangleright \neg Q}{\mathbf{i} + P \triangleright \perp} \\
\perp\mathbf{Elim}_{\triangleright} \quad \frac{\mathbf{i} + P \triangleright \perp}{\mathbf{i} + P \triangleright Q} \\
\diamond\mathbf{Intro}^{58} \quad \frac{\mathbf{i} + P \triangleright Q}{\neg \mathbf{i} + P \triangleright \neg Q} \\
\diamond\mathbf{Elim} \quad \frac{\neg \mathbf{i} \triangleright \neg P \quad \mathbf{i} + P \triangleright Q}{\neg \mathbf{i} \triangleright \neg Q} \\
\Rightarrow\mathbf{Intro} \quad \frac{\mathbf{i} \triangleright P \vdash_s^* \mathbf{i} \triangleright Q}{\mathbf{i} + P \triangleright Q} \\
\Rightarrow\mathbf{Elim} \quad \frac{\mathbf{i} \triangleright P \quad \mathbf{i} + P \triangleright Q}{\mathbf{i} \triangleright Q}
\end{array}$$

⁵⁸Though \diamond and \Rightarrow are not in \mathcal{L}_2 , I label the remaining four rules as introduction and elimination rules for these symbols given the corresponding transitions in \mathcal{L}_1

These proof rules fix the *outer logic* of \triangleright , where $|\varphi|, |\psi|, |\chi| \in S_{\mathcal{L}_2}$ and $|\varphi| \vdash_s |\psi|$ designates that $|\psi|$ has been proven in subproof s initiated with $|\varphi|$ using any sentences:

$$\begin{array}{l}
\mathbf{Reit} \quad \frac{|\varphi|}{|\varphi|} \\
\neg\mathbf{Intro} \quad \frac{|\varphi| \vdash_s \mathbf{i} \triangleright \perp}{\neg|\varphi|} \\
\neg\mathbf{Elim} \quad \frac{\neg\neg|\varphi|}{|\varphi|} \\
\forall\mathbf{Intro} \quad \frac{\frac{|\varphi|}{|\varphi| \vee |\psi|} \quad \frac{|\varphi|}{|\psi| \vee |\varphi|}}{|\varphi| \vee |\psi|} \\
\forall\mathbf{Elim} \quad \frac{|\varphi| \vee |\psi| \quad |\varphi| \vdash_{s_1} |\chi| \quad |\psi| \vdash_{s_2} |\chi|}{|\chi|} \\
\wedge\mathbf{Intro} \quad \frac{\frac{|\varphi|}{|\varphi| \wedge |\psi|} \quad \frac{|\psi|}{|\varphi| \wedge |\psi|}}{|\varphi| \wedge |\psi|} \\
\wedge\mathbf{Elim} \quad \frac{|\varphi| \wedge |\psi|}{|\varphi|} \quad \frac{|\varphi| \wedge |\psi|}{|\psi|} \\
\perp\mathbf{Intro} \quad \frac{|\varphi| \quad \neg|\varphi|}{\mathbf{i} \triangleright \perp} \\
\perp\mathbf{Elim} \quad \frac{\mathbf{i} \triangleright \perp}{|\varphi|}
\end{array}$$

Note that if the inner logic rules $\neg\mathbf{Intro}_{\triangleright}$, $\forall\mathbf{Elim}_{\triangleright}$, and $\Rightarrow\mathbf{Intro}$ were formulated with \vdash_s instead of \vdash_s^* , then \mathbf{Info} would deliver some horrible results. Consider something like the first example in §6:

1	$\diamond\neg W$	1	$\neg\mathbf{i} \triangleright \neg\neg W$	Premise
2	$R \Rightarrow W$	2	$\mathbf{i} + R \triangleright W$	Premise
3	\overline{R}	3	$\overline{\mathbf{i} \triangleright R}$	Supposition
4	\overline{W}	4	$\overline{\mathbf{i} \triangleright W}$	$\Rightarrow\mathbf{Elim}$: 2,3
5	$\neg\Box W$	5	$\neg\mathbf{i} \triangleright W$	$\neg\mathbf{Elim}_{\triangleright}$: 1
6	\perp	6	$\mathbf{i} \triangleright \perp$	$\perp\mathbf{Intro}$: 4,5
7	$\neg R$	7	$\mathbf{i} \triangleright \neg R$	$\neg\mathbf{Intro}_{\triangleright}$: 3-6

that these rules effectively license. Since $|\Box P| = |P|$, $\Box\mathbf{Intro}$ and $\Box\mathbf{Elim}$ would be reiteration rules.

This *reductio* is infelicitous. However, its \mathcal{L}_2 -counterpart would be a proof in **Info** given the \vdash_s -form of $\neg\mathbf{Intro}_\triangleright$ together with some of the system's other rules. I argued in §6 that step 7 is problematic given that the non-persistent $\diamond\neg W$ is imported into the subproof at step 5. Non-persistence is a semantic property but the syntactic condition (\star) also does the job.

In fact, (\star) is stronger than required. First, the appeal to some non-persistent sentences in some subproofs is harmless—for example, the appeal to $\diamond A$ inside a subproof beginning with supposition A . Second, $|\varphi| \vdash_s^* |\psi|$ when a sentence like $|\diamond W \vee \neg\diamond W|$ is appealed to inside of s though $\diamond W \vee \neg\diamond W$ is a logical truth so is clearly persistent. However, I prove in A.4 that this excess vigilance does not undermine **Info**'s completeness.

A.3 Soundness

I prove the following soundness theorem, where $\varphi_1, \dots, \varphi_n, \psi \in S_{\mathcal{L}_1}$ and $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\mathbf{Info}} |\psi|$ designates that $|\psi|$ is provable in **Info** from premises $|\varphi_1|, \dots, |\varphi_n|$:

Thm 1. If $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\mathbf{Info}} |\psi|$, then $\{\varphi_1, \dots, \varphi_n\} \models_I \psi$.

Proof: Assume $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\mathbf{Info}} |\psi|$ and $i \triangleright \varphi_1, \dots, i \triangleright \varphi_n$. To show that $i \triangleright \psi$, I first show that **Info**'s simpler inferential rules licensing transitions from input sentences $|\varphi_1^I|, \dots, |\varphi_n^I|$ to output sentence $|\varphi^O|$ preserve incorporation in \mathcal{L}_1 . That is, $i \triangleright \varphi^O$ if $i \triangleright \varphi_1^I \wedge \dots \wedge \varphi_n^I$. Keep in mind that $i \triangleright \varphi$ if and only if $\forall w \in i (\llbracket |\varphi| \rrbracket^{w,i} = T)$, so it suffices to show that $\llbracket |\varphi^O| \rrbracket^{w,i} = T$ if $\llbracket |\varphi_1^I| \rrbracket^{w,i} = \dots = \llbracket |\varphi_n^I| \rrbracket^{w,i} = T$.

Here are a few cases:

$\neg\mathbf{Elim}_\triangleright$: $\llbracket |i + P \triangleright \neg\neg Q| \rrbracket^{w,i} = T$ iff $\forall v \in i + P (\llbracket |\neg\neg Q| \rrbracket^{v,i+P} = T)$ iff $\forall v \in i + P (\llbracket |Q| \rrbracket^{v,i+P} = T)$ iff $\llbracket |i + P \triangleright Q| \rrbracket^{w,i} = T$.

$\vee\mathbf{Intro}_\triangleright$: $\llbracket |i + P \triangleright Q| \rrbracket^{w,i} = T$ iff $\forall v \in i + P (\llbracket |Q| \rrbracket^{v,i+P} = T)$ only if $\forall v \in i + P (\llbracket |Q \vee R| \rrbracket^{v,i+P} = T)$ iff $\llbracket |i + P \triangleright Q \vee R| \rrbracket^{w,i} = T$.

$\perp\mathbf{Intro}_\triangleright$: $\llbracket |i + P \triangleright Q| \rrbracket^{w,i} = \llbracket |i + P \triangleright \neg Q| \rrbracket^{w,i} = T$ iff $i + P = \emptyset$ iff $i + P \triangleright \perp$ iff $\llbracket |i + P \triangleright \perp| \rrbracket^{w,i} = T$.

$\diamond\mathbf{Elim}$: $\llbracket |\neg i \triangleright \neg P| \rrbracket^{w,i} = T$ iff $\exists v \in i (\llbracket |P| \rrbracket^{v,i} = T)$ iff $i + P \neq \emptyset$. So $\llbracket |\neg i \triangleright \neg P| \rrbracket^{w,i} = \llbracket |i + P \triangleright Q| \rrbracket^{w,i} = T$ iff $\forall v \in i + P (\llbracket |Q| \rrbracket^{v,i+P} = T)$ only if (since $i + P \neq \emptyset$) $\exists v \in i (\llbracket |\neg Q| \rrbracket^{v,i} = F)$ iff $\llbracket |\neg i \triangleright \neg Q| \rrbracket^{w,i} = T$.

$\wedge\mathbf{Intro}$: $\llbracket |\varphi| \rrbracket^{w,i} = \llbracket |\psi| \rrbracket^{w,i} = T$ iff $\llbracket |\varphi \wedge \psi| \rrbracket^{w,i} = T$.

$\perp\mathbf{Elim}$: $i \triangleright \varphi_1^I$ iff $\forall w \in i (\llbracket |i \triangleright \perp| \rrbracket^{w,i} = T)$ iff $i = \emptyset$ only if $i \triangleright \varphi^O$.

The remaining cases are similar.

Next consider **Info**'s complex inferential rules involving subproofs. Since these rules license transitions from facts of the form $|\varphi^{s_i}| \vdash_{s_i}^* |\psi^{s_i}|$ or $|\varphi^{s_i}| \vdash_{s_i} |\psi^{s_i}|$ and (in some cases) input sentences $|\varphi_1^I|, \dots, |\varphi_n^I|$ to output sentence $|\varphi^O|$, it must be shown that $i \triangleright \varphi^O$ if $\Gamma \cup \{\varphi^{s_i}\} \models_I \psi^{s_i}$ for each s_i (where Γ is a set of sentences in $S_{\mathcal{L}_1}$ incorporated by i) and $i \triangleright \varphi_1^I \wedge \dots \wedge \varphi_n^I$. For \neg **Intro** $_{\triangleright}$, \vee **Elim** $_{\triangleright}$, and \Rightarrow **Intro**, Γ is a set of persistent sentences. Since $\emptyset \triangleright \varphi^O$, it thus suffices to show that $\llbracket |\varphi^O| \rrbracket^{w,i} = T$ when both $\llbracket |\varphi^{s_i}| \rrbracket^{w,i'} = T$ only if $\llbracket |\psi^{s_i}| \rrbracket^{w,i'} = T$ for each s_i and non-empty $i' \subseteq i$, and $\llbracket |\varphi_1^I| \rrbracket^{w,i} = \dots = \llbracket |\varphi_n^I| \rrbracket^{w,i} = T$.

\neg **Intro** $_{\triangleright}$: $i + P = \emptyset$ only if $i + P \triangleright \neg Q$ iff $\llbracket i + P \triangleright \neg Q \rrbracket^{w,i} = T$, so assume $i + P \neq \emptyset$. Consider arbitrary $w \in i + P$ and let $i = \{w\}$. $\llbracket Q \rrbracket^{w,i+P} = T$ iff $i \triangleright Q$ iff $i + P \triangleright Q$ iff $\llbracket i + P \triangleright Q \rrbracket^{w,i} = T$ only if $\llbracket i + P \triangleright \perp \rrbracket^{w,i} = T$ iff $i \triangleright \perp$. Hence $\llbracket Q \rrbracket^{w,i+P} = F$, and since w was arbitrary, $\forall v \in i + P (\llbracket \neg Q \rrbracket^{v,i+P} = T)$, so $\llbracket i + P \triangleright \neg Q \rrbracket^{w,i} = T$.

\vee **Elim** $_{\triangleright}$: $i + P = \emptyset$ only if $\llbracket i + P \triangleright S \rrbracket^{w,i} = T$, so assume $i + P \neq \emptyset$. Consider $w \in i + P$ and let $i = \{w\}$. $\llbracket i + P \triangleright Q \vee R \rrbracket^{w,i} = T$ iff $i + P \triangleright Q \vee R$ only if $(i + P \triangleright Q$ or $i + P \triangleright R)$ iff $(\llbracket i + P \triangleright Q \rrbracket^{w,i} = T$ or $\llbracket i + P \triangleright R \rrbracket^{w,i} = T)$ only if $\llbracket i + P \triangleright S \rrbracket^{w,i} = T$ iff $i \triangleright S$ iff $\llbracket S \rrbracket^{w,i+P} = T$. Since w was arbitrary, $\forall v \in i + P (\llbracket S \rrbracket^{v,i+P} = T)$, so $\llbracket i + P \triangleright S \rrbracket^{w,i} = T$.

\Rightarrow **Intro**: $i + P = \emptyset$ only if $\llbracket i + P \triangleright Q \rrbracket^{w,i} = T$, so assume $i + P \neq \emptyset$. $\llbracket i \triangleright P \rrbracket^{w,i+P} = T$, so $\llbracket i \triangleright Q \rrbracket^{w,i+P} = T$, and so $\llbracket i + P \triangleright Q \rrbracket^{w,i} = T$.

For \neg **Intro** and \vee **Elim**, Γ can include non-persistent sentences. It then suffices to show that $\llbracket |\varphi^O| \rrbracket^{w,i} = T$ when both $\llbracket |\varphi^{s_i}| \rrbracket^{w,i} = T$ only if $\llbracket |\psi^{s_i}| \rrbracket^{w,i} = T$ for each s_i , and $\llbracket |\varphi_1^I| \rrbracket^{w,i} = \dots = \llbracket |\varphi_n^I| \rrbracket^{w,i} = T$.

\neg **Intro**: $\llbracket |\varphi| \rrbracket^{w,i} = T$ only if $\llbracket i \triangleright \perp \rrbracket^{w,i} = T$ iff $i \triangleright \perp$ iff $i = \emptyset$, contradicting the working assumption that $i \neq \emptyset$. Hence $\llbracket \neg |\varphi| \rrbracket^{w,i} = T$.

\vee **Elim**: $\llbracket |\varphi| \vee |\psi| \rrbracket^{w,i} = T$ iff either $\llbracket |\varphi| \rrbracket^{w,i} = T$ or $\llbracket |\psi| \rrbracket^{w,i} = T$. $\llbracket |\varphi| \rrbracket^{w,i} = T$ only if $\llbracket |\chi| \rrbracket^{w,i} = T$. $\llbracket |\psi| \rrbracket^{w,i} = T$ only if $\llbracket |\chi| \rrbracket^{w,i} = T$. Hence $\llbracket |\chi| \rrbracket^{w,i} = T$. \square

A.4 Completeness

I prove the following completeness theorem:

Thm 2. If $\{\varphi_1, \dots, \varphi_n\} \models_I \psi$, then $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\mathbf{Info}} |\psi|$.

Proof: Assume $\{|\varphi_1|, \dots, |\varphi_n|\} \not\vdash_{\mathbf{Info}} |\psi|$. I show $\{\varphi_1, \dots, \varphi_n\} \not\models_I \psi$. Where **Sent** is a standard proof system for sentential logic (such as the Fitch-style proof system \mathcal{F}_T in Barwise and Etchemendy [1999] minus

the rules for \supset and \equiv), I follow the strategy of van der Does, Groeneveld, and Veltman [1997] and leverage the completeness of **Sent** to prove the general theorem.⁵⁹

First consider the simple case where $\varphi_1, \dots, \varphi_n, \psi \in S_{\mathcal{L}_0}$. Given **Info**'s inner logic rules, $\{|\varphi_1|, \dots, |\varphi_n|\} \not\vdash_{\text{Info}} |\psi|$ only if $\{\varphi_1, \dots, \varphi_n\} \not\vdash_{\text{Sent}} \psi$ iff $\{w\} \triangleright \varphi_1 \wedge \dots \wedge \varphi_n \wedge \neg\psi$ for some w in logical space \mathbb{W} .

Next consider the case where $\varphi_1, \dots, \varphi_n, \psi$ are either sentences in $S_{\mathcal{L}_0}$ or are of the form $\Box P$, $\Diamond P$, or $P \Rightarrow Q$, where $P, Q \in S_{\mathcal{L}_0}$. We can define the following two functions:

$$\begin{array}{ll} P^\bullet & = P & P^{\bullet\bullet} & = P \\ \Box P^\bullet & = P & \Box P^{\bullet\bullet} & = P \\ \Diamond P^\bullet & = \neg\perp & \Diamond P^{\bullet\bullet} & = P \\ P \Rightarrow Q^\bullet & = \neg P \vee Q & P \Rightarrow Q^{\bullet\bullet} & = \neg P \vee Q \end{array}$$

$\{|\varphi|\} \vdash_{\text{Info}} |\varphi^\bullet|$ and $\{|\varphi^{\bullet\bullet}|\} \vdash_{\text{Info}} |\varphi|$. Hence $\{|\varphi_1|, \dots, |\varphi_n|\} \not\vdash_{\text{Info}} |\psi|$ only if $\{|\varphi_1^\bullet|, \dots, |\varphi_n^\bullet|\} \not\vdash_{\text{Info}} |\psi^{\bullet\bullet}|$ only if $\{\varphi_1^\bullet, \dots, \varphi_n^\bullet\} \not\vdash_{\text{Sent}} \psi^{\bullet\bullet}$ iff $\llbracket \varphi_1^\bullet \rrbracket^{w, \{w\}} = \dots = \llbracket \varphi_n^\bullet \rrbracket^{w, \{w\}} = T$ and $\llbracket \psi^{\bullet\bullet} \rrbracket^{w, \{w\}} = F$ for some w in logical space \mathbb{W} . Let $i_1 = \{w \in \mathbb{W} : \llbracket \varphi_1^\bullet \rrbracket^{w, \{w\}} = \dots = \llbracket \varphi_n^\bullet \rrbracket^{w, \{w\}} = T\}$ and $i_2 = i_1 \cap \{w \in \mathbb{W} : \llbracket \psi^{\bullet\bullet} \rrbracket^{w, \{w\}} = F\}$. Note that $w \in i_2 \subseteq i_1$.

i_1 and i_2 are the information states we are after. Suppose ψ is *not* of the form $\Diamond P$:

—If φ_i is of the form P or $\Box P$, $\forall w \in i_1 (\llbracket P \rrbracket^{w, i_1} = T)$, so $i_1 \triangleright \varphi_i$.

—If φ_i is of the form $P \Rightarrow Q$, $\forall w \in i_1 (\llbracket \neg P \vee Q \rrbracket^{w, i_1} = T)$, so $i_1 \triangleright \varphi_i$.

—If φ_i is of the form $\Diamond P$ and $i_1 \not\triangleright \varphi_i$, then $\forall w \in i_1 (\llbracket P \rrbracket^{w, i_1} = F)$, so $\{\varphi_1^\bullet, \dots, \varphi_n^\bullet\} \vdash_{\text{Sent}} \neg P$ by completeness. But $\{\varphi_1^\bullet, \dots, \varphi_n^\bullet\} \vdash_{\text{Sent}} \neg P$ only if $\{|\varphi_1^\bullet|, \dots, |\varphi_n^\bullet|\} \vdash_{\text{Info}} |\neg P|$ only if $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\text{Info}} |\neg P|$ only if $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\text{Info}} |\psi|$ (note $\{|\Diamond P|, |\neg P|\} \vdash_{\text{Info}} \mathbf{i} \triangleright \perp$). Since this contradicts the assumption that $\{|\varphi_1|, \dots, |\varphi_n|\} \not\vdash_{\text{Info}} |\psi|$, $i_1 \triangleright \varphi_i$.

—If ψ is of the form P or $\Box P$, $\llbracket P \rrbracket^{w, i_1} = F$, so $i_1 \not\triangleright \psi$.

—If ψ is of the form $P \Rightarrow Q$, $\llbracket P \rrbracket^{w, i_1} = T$ and $\llbracket Q \rrbracket^{w, i_1} = F$, so $i_1 \not\triangleright \psi$.

Thus, $i_1 \triangleright \varphi_1 \wedge \dots \wedge \varphi_n$ and $i_1 \not\triangleright \psi$.

Next suppose ψ is of the form $\Diamond P$:

—If φ_i is of the form P or $\Box P$, $\forall w \in i_2 (\llbracket P \rrbracket^{w, i_2} = T)$, so $i_2 \triangleright \varphi_i$.

—If φ_i is of the form $P \Rightarrow Q$, $\forall w \in i_2 (\llbracket \neg P \vee Q \rrbracket^{w, i_2} = T)$, so $i_2 \triangleright \varphi_i$.

⁵⁹van der Does, Groeneveld, and Veltman work with only the weak language with the following syntax:

Syntax of $\mathcal{L}_{0.5}$. The symbols of $\mathcal{L}_{0.5}$ are $A, B, C, \dots, \neg, \wedge, \Diamond$, and parentheses. If P is a sentence in $S_{\mathcal{L}_0}$ not involving \vee and \perp , then $P, \Diamond P \in S_{\mathcal{L}_{0.5}}$. Nothing else is in $S_{\mathcal{L}_{0.5}}$.

—If φ_i is of the form $\diamond P$ and $i_2 \not\triangleright \varphi_i$, then $\forall w \in i_2(\llbracket P \rrbracket^{w,i_2} = F)$, so $\{\varphi_1^\bullet, \dots, \varphi_n^\bullet, \neg\psi^{\bullet\bullet}\} \vdash_{\text{Sent}} \neg P$. But $\{\varphi_1^\bullet, \dots, \varphi_n^\bullet, \neg\psi^{\bullet\bullet}\} \vdash_{\text{Sent}} \neg P$ only if $\{|\varphi_1^\bullet|, \dots, |\varphi_n^\bullet|, |P|\} \vdash_{\text{Info}} |\psi^{\bullet\bullet}|$ only if $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\text{Info}} |P \Rightarrow \psi^{\bullet\bullet}|$ only if $\{|\varphi_1|, \dots, |\varphi_n|\} \vdash_{\text{Info}} |\diamond\psi^{\bullet\bullet}|$ given $\diamond\mathbf{Elim}$. Since this contradicts the assumption that $\{|\varphi_1|, \dots, |\varphi_n|\} \not\vdash_{\text{Info}} |\psi|$, $i_2 \triangleright \varphi_i$.

—Since ψ is of the form $\diamond P$, $\forall w \in i_2(\llbracket P \rrbracket^{w,i_2} = F)$, so $i_2 \not\triangleright \psi$.

Thus, $i_2 \triangleright \varphi_1 \wedge \dots \wedge \varphi_n$ and $i_2 \not\triangleright \psi$.

Finally, consider the general case where $\varphi_1, \dots, \varphi_n, \psi \in S_{\mathcal{L}_1}$. Note the following facts, where $\Gamma \subseteq S_{\mathcal{L}_1}$ and $|\Gamma| \subseteq S_{\mathcal{L}_2}$ is obtained from Γ by applying $|\cdot|$ to each of its members:

—If $\chi_1 \notin S_{\mathcal{L}_0}$, then $|\Gamma| \not\vdash_{\text{Info}} |\chi_1|$ only if $|\Gamma| \cup \{|\neg\chi_1|\} \not\vdash_{\text{Info}} |\perp|$. Also, $i \triangleright \neg\chi_1$ iff $i \not\triangleright \chi_1$ when $i \neq \emptyset$, since $\llbracket \neg\chi_1 \rrbracket^{w,i} = T$ iff $\llbracket \neg|\chi_1| \rrbracket^{w,i} = T$ iff $\llbracket |\chi_1| \rrbracket^{w,i} = F$. Hence $\Gamma \cup \{\neg\chi_1\} \not\vdash_I \perp$ only if $\Gamma \not\vdash_I \chi_1$.

—If χ_1 is not of the form P , $\Box P$, $\diamond P$, or $P \Rightarrow Q$, $|\Gamma| \cup \{|\chi_1|\} \not\vdash_{\text{Info}} |\chi_2|$ only if there is a sentence χ'_1 which is a disjunction of conjunctions of sentences or negated sentences of the form $\Box P$, $\diamond P$, or $P \Rightarrow Q$, and $|\Gamma| \cup \{|\chi'_1|\} \not\vdash_{\text{Info}} |\chi_2|$. Also, $i \triangleright \chi_1$ iff $i \triangleright \chi'_1$, since $\llbracket |\chi_1| \rrbracket^{w,i} = \llbracket |\chi'_1| \rrbracket^{w,i}$. Hence $\Gamma \cup \{\chi'_1\} \not\vdash_I \chi_2$ only if $\Gamma \cup \{\chi_1\} \not\vdash_I \chi_2$.

—If $\chi_1 \notin S_{\mathcal{L}_0}$ and $\chi_2 \notin S_{\mathcal{L}_0}$, then $|\Gamma| \cup \{|\chi_1 \vee \chi_2|\} \not\vdash_{\text{Info}} |\chi_3|$ only if either $|\Gamma| \cup \{|\chi_1|\} \not\vdash_{\text{Info}} |\chi_3|$ or $|\Gamma| \cup \{|\chi_2|\} \not\vdash_{\text{Info}} |\chi_3|$. Also, $i \triangleright \chi_1 \vee \chi_2$ iff either $i \triangleright \chi_1$ or $i \triangleright \chi_2$, since $\llbracket |\chi_1 \vee \chi_2| \rrbracket^{w,i} = T$ iff $\llbracket |\chi_1| \vee |\chi_2| \rrbracket^{w,i} = T$ iff either $\llbracket |\chi_1| \rrbracket^{w,i} = T$ or $\llbracket |\chi_2| \rrbracket^{w,i} = T$. Hence $\Gamma \cup \{\chi_1\} \not\vdash_I \chi_3$ only if $\Gamma \cup \{\chi_1 \vee \chi_2\} \not\vdash_I \chi_3$, and $\Gamma \cup \{\chi_2\} \not\vdash_I \chi_3$ only if $\Gamma \cup \{\chi_1 \vee \chi_2\} \not\vdash_I \chi_3$.

—If $\chi_1 \notin S_{\mathcal{L}_0}$ and $\chi_2 \notin S_{\mathcal{L}_0}$, then $|\Gamma| \cup \{|\chi_1 \wedge \chi_2|\} \not\vdash_{\text{Info}} |\chi_3|$ only if $|\Gamma| \cup \{|\chi_1|, |\chi_2|\} \not\vdash_{\text{Info}} |\chi_3|$. Also, $i \triangleright \chi_1 \wedge \chi_2$ iff $i \triangleright \chi_1$ and $i \triangleright \chi_2$, since $\llbracket |\chi_1 \wedge \chi_2| \rrbracket^{w,i} = T$ iff $\llbracket |\chi_1| \wedge |\chi_2| \rrbracket^{w,i} = T$ iff $\llbracket |\chi_1| \rrbracket^{w,i} = \llbracket |\chi_2| \rrbracket^{w,i} = T$. Hence $\Gamma \cup \{\chi_1, \chi_2\} \not\vdash_I \chi_3$ only if $\Gamma \cup \{\chi_1 \wedge \chi_2\} \not\vdash_I \chi_3$.

— $|\Gamma| \cup \{|\neg\Box P|\} \not\vdash_{\text{Info}} |\chi_1|$ only if $|\Gamma| \cup \{|\diamond\neg P|\} \not\vdash_{\text{Info}} |\chi_1|$. Also, $i \triangleright \neg\Box P$ iff $i \triangleright \diamond\neg P$, so $\Gamma \cup \{\diamond\neg P\} \not\vdash_I \chi_1$ only if $\Gamma \cup \{\neg\Box P\} \not\vdash_I \chi_1$.

— $|\Gamma| \cup \{|\neg\diamond P|\} \not\vdash_{\text{Info}} |\chi_1|$ only if $|\Gamma| \cup \{|\Box\neg P|\} \not\vdash_{\text{Info}} |\chi_1|$. Also, $i \triangleright \neg\diamond P$ iff $i \triangleright \Box\neg P$, so $\Gamma \cup \{\Box\neg P\} \not\vdash_I \chi_1$ only if $\Gamma \cup \{\neg\diamond P\} \not\vdash_I \chi_1$.

— $|\Gamma| \cup \{|\neg(P \Rightarrow Q)|\} \not\vdash_{\text{Info}} |\chi_1|$ only if $|\Gamma| \cup \{|\diamond(P \wedge \neg Q)|\} \not\vdash_{\text{Info}} |\chi_1|$. Also, $i \triangleright \neg(P \Rightarrow Q)$ iff $i \triangleright \diamond(P \wedge \neg Q)$, so $\Gamma \cup \{\diamond(P \wedge \neg Q)\} \not\vdash_I \chi_1$ only if $\Gamma \cup \{\neg(P \Rightarrow Q)\} \not\vdash_I \chi_1$.

Therefore, there are sentences χ_i^j of the form P , $\Box P$, $\diamond P$, or $P \Rightarrow Q$ such that $\{|\varphi_1|, \dots, |\varphi_n|\} \not\vdash_{\text{Info}} |\psi|$ only if $\forall_j(\{|\chi_1^j|, \dots, |\chi_{n_j}^j|\} \not\vdash_{\text{Info}} |\chi_{n_j+1}^j|)$, and $\{\chi_1^j, \dots, \chi_{n_j}^j\} \not\vdash_I \chi_{n_j+1}^j$ only if $\{\varphi_1, \dots, \varphi_n\} \not\vdash_I \psi$ for all j . From the

completeness result proven above, it follows that $\{\varphi_1, \dots, \varphi_n\} \not\models_I \psi$. \square

References

- Ernest W. Adams. *The Logic of Conditionals*. Reidel, Dordrecht, 1975.
- Alan Ross Anderson, Nuel Belnap Jr., and J. Michael Dunn. *Entailment, Volume II*. Princeton University Press, Princeton, 1992.
- Jon Barwise. *The Situation in Logic*. CSLI Publications, Stanford, 1989.
- Jon Barwise and John Etchemendy. Information, infons, and inference. In Robin Cooper, Kuniaki Mukai, and John Perry, editors, *Situation Theory and Its Applications, Volume 1*, pages 33–78. CSLI Publications, 1990.
- Jon Barwise and John Etchemendy. *Language, Proof and Logic*. CSLI Publications, Stanford, 1999.
- Justin Bledin. *Logic Informed*. PhD thesis, University of California, Berkeley, 2013.
- Lewis Carroll. A logical paradox. *Mind*, 3(11):436–438, 1894.
- Keith Devlin. *Logic and Information*. Cambridge University Press, Cambridge, 1991.
- Dorothy Edgington. Do conditionals have truth conditions? *Critica: Revista Hispanoamericana de Filosofía*, 18(52):3–39, 1986.
- Hartry Field. Truth and the unprovability of consistency. *Mind*, 115(459):567–606, 2006.
- Hartry Field. *Saving Truth From Paradox*. Oxford University Press, Oxford, 2008.
- Hartry Field. What is the normative role of logic? *Proceedings of the Aristotelian Society Supplementary Volume*, LXXXIII:251–268, 2009a.
- Hartry Field. Pluralism in logic. *Review of Symbolic Logic*, 2:342–359, 2009b.
- Hartry Field. What is logical validity? Unpublished manuscript.
- Gottlob Frege. Logik. 1879-1891. Translated as ‘Logic’ in *Posthumous Writings*, pages 1–8, 1979.

- Gottlob Frege. *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*. Pohle, Jena, 1893. Partially translated as *The Basic Laws of Arithmetic: Exposition of the System*, 1964.
- Gottlob Frege. Logik. 1897. Translated as ‘Logic’ in *Posthumous Writings*, pages 126–151, 1979.
- Gottlob Frege. Der Gedanke: Eine logische Untersuchung. *Beiträge zur Philosophie des Deutschen Idealismus*, 1, 1918. Translated as ‘Thoughts’ by Peter Geach and R. H. Stoothoff in *Collected Papers on Mathematics, Logic, and Philosophy*, pages 351–372, 1984.
- Gottlob Frege. *The Basic Laws of Arithmetic: Exposition of the System*. Trans. Montgomery Furth, University of California Press, Berkeley, 1964.
- Gottlob Frege. *Posthumous Writings*. Trans. Peter Long and Roger White, University of Chicago Press, Chicago, 1979.
- Gottlob Frege. *Collected Papers on Mathematics, Logic, and Philosophy*. Ed. Brian McGuinness, Blackwell, Oxford, 1984.
- Anthony Gillies. Iffiness. *Semantics & Pragmatics*, 3(4):1–42, 2010.
- Gilbert Harman. *Change in View: Principles of Reasoning*. MIT Press, Cambridge, 1986.
- David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes From Kaplan*, pages 481–563. Oxford University Press, 1989.
- Justin Khoo. Operators or restrictors? A reply to Gillies. *Semantics & Pragmatics*, 4(4):1–15, 2011.
- Niko Kolodny and John MacFarlane. Ifs and oughts. *Journal of Philosophy*, CVII(3):115–143, 2010.
- Angelika Kratzer. Conditionals. *Chicago Linguistics Society*, 22(2):1–15, 1986.
- David Lewis. Adverbs of quantification. In E. L. Keenan, editor, *Semantics of Natural Language*, pages 3–15. Cambridge University Press, 1975.
- David Lewis. Index, context, and content. In Stig Kanger and Sven Öhman, editors, *Philosophy and Grammar*, pages 79–100. Reidel, 1980.

- David Lewis. Logic for equivocators. *Nou̇s*, 16(3):431–441, 1982.
- John MacFarlane. Frege, Kant, and the logic in Logicism. *Philosophical Review*, 111(1):25–65, 2002.
- John MacFarlane. Logical constants. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition). <http://plato.stanford.edu/archives/fall2009/entries/logical-constants/>.
- John MacFarlane. In what sense (if any) is logic normative for thought? Unpublished manuscript.
- Vann McGee. A counterexample to Modus Ponens. *Journal of Philosophy*, 82(9):462–471, 1985.
- Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975.
- Robert Stalnaker. *Inquiry*. MIT Press, Cambridge, 1984.
- Alfred Tarski. O pojciu wynikania logicznego. *Przegląd Filozoficzny*, 39, 1936. Translated as ‘On the Concept of Logical Consequence’ in *Logic, Semantics, Metamathematics*, pp. 409–420, 1956.
- Alfred Tarski. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. Ed. and Trans. J. H. Woodger, Clarendon Press, Oxford, 1956.
- Jan Łukasiewicz. Philosophische bemerkungen zu mehrwertigen systemen des aussagenkalküls. *Comptes rendus des seances de la Societe des Sciences et des Lettre de Varsovie*, 23, 1930. Translated as ‘Philosophical Remarks on Many-Valued Systems of Propositional Logic’ in *Selected Works*, pp. 153–178, 1970.
- Jan Łukasiewicz. *Selected Works*. North Holland, Amsterdam, 1970.
- Johan van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, Cambridge, 2011.
- Jaap van der Does, Willem Groeneveld, and Frank Veltman. An update on ‘might’. *Journal of Logic, Language, and Information*, 6:361–380, 1997.
- Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.

Malte Willer. A remark on iff-oughts. *Journal of Philosophy*, 109(7): 449–461, 2012.

Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, 2007.

Seth Yalcin. Nonfactualism about epistemic modality. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 295–332. Oxford University Press, 2011.

Seth Yalcin. A counterexample to modus tollens. *Journal of Philosophical Logic*, 41(6):1001–1024, 2012.