Indicative conditionals and graded information*

Ivano Ciardelli

Preprint — Final version in the *Journal of Philosophical Logic*

Abstract

I propose an account of indicative conditionals that combines features of minimal change semantics and information semantics. As in information semantics, conditionals are interpreted relative to an information state in accordance with the Ramsey test idea: "if p then q" is supported at a state s iff q is supported at the hypothetical state $s[p]$ obtained by restricting s to the p-worlds. However, information states are not modeled as simple sets of worlds, but by means of a Lewisian system of spheres. Worlds in the inner sphere are considered possible; worlds outside of it are ruled out, but to different degrees. In this way, even when a state supports "not p", it is still possible to suppose p consistently. I argue that this account does better than its predecessors with respect to a set of desiderata concerning inferences with conditionals. In particular, it captures three important facts: (i) that a conditional is logically independent from its antecedent; (ii) that a sequence of antecedents behaves like a single conjunctive antecedent (the import-export equivalence); and (iii) that conditionals restrict the quantification domain of epistemic modals. I also discuss two ways to construe the role of a premise, and propose a generalized notion of entailment that keeps the two apart.

Keywords: Indicative conditionals · Epistemic modals · Information semantics · Non-monotonic reasoning · Import-export · Modus ponens

We take ourselves to know the following:

(1) Shakespeare wrote Hamlet.

What if we are wrong? As humans, we are perfectly able to entertain this assumption, and to form a coherent set of conditional beliefs about this hypothetical scenario. We can then use indicative conditionals to report and discuss such beliefs. We would, for instance, accept the statement (2) and reject (3).[1]

*For inspiring discussions and/or useful feedback on earlier drafts, I am grateful to Maria Aloni, Justin Bledin, Bruno Jacinto, Hannes Leitgeb, Vít Punčochář, Floris Roelofsen, Lorenzo Rossi, Katrin Schulz, Frank Veltman, Malte Willer, Seth Yalcin, and two anonymous reviewers.

[1] It is important to notice that, even when we accept (1), the indicative conditionals (2) and (3) do not mean the same as the counterfactual conditionals (i-a) and (i-b):

(i) a. If Shakespeare had not written Hamlet, someone else would have.

(2) If Shakespeare did not write Hamlet, someone else did.

(3) If Shakespeare did not write Hamlet, T. S. Eliot did.

This observation is problematic for the logic textbook, material analysis of conditionals. According to this analysis, an indicative conditional is true if the antecedent is false or the consequent true. This account predicts that (1) logically entails (2) and (3). Since we accept (1), and since (3) follows from it, according to this account we could not rationally reject (3). This seems wrong: it seems perfectly rational to accept (1) and reject (3) (cf. Edgington, 1986).

This problem no longer arises if we adopt the more sophisticated theory of indicative conditionals proposed by Stalnaker (1968, 1976). This view, which I will refer to as *minimal change semantics*, holds that a conditional "if A then C" expresses a proposition—construed as a set of possible worlds—but not the proposition expressed by the material conditional. Rather, "if A then C" is true if, among the worlds in which the antecedent is true, those that are most similar to the actual world are worlds where the consequent is true.[2]

Minimal change semantics does a very good job at predicting certain logical features that distinguish indicative conditionals from material conditionals. In particular, the problem mentioned above no longer arises. The falsity of an antecedent implies nothing about which antecedent worlds are most similar to the actual world; thus, "if A then C" is predicted to be logically independent of "not A". This explains why accepting "not A" is perfectly consistent both with accepting "if A then C", and with rejecting it.

Over the years, however, several important problems with this theory also became apparent. One such problem is that minimal change semantics does not deal well with iterated conditionals like (4).

(4) If Bob is in Paris, then if he's staying in a hotel, he's at the Ritz.

Intuitively, the two *if*-clauses can be collected into a single conjunctive *if*-clause, as in (5), without affecting the meaning of the conditional.

(5) If Bob is in Paris and he's staying in a hotel, he's at the Ritz.

The principle stating the equivalence of sentences like (4) and (5), known as the *import-export* principle, is invalid in minimal change semantics. Worse, the semantics cannot be strengthened to validate this principle without falling back

---

b. If Shakespeare had not written Hamlet, T. S. Eliot would have.

In the case of (i-a) and (i-b), we are holding fixed our belief that Shakespeare wrote Hamlet, and considering an alternative course of history which we take to be non-actual. By contrast, in the case of (2) and (3) we are supposing that our belief that Shakespeare wrote Hamlet is in fact wrong, and considering what things like in that case. In the first case, we assume that the world were different than it is; in the second case, we assume that it is different than we think (see Adams, 1970). In this paper I will mostly set aside counterfactuals, although I will come back briefly to the difference between the two types of conditionals in Footnote 24.

[2] Stalnaker also assumes that for any world w and assumption φ there is a unique φ-world which is maximally similar to w. Whether this assumption should be made is a contentious point, but one which is orthogonal to our concerns in this paper.

into the material analysis (Gibbard, 1980; McGee, 1985; Gillies, 2009).

Another important empirical point that has become salient in the literature since the work of Kratzer (1986) is that if-clauses can restrict the range of modal operators which appear in the consequent clause. Thus, for instance, on its most salient interpretation, (6) does not convey that if Bob is not in London, then all epistemically possible worlds are worlds where he is in Paris; rather, it conveys that all epistemically possible worlds where Bob is not in London are worlds where he is in Paris—i.e., that Bob's being in Paris is epistemically necessary *conditionally on his not being in London*.

(6) If Bob is not in London, he must be in Paris.

Minimal change semantics is not well-equipped to deal with this interaction, since in this account there is no direct relation between the interpretation of an if-clause and the domain of quantification of a modal.

Inspired by these observations and by earlier work on dynamic semantics (especially by Dekker, 1993; Veltman, 1996) in recent years several authors (Gillies, 2004, 2009, 2010; Yalcin, 2007; Kolodny and MacFarlane, 2010; Bledin, 2014; Starr, 2014; Willer, 2014) have taken a different perspective: they have proposed that indicative conditionals do not express properties of possible worlds (propositions) but rather properties of information states.

These proposals are based on the so-called *Ramsey Test* view of conditionals (after a comment in Ramsey, 1929): in order to assess whether an information state supports a conditional, we augment the state hypothetically with the antecedent, and then check whether the resulting state supports the consequent. If, and only if, this is the case, the conditional is supported in the original information state. Formally, an information state s is modeled as a set of possible worlds—those worlds that are compatible with the available information; to assess a conditional "if A then C" with respect to s is to assess C relative to the hypothetical state $s[A]$ which results from restricting s to those worlds where A is true.

This view, which I will refer to as *information semantics*, solves both problems that we discussed for minimal change semantics. First, since a conjunction is true just in case both conjuncts are, we have $s[A][B] = s[A \text{ and } B]$, which implies that "if A, then if B, C" and "if A and B, C" are supported in exactly the same circumstances. Moreover, if the set s of possibilities serves as the domain of quantification for a modal operator O, then the conditional "if A, then $O(B)$" is supported just in case $O(B)$ is supported relative to $s[A]$. Thus, the fact that the *if*-clause restricts the domain of a modal occurring in the consequent is explained straightforwardly. Hence, various pieces of the puzzle fall naturally into place when we take an informational perspective.

However, existing versions of information semantics do not retain some of the most appealing features of minimal change semantics. In particular, they do not make "if A then C" logically independent of "not A", and thus fail to account for our initial observation about (2) and (3). To see the problem, consider first an account governed by the simple Ramsey test clause described above. Suppose an information state s supports "not A"; then updating s with A results in the empty state—the inconsistent state—which supports any conclusion whatsoever; as a consequence, s supports "if A then C". Thus, "not A" is predicted to entail "if A then C", bringing back the problem of the material account. Many implementations of information semantics prevent this by assuming that conditionals cannot be supported vacuously; that is, they make a conditional supported in a state only in case the antecedent can be consistently supposed. This, however, leads to a different problem: now, in any state which supports "not A", "A" cannot be supposed consistently, so "if A then C" will not be supported. Thus "not A" is predicted to be inconsistent with "if A then B". Thus, we could not be rational in accepting (1) and (2).

My aim in this paper is to show that we can combine the attractive features of information semantics with those of minimal change semantics. What I will propose is that the fundamental idea of information semantics is right: indicative conditionals express features of information states, and they do so by claiming that the consequent holds in the hypothetical state that results from updating the current state with the antecedent. However, updating does not always amount to *augmenting* the current information state: sometimes—when the antecedent runs against the available information—updating requires hypothetically giving up some of the available information in order to consistently entertain the assumption.[3] One natural way to model this process is to refine the view of information adopted in standard information semantics. Identifying information states with sets of worlds captures the idea that information allows us to rule out certain states of affairs as candidates for the actual world. But, intuitively, the worlds that we rule out are not all on a par. Our world knowledge rules out a world where T.S. Eliot wrote Hamlet *more strongly* than it rules out a world where another Elizabethan author did. To capture this idea, we need a model of information in which possibility is *graded*, rather than *flat*. For this purpose, we can borrow the formal tools which have been developed in the minimal change semantics tradition (and applied to the analysis of belief revision since Grove (1988)): rather than having a relative similarity ordering $\leq_w$, now we will have a plausibility ordering $\leq_s$, which ranks worlds in terms of how plausible they are according to the available information. The set of maximally plausible worlds in s encodes the unconditional information in s, the set of ways the world might be according to the available information. A sentence A is supported by s if A holds in all the most plausible worlds. The process of making an assumption A can still be modeled by restricting the state to the A-worlds. Now, however, even if a state supports "not A", making the assumption A need not result in the inconsistent information state: even when all maximally plausible worlds make A false, we will, in general, be able to discriminate more and less plausible A-worlds. This ability is crucial in order to

[3] This insight also lies at the heart of approaches to conditionals developed by scholars working on belief revision (see Section 8.2). The account proposed here fits within this tradition. However, one should avoid linking too closely the process of making assumptions and the process of revising beliefs. While there are commonalities between the two, there are also important differences. I will come back to this point in Section 8.2.

assess "if A then C" non-trivially in a context supporting "not A".[4]

We will see that, in addition to combining the desirable features of minimal change semantics and standard information semantics, the resulting graded information semantics also accounts for some observations about conditionals which are not predicted by either of these approaches. In particular:

- the existence of counterexamples to *modus ponens* for nested conditionals of the form "if A, then if B, then C" (McGee, 1985);
- the fact that the *or*-to-*if* inference from "A or B" to "if not A, B" is *almost* valid, though not quite universally valid (Adams, 1965, 1975).

The paper is structured as follows: in Section 1 I describe the formal language and the models that I will use in the paper, and the intended interpretation of the relation of logical entailment. In Section 2 I lay out a number of desiderata, in the form of entailment predictions that a theory of conditionals should deliver. In Sections 3 and 4 I discuss minimal change semantics and information semantics, emphasizing their respective merits and shortcoming with respect to the desiderata. In Section 5 I introduce graded information semantics, showing that it satisfies the desiderata. In Section 6 I discuss some worries about the status of three of the desiderata. In Section 7 I discuss the connection between conditionals and entailment, and propose a generalized notion of entailment which allows us to restore a form of deduction theorem. In Section 8 I compare graded information semantics with three related lines of work on conditionals. In Section 9 I discuss the interpretation of negated conditionals and propose a variant of the system that predicts that the negation of "if A then B" is equivalent to "if A then not B". Section 10 concludes.

1 Preliminaries: language, models, entailment

Throughout the paper I will work with a simple formal language. The base layer of the language is a set $\mathcal{L}_0$ of *factual sentences*, whose semantics can be given in terms of truth-conditions relative to a state of affairs. For our purposes, we can assume that $\mathcal{L}_0$ is the language of propositional logic based on a set of atoms $\mathcal{P}$. Formally, $\mathcal{L}_0$ is given by the following BNF definition, where p ranges over atomic sentences in $\mathcal{P}$:

$$\alpha ::= p \mid \neg\alpha \mid \alpha \wedge \alpha$$

I will use α and β as meta-variables ranging over factual sentences.

The full language $\mathcal{L}$ that I will work with is obtained by extending $\mathcal{L}_0$ with an operator $\Rightarrow$ for indicative conditionals and an operator $\Box$ for the epistemic modal

[4] This is not the first time ranked states appear in the information semantics literature. Gillies (2007) used them to give a dynamic account of counterfactuals, while Willer (2017) used a set of such states for the analysis of indicative conditionals. Both accounts are quite different from the proposal I will make. For instance, in Willer's account making an assumption can trigger certain non-monotonic changes, but nevertheless A is not consistently supposable in a state in which "not A" is accepted.

must (a corresponding operator $\Diamond$ for *might* is defined by duality). I restrict to the case in which the antecedent of a conditional, as well as the argument of a modal, are factual sentences. On the other hand, I allow the consequent of a conditional to be any sentence of our language; so, nested conditionals of the form $p \Rightarrow (q \Rightarrow r)$ are allowed, as well as conditionalized modal statements like $p \Rightarrow \Diamond q$. More formally, the syntax of $\mathcal{L}$ is given by the following BNF definition, where α ranges over sentences in $\mathcal{L}_0$:

$$\varphi ::= \alpha \mid \alpha \Rightarrow \varphi \mid \Box\alpha \mid \neg\varphi \mid \varphi \wedge \varphi$$

I will use φ, ψ, χ as meta-variables ranging over all formulas in $\mathcal{L}$. Disjunction, the falsum constant, and the possibility modal *might* are defined operators:

- $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$
- $\bot := p \wedge \neg p$
- $\Diamond\alpha := \neg\Box\neg\alpha$

Semantically, I will assume a model $M = \langle W, V \rangle$ that provides a universe W of possible worlds—a set whose elements stand for possible states of affairs—together with a valuation function $V : \mathcal{P} \times W \to \{0, 1\}$ which specifies the truth-values of atomic sentence at each possible world. In order to simplify the discussion I will assume that W is finite, although this restriction is not essential. The valuation V extends in the usual way to a valuation $V : \mathcal{L}_0 \times W \to \{0, 1\}$ defined on all factual sentences. If $\alpha \in \mathcal{L}_0$ and $w \in W$, I also write $w(\alpha)$ instead of $V(\alpha, w)$, and denote by $|\alpha|$ the set of worlds where α is true:

$$|\alpha| := \{w \in W \mid w(\alpha) = 1\}$$

The analysis that we are after should deliver a formal relation of entailment $\models \subseteq \wp(\mathcal{L}) \times \mathcal{L}$, which is intended to characterize the pre-theoretical notion of consequence, understood as follows: ψ is a consequence of $\varphi_1, \ldots, \varphi_n$ if and only if ψ is acceptable in any context in which $\varphi_1, \ldots, \varphi_n$ are acceptable. I use the symbol $\equiv$ for logical equivalence, defined as mutual entailment.

2 Desiderata

In this section I lay out some desiderata for a logic of conditionals that is intended to track the pre-theoretical notion of consequence, construed as preservation of acceptability.

2.1 Conditionals and their constituents

The first set of desiderata concerns the interactions between a conditional and its constituent parts. First, a conditional should be compatible with, but should not follow from, the negation of its antecedent.

Desideratum 2.1 (No trivialization)**.**

- $\neg p \not\models p \Rightarrow q$
- $\neg p, p \Rightarrow q \not\models \bot$

The reason for this desideratum is that accepting (7-a) does not commit us either way with respect to (7-b): in a context where we accept (7-a), it seems equally consistent to accept and to reject (7-b) (cf. Edgington, 1986).

(7) a. Alice will not come to the party.
b. If Alice comes to the party, she will come with Bob.

Our second desideratum is that plain indicative conditionals, where antecedent and consequent are factual sentences, obey *modus ponens* and *modus tollens*. The motivation for this desideratum comes from the strong intuitive appeal of instances of these rules, and the apparent lack of counterexamples.

Desideratum 2.2 (Factual modus ponens and modus tollens)**.**
For all $\alpha, \beta \in \mathcal{L}_0$:

- $\alpha \Rightarrow \beta, \alpha \models \beta$
- $\alpha \Rightarrow \beta, \neg\beta \models \neg\alpha$

2.2 Iterated conditionals

The second set of desiderata concerns the behaviour of iterated conditionals, i.e., conditionals whose consequent is itself a conditional. The first desideratum is the so-called *import-export* principle, which says that a sequence of two if-clauses is equivalent to a single, conjunctive if-clause.

Desideratum 2.3 (Import-export)**.**
For all $\alpha, \beta \in \mathcal{L}_0$ and all $\varphi \in \mathcal{L}$:

- $\alpha \Rightarrow (\beta \Rightarrow \varphi) \equiv \alpha \wedge \beta \Rightarrow \varphi$

The argument for this desideratum comes from the observation that (8-a) and (8-b) seem to express the same thing, and that such examples can be multiplied without running into counterexamples (cf. McGee, 1985; Gillies, 2009).

(8) a. If Bob is in Paris, then if he's staying in a hotel, he's at the Ritz.
b. If Bob is in Paris and he's staying in a hotel, he's at the Ritz.

The second desideratum stems from a famous example by McGee (1985). The example is set within the context of the 1980 American presidential campaign. Three candidates are running for President: Reagan, a Republican, is leading in the polls; Carter, a democrat, is lagging behind; Anderson, a Republican running as an independent, has virtually no chances of winning. Now consider:

(9) a. If a Republican wins, then if Reagan doesn't win, Anderson will.

b. A Republican will win.
c. If Reagan does not win, Anderson will.

In the given situation, it seems fully rational to accept (9-a) (since Reagan and Anderson are the only Republican candidates) and (9-b) (since the polls indicate that Reagan will win), while rejecting (9-c) (since Carter is well ahead of Anderson). Thus, our logic should not make (9-c) a consequence of (9-a) and (9-b). This leads to the following desideratum: *modus ponens* should not be a valid principle when applied to iterated conditionals.[5,6]

Desideratum 2.4 (No modus ponens for iterated conditionals)**.**

- $p \Rightarrow (q \Rightarrow r), p \not\models q \Rightarrow r$

2.3 Conditionals and modals

The last set of desiderata has to do with the interaction between conditionals and epistemic modals. As we mentioned in the introduction, a conditional antecedent can restrict the scope of a modal occurring in the consequent. Thus, e.g., in the configuration $p \Rightarrow \Box q$, the necessity modal does not quantify over all the epistemically possible worlds, but only over the epistemically possible p-worlds. As a consequence, the consequent of the conditional does not make the same modal claim as the un-embedded modal sentence $\Box q$.

This fact has important logical repercussions. To see why, take a variation on an example used by Veltman (1985) and Yalcin (2012). A marble is extracted at random from an urn filled with black and white marbles. The color of the marble has not yet been revealed. Intuitively, (10-a) and (10-b) are acceptable in the described situation, but we have no reason to accept (10-c).

(10) a. If the marble is not white, it must be black.
b. It is not the case that the marble must be black.
c. The marble is white.

This leads to the following desideratum: *modus tollens* should not be generally valid for conditionals whose consequent is of the form $\Box q$.[7]

Desideratum 2.5 (No modus tollens for modal consequents)**.**

- $p \Rightarrow \Box q, \neg\Box q \not\models \neg p$

[5]Some proponents of information semantics, notably Bledin (2015), have argued that this example does not provide a counterexample to *modus ponens*. I will discuss their objection and defend Desideratum 2.4 in Section 6.2.

[6]Analogous violations of *modus ponens* for iterated conditionals have been pointed out by Briggs (2012) in the case of counterfactuals. Briggs' example has nothing to do with epistemic uncertainty, yet from a structural point of view it closely parallels McGee's example.

[7]In the literature, one can find examples of the failure of *modus tollens* with various kinds of consequents. Veltman (1985), Cantwell (2008), and Yalcin (2012) discuss respectively the case in which the consequent is a conditional, a deontic modal, and a probability operator.

The next desideratum concerns the so-called *or*-to-*if* inference—the inference from a disjunction $p \vee q$ to a conditional $\neg p \Rightarrow q$. In most contexts, an inference from (11-a) to (11-b) seems intuitively impeccable.

(11) a. Either the butler or the gardener did it.
b. If the butler didn't do it, the gardener did.

Yet, as Adams (1965, 1975) and Edgington (1986) pointed out, this inference can be unwarranted in a situation in which our acceptance of the disjunction is based solely on the acceptance of the first disjunct. For instance, suppose the investigation started out with three suspects: the butler, the gardener, and the chauffeur. Convincing evidence has been found that indicates the butler as the culprit, while nothing has emerged against the gardener or the chauffeur. Based on this information, the detective accepts (11-a), since she accepts the first disjunct; but she has no reason to accept (11-b). Examples of this kind show that the *or*-to-*if* inference does not *always* preserve acceptability.

Nevertheless, in those cases where our acceptance of the disjunction is *not* based on the acceptance of the first disjunct, the argument from (11-a) to (11-b) seems impeccable—and we would like it to be justified by our logic. Assuming that lack of acceptance of p can be expressed by the modal statement $\Diamond\neg p$, this leads to the following desideratum.

Desideratum 2.6 (Cautious or-to-if)**.**

- $p \vee q \not\models \neg p \Rightarrow q$
- $p \vee q, \Diamond\neg p \models \neg p \Rightarrow q$

The final desideratum concerns the interaction of conditionals and *might*. Consider the following sentences:

(12) a. If Hamlet was not written by Shakespeare, it might have been written by someone else.
b. It might be that Hamlet was not written by Shakespeare, but by someone else.

Someone uttering (12-b) is calling the authorship of Hamlet into question in a way that someone uttering (12-a) is not. Indeed, (12-a) sounds quite obvious—even slightly odd, since something stronger seems assertible: if Hamlet was not written by Shakespeare, then it *must* have been written by someone else. By contrast, before accepting (12-b) we would expect the speaker to provide evidence undermining our belief that Shakespeare wrote Hamlet.

This motivates the following desideratum: our theory of conditionals should tease apart $p \Rightarrow \Diamond q$ and $\Diamond(p \wedge q)$; in particular, it should predict that $\Diamond(p \wedge q)$ entails $\Diamond p$, while $p \Rightarrow \Diamond q$ does not.[8]

Desideratum 2.7 (If-might interaction)**.**

[8]Further evidence of the difference between $p \Rightarrow \Diamond q$ and $\Diamond(p \wedge q)$ will be given in Sec. 6.3.

- $p \Rightarrow \Diamond q \not\models \Diamond p$
- $\Diamond(p \wedge q) \models \Diamond p$

3 Minimal change semantics

3.1 Account

Minimal change semantics (Stalnaker, 1968; Lewis, 1973)[9] is based on the following assumptions: a conditional $\alpha \Rightarrow \varphi$ has truth-conditions with respect to possible worlds; the truth of $\alpha \Rightarrow \varphi$ at a world w depends on whether φ holds at certain α-worlds, which are construed as those α-worlds which are most similar to the world of evaluation in the relevant respects.[10]

In order to articulate this proposal, the model M needs to be augmented with a further parameter, namely, a function which yields for each world $w \in W$ a total pre-ordering $\leq_w$ defined over a subset of W.[11] Intuitively, $v \leq_w u$ holds if v is at least as similar to w as u is in the relevant respects. The relation $\leq_w$ is required to satisfy certain constraints. In particular, any world w is assumed to be strictly more similar to itself than any other world is. This ensures that if α is true at w, then the truth of a conditional $\alpha \Rightarrow \varphi$ depends only on whether φ is true at w. In Stalnaker's version of the theory, but not in Lewis's version, it is furthermore required that $\leq_w$ should be a total ordering—i.e., that there be no ties in relative similarity.

To state the semantics formally, let us denote by $\mathsf{Best}(w, \alpha)$ the set of α-worlds which are most similar to w, i.e., the set of worlds in $|\alpha|$ which are minimal with respect to $\leq_w$. Moreover, let R be an epistemic accessibility relation: intuitively, wRv holds if v is compatible with the information available at w to the relevant agent or group.[12] Then the semantics for our language $\mathcal{L}$ can be stated as follows.

- $w \models p$ iff $w(p) = 1$
- $w \models \neg\varphi$ iff $w \not\models \varphi$
- $w \models \varphi \wedge \psi$ iff $w \models \varphi$ and $w \models \psi$
- $w \models \Box\alpha$ iff $v \models \alpha$ for all $v \in R[w]$

[9] It should be noted that, while Stalnaker viewed his theory as concerning both indicative and counterfactual conditionals, Lewis viewed his proposal only as a theory of counterfactuals. For indicatives, he endorsed the material analysis. Nevertheless, Lewis's variant of minimal change semantics can be viewed as a theory of indicative conditionals just as well as Stalnaker's, which is why I cite both of them as sources for this approach.

[10] Clarifying what the "relevant respects" is a thorny problem, but we need not get into it here, as it is immaterial to the entailment predictions made by the account.

[11] A total pre-ordering is a relation R which which is reflexive, transitive, and such that for every two items x, y in its domain, either xRy or yRx.

[12] Whose information matters for the truth of epistemic modal statements is another thorny question (see, e.g., Price, 1983; von Fintel and Gillies, 2010; Yalcin, 2011) which, fortunately, can be set aside for our purposes.

- $w \models \alpha \Rightarrow \varphi$ iff $v \models \varphi$ for all $v \in \mathsf{Best}(w, \alpha)$

The notion of entailment stemming from minimal change semantics is just the standard truth-conditional one: an entailment is valid if the conclusion is true in any world of any model where the premises are all true.

3.2 Predictions

Let us assess how well minimal-change semantics does in terms of the desiderata laid out in the previous section. First, suppose that in the actual world, $\neg p$ is true. This tells us nothing about whether q is true in the closest p-worlds. So, the prediction is that $p \Rightarrow q$ is consistent with, but does not follow from, the negation $\neg p$. Thus, Desideratum 2.1 is satisfied.

Desideratum 2.2 is satisfied as well. To see that the theory validates *modus ponens*, suppose α and $\alpha \Rightarrow \varphi$ are both true at w. Since α is true, we have $\mathsf{Best}(w, \alpha) = \{w\}$. Since $\alpha \Rightarrow \varphi$ is true, φ must be true at each world in $\mathsf{Best}(w, \alpha)$, which means that it must be true at w. To see that *modus tollens* is also valid, suppose $\alpha \Rightarrow \varphi$ and $\neg\varphi$ are both true at w. If α was also true at w, then by *modus ponens* φ would be true at w, which would contradict the truth of $\neg\varphi$. Therefore, $\neg\alpha$ must be true at w.

Notice, however, that the proof we have just seen shows that *modus ponens* and *modus tollens* are valid generally, for arbitrary consequents. This means that Desiderata 2.4 and 2.5, which require these principles to fail for certain kinds of consequents, are not satisfied. To get a better sense of why they are not, let us take a closer look at the predictions of minimal change semantics about the examples that we used to motivate those desiderata.

Let us first look at McGee's election scenario. Intuitively, the following sentence seems true at the actual world w.

(13) If Reagan does not win, Carter will.

The actual world is also one where a Republican, namely Reagan, wins. So, the set of closest worlds where a Republican wins is the singleton $\{w\}$, and in all the worlds in this set, (13) is true. This means that (14) is true at w:

(14) If a Republican wins, then if Reagan does not win, Carter will.

This is a bad prediction, since (14) sounds not simply false, but plainly absurd, given that Carter is not a Republican.

Notice that, by contrast, the conditional (15) is rightly predicted to be false: all the closest worlds where a Republican wins and Reagan does not win are worlds where a Republican wins, and thus where Carter does not win.

(15) If a Republican wins and Reagan does not win, Carter will.

This shows that (14) and (15) are not equivalent in minimal change semantics: thus, the *import-export* principle fails, and Desideratum 2.3 is not satisfied.

Conceptually, what this example shows is that minimal change semantics does not give a satisfactory account of iterated conditionals of the form $\alpha \Rightarrow (\beta \Rightarrow \varphi)$, since a sequence of two if-clauses does not result in φ being assessed in a context where both α and β hold, as it intuitively should.

Next consider the marbles example. The context specifies that it is unknown whether the extracted marble is black or white. Thus, among the worlds $v \in R[w]$ which are epistemically possible at the actual world w there are some where the extracted marble is white, and others where it is black. This correctly predicts that (16-a) is false. But suppose the actual world is one where the extracted marble happens to be black. Then the antecedent of (16-b) is true. So the truth of (16-b) just turns on whether the consequent, which coincides with (16-a), is true; since (16-a) is false, (16-b) is predicted to be false.

(16) a. The marble must be black.
b. If the marble is not white, it must be black.

However, as we discussed above, (16-b) is intuitively true in the described situation. The reason is that, in assessing (16-b), we take the modal to quantify not over all epistemic possibilities, but only over those where the antecedent is true. In other words, we take the domain of the modal to be restricted by the antecedent. This interaction between conditional antecedents and the domains of modals is not captured, and is not expected, in minimal change semantics.

Due to the lack of interactions between conditionals and modals, Desideratum 2.6 is not satisfied either. As required, minimal change semantics invalidates the plain *or*-to-*if* entailment, $p \vee q \models \neg p \Rightarrow q$: if p is true, then $p \vee q$ is true, but nothing guarantees that the closest $\neg p$ world is a q world. But adding the extra assumption $\Diamond\neg p$ does not change the situation, so the cautious version of the principle, $p \vee q, \Diamond\neg p \models \neg p \Rightarrow q$ is also invalid. Thus, the apparent status of the *or*-to-*if* inference as a sound piece of reasoning is not accounted for.[13]

Finally, consider Desideratum 2.7. In this case, minimal change semantics makes the desired predictions: if a conjunction is epistemically possible at a world, so are the conjuncts; so, $\Diamond(p \wedge q)$ is predicted to entail $\Diamond p$. On the other hand, if p is false then the fact that q is epistemically possible at the most similar p-worlds implies nothing about what is epistemically possible at the actual world. So, $p \Rightarrow \Diamond q$ is not predicted to entail $\Diamond p$.

The predictions of minimal change semantics with respect to our desiderata are summarized in the following table.

Desideratum	2.1	2.2	2.3	2.4	2.5	2.6	2.7
Minimal change semantics	✓	✓	×	×	×	×	✓

[13] See Stalnaker (1976) for a pragmatic account of the *or*-to-*if* inference, and Gillies (2004) for why such a pragmatic explanation is not quite satisfactory if our focus is on acceptance rather than assertion.

4 Flat information semantics

4.1 Account

In recent work, a radically different view on the semantics of indicative conditionals has emerged. In this view, an indicative conditional does not express a property of states of affairs, but rather a property of states of information.[14] This account of conditionals, ultimately rooted in the work of Veltman (1985, 1996) and Dekker (1993), has been defended most prominently by Gillies (2004, 2009, 2010). However, the specific version of this view that I discuss here is due to Yalcin (2007) and Bledin (2014).

In this account, an information state s is identified with a set of possible worlds: a set $s \subseteq W$ models a body of information which is compatible with all worlds $w \in s$ and incompatible with all worlds $w \notin s$. Since an information state is modeled as a simple set of worlds, with no further structure on it, I will refer to this account as *flat* information semantics; this term is meant to contrast with the *graded* information semantics developed in the next section.

The semantics evaluates sentences as true or false relative to two parameters, a possible world w—relevant for factual sentences—and an information state s—relevant for conditionals and epistemic modals. Thus, truth is given as a relation $\langle w, s\rangle \models \varphi$ between world-state pairs and sentences of $\mathcal{L}$. In addition to the primitive notion of truth we also have a derived notion of *support* (also known as *incorporation*) relative to an information state s, which captures which sentences are established in s. This is defined by checking what is guaranteed to be true regardless of which world within s is actual:

Definition 4.1 (Support). $s \models \varphi$ iff $\forall w \in s : \langle w, s\rangle \models \varphi$

Notice that the empty set $\emptyset$—which models a state of inconsistent information—trivially supports all formulas. Also notice that, if the truth-conditions of φ depend only on the world parameter (as will be the case for factual sentences $\alpha \in \mathcal{L}_0$) then support at s amounts to truth at each world in s. On the other hand, if $s \neq \emptyset$ and the truth-conditions of φ depend only on the information component (as will be the case for conditionals and modal sentences), then support at s coincides with truth at $\langle w, s\rangle$ for an arbitrary $w \in W$.

Assessing a conditional involves making an supposition—an operation that modifies the information available in an information state. Technincally, supposing α in a state s amounts to zooming in on the α-worlds in s. This is captured by defining the *update* $s[\alpha]$ of an information state s with a factual sentence $\alpha \in \mathcal{L}_0$ as follows.

[14]The information state which is described by uttering a conditional is often assumed to be the speaker's own information state. However, the semantics is compatible with different interpretations. Observations concerning cases of disagreement about conditionals seem to point at a more complex story: conversational participants use epistemic expressions to align their information states and debate the available evidence, rather than merely describing their respective private beliefs. Although this is an important issue, I will set it aside in this paper, since my focus here is on the logical predictions delivered by such an account.

Definition 4.2 (Update)**.** For $\alpha \in \mathcal{L}_0$, $s[\alpha] := s \cap |\alpha|$

Then the semantics for the language $\mathcal{L}$ is given by the following clauses.

Definition 4.3 (Flat information semantics)**.**

- $\langle w, s\rangle \models p$ iff $w(p) = 1$
- $\langle w, s\rangle \models \neg\varphi$ iff $\langle w, s\rangle \not\models \varphi$
- $\langle w, s\rangle \models \varphi \wedge \psi$ iff $\langle w, s\rangle \models \varphi$ and $\langle w, s\rangle \models \psi$
- $\langle w, s\rangle \models \Box\alpha$ iff $s \models \alpha$
- $\langle w, s\rangle \models \alpha \Rightarrow \varphi$ iff $s[\alpha] \models \varphi$[15]

It is clear from the clauses that factual sentences $\alpha \in \mathcal{L}_0$ depend only on the world component for their truth: for all w, s, s', $\langle w, s\rangle \models \alpha$ iff $\langle w, s'\rangle \models \alpha$. On the other hand, conditionals and epistemic modal sentences φ depend only on the informational component: for all w, w', s, $\langle w, s\rangle \models \varphi$ iff $\langle w', s\rangle \models \varphi$.

The relation of logical entailment is cashed out as preservation of support: an entailment holds if the conclusion is supported in any information state that supports the premises (for discussion, see Yalcin, 2007; Bledin, 2014).

Definition 4.4 (Entailment)**.**

$$\varphi_1, \ldots, \varphi_n \models \psi \quad \text{iff} \quad \begin{array}{l} \text{for any model } M \text{ and information state } s: \\ \text{if } s \models \varphi_i \text{ for } 1 \leq i \leq n, \text{ then } s \models \psi \end{array}$$

Equipped with this new definition of entailment, let us now look at how flat information semantics fares with respect to our desiderata.

4.2 Predictions

Interestingly, the issues that were problematic for minimal change semantics fall naturally into place in this approach. Consider the import-export principle. Since a conjunction is true at those worlds where both conjuncts are true, for any state s we have $s[\alpha \wedge \beta] = s[\alpha][\beta]$. This implies that for any state s:

$$\begin{array}{rll} s \models \alpha \Rightarrow (\beta \Rightarrow \varphi) & \text{iff} & s[\alpha][\beta] \models \varphi \\ & \text{iff} & s[\alpha \wedge \beta] \models \varphi \\ & \text{iff} & s \models \alpha \wedge \beta \Rightarrow \varphi \end{array}$$

Thus, the import-export equivalence $\alpha \Rightarrow (\beta \Rightarrow \varphi) \equiv \alpha \wedge \beta \Rightarrow \varphi$ is logically valid, and Desideratum 2.3 is satisfied.

[15] Yalcin (2007), but not Bledin (2014), also requires $s[\alpha] \neq \emptyset$. I will discuss this modification in Section 4.3.

Moreover, unlike in minimal change semantics, in this approach conditional antecedents restrict the domain of epistemic modals: this is because a conditional antecedent updates the information state parameter, which then provides the domain of quantification for the modal. Indeed, we have:

$$\begin{array}{lll} s \models \alpha \Rightarrow \Box\beta & \text{iff} & s[\alpha] \models \Box\beta \\ & \text{iff} & \forall w \in s \cap |\alpha| : w(\beta) = 1 \\[1ex] s \models \alpha \Rightarrow \Diamond\beta & \text{iff} & s[\alpha] \models \Diamond\beta \\ & \text{iff} & \exists w \in s \cap |\alpha| : w(\beta) = 1 \end{array}$$

This allows us to satisfy Desideratum 2.5. To see this, consider again the urn scenario. Let s be an information state which captures the information available in this scenario: s contains only worlds where the extracted marble is white or black, and it contains worlds of both kinds. Let w and b stand respectively for "the marble is white" and "the marble is black". Then $s[\neg w]$ consists entirely of worlds where the marble is black; so, $s[\neg w]$ supports b, and thus also $\Box b$; this means that $s \models \neg w \Rightarrow \Box b$. Also, since s contains worlds where the marble is not black we have $s \not\models b$, which implies that $s \models \neg\Box b$. On the other hand, since s contains worlds where the marble is white, $s \not\models \neg w$. Thus, the intuitive judgments that we discussed about the sentences in (10) are predicted. This also means that *modus tollens* fails for modal consequents, as required by Desideratum 2.5.

Thus, flat information semantics provides an elegant and insightful account of the data concerning nested conditionals and the conditional-modal interaction, just those data that were puzzling for minimal change semantics. Unfortunately, however, flat information semantics fails to satisfy some of the desiderata that flat information semantics *did* satisfy.

In particular, the entailment $\neg p \models p \Rightarrow q$ is valid. To see this, suppose $s \models \neg p$. This means that s contains no p-worlds, and so $s[p] = \emptyset$ trivially supports q. So, $s \models p \Rightarrow q$. Thus, Desideratum 2.1 is not satisfied.

It not hard to see that *modus ponens* and *modus tollens* come out as valid for factual conditionals; therefore Desideratum 2.2 is satisfied. However, *modus ponens* is in fact valid in full generality. This means that Desideratum 2.4, which requires the failure of *modus ponens* for nested conditionals, is not satisfied. To see what the problem is, let us look at the predictions of the semantics for McGee's example. Suppose state s supports r = "Reagan will win". Then s contains only worlds where Reagan wins. But then updating s with the assumption $\neg r$ that Reagan does not win results in an inconsistent state, which trivially supports any sentence, including the conclusion that Anderson will win (as well as the opposite conclusion that he won't). Thus, s trivially supports "If Reagan does not win, Anderson will".

Next, consider the *or*-to-*if* inference. We saw above that minimal change semantics renders both the plain and the cautious version of this inference invalid. By contrast, flat information semantics makes them both valid. To see why, suppose $s \models p \vee q$. This means that $p \vee q$ is true at all worlds in s. Thus, at every world in s where p is not true, q is true. This implies that $s[\neg p] \models q$, which means that $s \models \neg p \Rightarrow q$. Thus, $p \vee q \models \neg p \Rightarrow q$ is valid. To better see the problem, consider the special case in which our state s supports $p \vee q$ because it supports p, while it contains no information about q. Intuitively, s provides then no support for $\neg p \Rightarrow q$. However, in flat information semantics, in this case $s[\neg p]$ is the empty state, which supports any formula whatsoever. Thus, in this case, s trivially supports $\neg p \Rightarrow \varphi$ for any φ.

Finally, consider the conditional $p \Rightarrow \Diamond q$. If a state $s \neq \emptyset$ contains no p-worlds, then $p \Rightarrow \Diamond q$ is vacuously supported, while $\Diamond p$ is not. This shows that $p \Rightarrow \Diamond q$ does not entail $\Diamond p$; since the latter is obviously entailed by $\Diamond(p \wedge q)$, Desideratum 2.7 is satisfied. The following table summarizes the situation.

Desideratum	2.1	2.2	2.3	2.4	2.5	2.6	2.7
Flat information semantics	×	✓	✓	×	✓	×	✓

4.3 Antecedent compatibility

The problems we diagnosed for flat information semantics stem from the fact that, when assuming α leads to an inconsistent information state, all conditionals of the form $\alpha \Rightarrow \varphi$ are trivially supported. Many versions of information semantics (e.g., Yalcin, 2007; Gillies, 2009; Starr, 2014; Willer, 2014), however, assume that an indicative conditional can never be trivially supported; that is, in order for $\alpha \Rightarrow \varphi$ to be supported at s, $s[\alpha]$ must be non-empty:[16]

- $s \models \alpha \Rightarrow \varphi$ iff $s[\alpha] \neq \emptyset$ and $s[\alpha] \models \varphi$

This move prevents the problematic entailment $\neg p \models p \Rightarrow q$. However, it does not rescue Desideratum 2.1, since it goes too far in the other direction: now, no state can support both $\neg p$ and $p \Rightarrow q$, which makes these two formulas inconsistent, contrary to the second half of Desideratum 2.1. Thus, the resulting logic would predict that one cannot accept both $\neg p$ and $p \Rightarrow q$, on pain of inconsistency. Intuitively, however, it seems that we can; for instance, it seems perfectly consistent to accept both (1) = "Shakespeare wrote Hamlet" and (2) = "If Shakespeare did not write Hamlet, someone else did".

The antecedent compatibility condition helps with Desideratum 2.6, which will now be satisfied: the conclusion $\neg p \Rightarrow q$ is not entailed by $p \vee q$ alone, but it is jointly entailed by $p \vee q$ together with $\Diamond \neg p$, where the latter premise serves to ensure that the compatibility condition is satisfied.

However, the move also creates a new problem: now, any state supporting a conditional of the form $\alpha \Rightarrow \varphi$ must be compatible with α, and therefore must support $\Diamond \alpha$. Thus, the entailment $\alpha \Rightarrow \varphi \models \Diamond \alpha$ becomes valid for all α and φ. In particular, we get $p \Rightarrow \Diamond q \equiv \Diamond(p \wedge q)$, in contrast with Desideratum 2.7.

Finally, the move does not help with Desideratum 2.4. For suppose that s supports $p \Rightarrow (q \Rightarrow r)$. Then $s[p][q]$ must be a non-empty set which supports r.

[16]The underlying idea, which can be traced back to Von Fintel (1998), is that (i) indicative conditionals quantify over antecedent worlds in the context set, and (ii) quantifiers generally presuppose that their domain is non-empty.

If furthermore $s \models p$, then $s[p] = s$, so $s[p][q] = s[q]$. Thus, $s[q]$ is a non-empty state which supports r, which guarantees that $s \models q \Rightarrow r$.

The predictions of flat information semantics augmented with an antecedent compatibility requirement can thus be summarized as follows.

Desideratum	2.1	2.2	2.3	2.4	2.5	2.6	2.7
Flat info semantics + AC	×	✓	✓	×	✓	✓	×

In sum, while adding an antecedent compatibility requirement solves some of the problems pointed out above, it does not solve all of them, and it creates some new ones. The deeper issue is that, as pointed out in the beginning, we are capable of meaningfully assessing indicative conditionals involving antecedents which run against our beliefs, so long as these beliefs are not absolutely certain. What we want is a semantics that accounts for how we interpret such conditionals, rather than rendering them defective in some way.[17]

5 Graded information semantics

As we discussed in the beginning of the paper, we can, and often do, entertain (epistemic, non-counterfactual) assumptions that are contrary to what we believe. For instance, we can meaningfully debate how Hamlet might have come about if Shakespeare did not write it. This means that making an assumption does not always amount to *extending* the current information with the antecedent; sometimes, part of the available information must be hypothetically given up in order to consistently entertain the assumption. In order to model this process, I will borrow the formal tools developed in the minimal change semantics tradition, and later applied to the modeling of belief revision. For reasons that will become clear soon, I will refer to the resulting system as *graded information semantics*, abbreviated as GrIS.

5.1 Account

Our starting point is a more fine-grained modeling of information states. This modeling reflects the fact that information is not an all-or-nothing matter: between worlds that are regarded as absolutely possible and worlds that are regarded as absolutely impossible there is a gradient of intermediate degrees. Following Grove (1988), we can model this by taking the worlds in an information

[17] This is not to reject the plausible idea that the antecedent of an indicative conditional must in some sense be epistemically possible for the conditional to be felicitous. However, the considerations above indicate that the kind of possibility required for the felicity of an indicative conditional $\alpha \Rightarrow \varphi$ is very weak, weaker than the kind of possibility needed to support $\Diamond\alpha$, and weak enough to be compatible with the acceptance of $\neg\alpha$. In the framework presented below, this difference can be formulated precisely as follows: in order for an indicative conditional $\alpha \Rightarrow \varphi$ to be felicitous, α must be true at some world which is within the purview of one's epistemic state, although this world need not be viewed as a serious candidate for the actual world; by contrast, in order for an epistemic modal claim $\Diamond\alpha$ to be accepted, α must be true in some world which is a serious candidate for the actual world.

s to be ranked by a total pre-order $\leq$, encoding *implausibility*. Implausibility should be understood in the following technical sense: worlds which are minimal with respect to $\leq$ represent the ways the world might be according to the available information. Non-minimal worlds are ruled out by the available information, but they are not all ruled out with the same strength: if $w \leq w'$, then w' is ruled out at least as strongly as w.[18] This leads us to the notion of a *graded information state*.

Definition 5.1 (Graded information states)**.**
A graded information state is a pair $s = \langle D_s, \leq_s \rangle$ where $D_s \subseteq W$ and $\leq_s$ is a total pre-ordering over D_s. We write $\mathsf{Best}(s)$ for the set of $\leq_s$-minimal elements in D_s.[19]

Example 5.2. Consider whether Alice and Bob went to the party last night. They told us they would go, and we have no reason to doubt that. But since we were not there, we are not absolutely sure. Since Alice and Bob always hang out together, we think that if they didn't both go, neither of them went. In the unlikely event that only one of them went, we have no reason to think that it was Alice rather than Bob or vice versa. This body of information can be captured as a graded information state, and represented graphically as follows:

$$s = \begin{array}{|c|c|c|}\hline ab & \emptyset & a, b \\ \hline \end{array}$$

Here, the cells correspond to different degrees of implausibility, ordered from the lowest (to the left) to the highest (to the right). In each cell we have the corresponding worlds, separated by a comma. When this is unproblematic, like here, I will display worlds by listing the atomic sentences which are true in them ($\emptyset$ denotes the world in which neither comes).

Notice that graded states are, in a precise sense, a generalization of information states modeled as sets of worlds, since any set of worlds $D \subseteq W$ can be identified with the graded state $\langle D, \mathsf{tot}_D \rangle$ where $\mathsf{tot}_D = D \times D$ is the total relation on D. Graded states of the form $\langle D, \mathsf{tot}_D \rangle$ will be called *flat states*.

The process of making an assumption α in state s can still be captured in terms of restricting s to the worlds where α is true.

Definition 5.3 (Updating a graded state)**.**
If $s = \langle D, \leq \rangle$ is a graded state and $\alpha \in \mathcal{L}_0$, the update of s with α is the state $s[\alpha] = \langle D_\alpha, \leq_\alpha \rangle$ where $D_\alpha = D \cap |\alpha|$ and $\leq_\alpha$ is the restriction of $\leq$ to D_α.

[18] Using pre-orders in this way is standard in the belief revision tradition. It goes back to Grove (1988), who cites Lewis (1973) for inspiration, and it has been used, among others, by Gärdenfors (1988); Segerberg (1998); Board (2004); Baltag and Smets (2006).

[19] Notice that, as long as $D_s \neq \emptyset$, the set $\mathsf{Best}(s)$ is guaranteed to be non-empty since W is finite. In the infinite case, one may either complicate the semantic definition of support in a way analogous to the truth-conditions given in Lewis (1973), or make the limit assumption, i.e., assume that the order $\leq_s$ is well-founded, allowing no infinite strictly descending chains. My preference is for the latter option, since the limit assumption is needed to guarantee that an entertainable supposition has a consistent set of consequences (see Herzberger, 1979).

Notice that making an assumption is an irreversible operation: after we have assumed that α, the $\neg\alpha$-worlds are completely dropped from the state, and no further assumption can bring them back. This is an important feature of the model: it corresponds to the fact that, in a context where α is being assumed, α is treated as being certainly true ("true by assumption"). In this respect, making an assumption is different from actually revising an information state to integrate new information. We will come back to this point in Section 8.2.

Example 5.4. Consider the state s of Example 5.2. Updating s with the assumption $a =$ "Alice didn't go" yields the following state:

$$s[\neg a] = \begin{array}{|c|c|}\hline \emptyset & b \\ \hline\end{array}$$

As in flat information semantics, formulas are assessed with respect to two parameters: a world of evaluation, relevant for factual sentences, and an information state, relevant for conditionals and epistemic modal claims; the only difference is that the information state will be graded rather than flat. As in information semantics, in addition to the basic notion of truth relative a world-state pair we have a defined notion of *support* relative to an information state, which captures what is settled in it. This is defined as what is the case in all the worlds in $\mathsf{Best}(s)$, which represent the possibilities which are left open by the available information:

$$s \models \varphi \quad \text{iff} \quad \forall w \in \mathsf{Best}(s) : \langle w, s \rangle \models \varphi$$

With the revised notions of information state, update, and support in place, the clauses of flat information semantics can be taken over without any changes. I repeat these clauses below for convenience.

Definition 5.5 (Graded information semantics)**.**
Let w be a possible world and s a graded information state. We let:

- $\langle w, s \rangle \models p \;$ iff $\; w(p) = 1$
- $\langle w, s \rangle \models \neg\varphi \;$ iff $\; \langle w, s \rangle \not\models \varphi$
- $\langle w, s \rangle \models \varphi \wedge \psi \;$ iff $\; \langle w, s \rangle \models \varphi$ and $\langle w, s \rangle \models \psi$
- $\langle w, s \rangle \models \Box\alpha \;$ iff $\; s \models \alpha$
- $\langle w, s \rangle \models \alpha \Rightarrow \varphi \;$ iff $\; s[\alpha] \models \varphi$[20]

As in standard information semantics, the truth-conditions of a factual sentence depend only on the world of evaluation, while those of epistemic modal sentences and indicative conditionals depend only on the information state.

[20] Following Yalcin (2007) and most work in information semantics, we may also add an antecedent compatibility condition, requiring that $s[\alpha] \neq \emptyset$. In GrIS, unlike in flat information semantics, such a requirement would not entail that $\neg p$ and $p \Rightarrow q$ are inconsistent. I think natural language does implement such a requirement, but I will leave it out for simplicity.

As in flat information semantics, the inconsistent information state, which now amounts to $s_\perp = \langle\emptyset, \emptyset\rangle$, trivially supports every formula. On the other hand, relative to consistent states s, the support conditions for factual sentences, modal sentences, and conditionals can be written as follows:

Fact 5.6 (Derived support conditions, for a consistent state s)**.**

- $s \models \alpha$ iff $\mathsf{Best}(s) \subseteq |\alpha|$, for $\alpha \in \mathcal{L}_0$
- $s \models \Box\alpha$ iff $\mathsf{Best}(s) \subseteq |\alpha|$
- $s \models \Diamond\alpha$ iff $\mathsf{Best}(s) \cap |\alpha| \neq \emptyset$
- $s \models \alpha \Rightarrow \varphi$ iff $s[\alpha] \models \varphi$

Entailment is still defined as preservation of support, as in Definition 4.4.[21]

Graded information semantics is, in a precise sense, a generalization of flat information semantics: the latter can be retrieved as a special case by restricting our attention to flat states, where the plausibility ordering is total. Starting with a state s modeled as a set of worlds, we can think of it as a flat state $\langle s, \mathsf{tot}_s\rangle$ and check that flat information semantics on s coincides with graded information semantics on $\langle s, \mathsf{tot}_s\rangle$.

What about the other direction, i.e., simulating a graded state by means of a flat state? In a graded information state s, the unconditional information about how the world might be is captured by the set $\mathsf{Best}(s)$. So, it is natural to simulate a graded state s by identifying it with the set of worlds $\mathsf{Best}(s)$. It is easy to check that, for all $\Rightarrow$-free sentences, support at s in GrIS coincides with support at $\mathsf{Best}(s)$ in flat information semantics. Moreover, the same is true for formulas containing $\Rightarrow$, as long as all updates required to interpret the relevant conditionals can be effected consistently. Thus, the extra structure provided by GrIS only becomes relevant when interpreting conditionals whose antecedents run against some of the information available in the state. In GrIS, unlike in flat information semantics, such conditionals can receive a non-trivial interpretation. This is possible because, in addition to the unconditional information encoded by the set $\mathsf{Best}(s)$, the graded set s also keeps track of some conditional information, captured by the division of $W - \mathsf{Best}(s)$ into different layers of implausibility.

[21] Notice that, like flat information semantics, GrIS has two semantic levels: the basic level of truth, where semantics is recursively defined, and the derived level of support, which matters for defining the relation of entailment. Two comments about this. First, given this architecture, the fact that two formulas are logically equivalent (i.e., are supported at the same states) does not guarantee that they are inter-substitutable in embedded contexts; thus, e.g., $p \equiv \Box p$, but $\neg p \not\equiv \neg\Box p$; however, formulas which are semantically equivalent (i.e., have the same truth-conditions) are of course inter-substitutable, since the semantics is compositional. Second, notice that there exists another natural notion of entailment for GrIS, namely, preservation of truth at a world-state pair (possibly restricted to pairs $\langle w, s\rangle$ with $w \in s$). While the support-based relation that we explore here is connected to preservation of acceptability at an information state—and thus to argumentation—the truth-based relation could matter for attitudes other than plain acceptance, e.g., judgments of probability. This would give us a natural way to address the concerns of Schulz (2010), who criticizes information semantics based on the fact that it makes p and $\Box p$ logically equivalent, although we might have different credence in these two sentences.

5.2 Predictions

5.2.1 Interactions between conditionals and their constituents

Desideratum 2.1 requires that a conditional $p \Rightarrow q$ should be consistent with, but should not follow from, the negation of its antecedent. Consider the following two states, where each world is represented by the set of atomic sentences true in it (thus '$\emptyset$' stands for a world w where p and q are false, 'p' for a world where p is true and q false, and 'pq' for a world in which both are true).

$$s_1 = \boxed{\emptyset \mid pq} \qquad s_2 = \boxed{\emptyset \mid p}$$

Both states supports $\neg p$, while only s_1 supports $p \Rightarrow q$. Thus, s_1 witnesses $\neg p, p \Rightarrow q \not\models \bot$, while s_2 witnesses $\neg p \not\models p \Rightarrow q$. This shows that Desideratum 2.1 is satisfied: unlike flat information semantics, GrIS is capable of interpreting in a non-trivial way conditionals whose antecedents are rejected in the state.

Desideratum 2.2 requires that we validate *modus ponens* and *modus tollens* for conditionals with factual antecedent and consequent. This is the case.

Proposition 5.7 (Modus ponens for factual conditionals)**.**
For all $\alpha, \beta \in \mathcal{L}_0 : \ \alpha, \alpha \Rightarrow \beta \models \beta$

Proof. Suppose s supports α and $\alpha \Rightarrow \beta$. Since $s \models \alpha$, $\mathsf{Best}(s) \subseteq |\alpha|$. Thus, updating s with α does not remove any of the worlds in $\mathsf{Best}(s)$, i.e., we have $\mathsf{Best}(s[\alpha]) = \mathsf{Best}(s)$. Since $s \models \alpha \Rightarrow \beta$, we have $s[\alpha] \models \beta$, so $\mathsf{Best}(s[\alpha]) \subseteq |\beta|$. This implies that $\mathsf{Best}(s) \subseteq |\beta|$, which means that s supports β. □

In fact, as we will discuss in more detail below, *modus ponens* is valid not only for factual consequents, but for all consequents which do not contain $\Rightarrow$, including those of the form $\Box\alpha$ and $\Diamond\alpha$. Next, let us turn to *modus tollens*.

Proposition 5.8 (Modus tollens for factual conditionals)**.**
For all $\alpha, \beta \in \mathcal{L}_0 : \ \alpha \Rightarrow \beta, \neg\beta \models \neg\alpha$

Proof. Suppose s supports $\alpha \Rightarrow \beta$ and $\neg\beta$. Towards a contradiction, suppose s does not support $\neg\alpha$. Then $\mathsf{Best}(s)$ contains some α-worlds. This means that updating s with α does not entirely eliminate the top layer of s, so we have $\emptyset \neq \mathsf{Best}(s[\alpha]) \subseteq \mathsf{Best}(s)$. Now take some $w \in \mathsf{Best}(s[\alpha])$. Since $w \in \mathsf{Best}(s[\alpha])$ and $s \models \alpha \Rightarrow \beta$, w must be a β-world. But since $w \in \mathsf{Best}(s[\alpha]) \subseteq \mathsf{Best}(s)$ and $s \models \neg\beta$, w must be a $\neg\beta$-world—a contradiction. □

5.2.2 Iterated conditionals

Let us now turn to the second set of desiderata, having to do with the interpretation of iterated conditionals. Desideratum 2.3 requires that a sequence of two antecedents may be collected into a single conjunctive antecedent without affecting the semantics of the conditional. This is indeed the case.

Proposition 5.9 (Import-export)**.**
$\alpha \wedge \beta \Rightarrow \varphi \equiv \alpha \Rightarrow (\beta \Rightarrow \varphi)$.

Proof. As in the case of flat information semantics, the proof just uses the fact that $|\alpha \wedge \beta| = |\alpha| \cap |\beta|$, so that $s[\alpha][\beta] = s[\alpha \wedge \beta]$. $\square$

Thus, like flat information semantics, but unlike minimal change semantics, GrIS provides a suitable treatment of iterated conditionals.

The second desideratum in this set is that the theory should allow us to make good sense of McGee's counterexample to *modus ponens*. This is something that neither minimal change semantics nor flat information semantics achieves. In order to analyze McGee's example, let the atomic sentences r, c, a stand respectively for *Reagan wins*, *Carter wins*, and *Anderson wins*, and let R and D stand respectively for *a Republican wins* and *a Democrat wins*. The natural representation of the relevant context is given by the following state:

$s =$ | rR | cD | aR |

In the most plausible worlds in s, a Republican wins, so we have $s \models R$. Updating the state with the assumption that a Republican wins leads to dropping the world where Carter wins, resulting in the following state:

$s[R] =$ | rR | aR |

Further updating with the assumption that Reagan does not win leaves us with a victory of Anderson as the only possibility:

$s[R][\neg r] =$ | aR |

Thus, the original state s supports the implication $R \Rightarrow (\neg r \Rightarrow a)$.

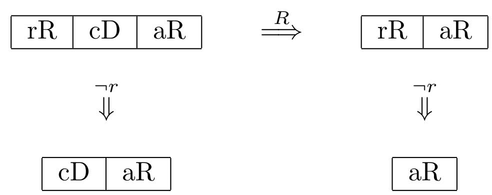
But now consider what happens when we update s with the assumption that Reagan does not not win. This yields:

$s[\neg r] =$ | cD | aR |

This state does not support the conclusion that Anderson will win (it supports the opposite conclusion), so the original state s does not support $\neg r \Rightarrow a$. Thus, the failure of *modus ponens* is predicted, and Desideratum 2.4 is satisfied. Moreover, it seems that the predictions align closely with our intuitions about the way we assess the relevant conditionals in the given context.

It is worth stopping to reflect for a moment on why *modus ponens* can fail for nested conditionals, while it is valid for all conditional-free consequents. The key remark is that, if $s \models \alpha$, then updating s with α does not remove any world from the top layer of s. Thus, the states s and $s[\alpha]$ are identical in their top layer. This layer, which encodes the state's unconditional information, is all that matters for the assessment of both factual sentences and epistemic modal sentences at s. Thus, if $s \models \alpha$, then for any $\Rightarrow$-free sentence φ we have $s \models \alpha \Rightarrow \varphi$ iff $s \models \varphi$. This guarantees that for all conditional-free consequents, *modus ponens* is valid.

At the same time, even when $s \models \alpha$, updating s with α is not a vacuous operation: in general, it affects the structure of the graded state by eliminating some of the lower-ranked possibilities. The difference between s and $s[\alpha]$ may well matter for the assessment of a conditional $\beta \Rightarrow \varphi$: if the antecedent β eliminates all worlds in the previous top layer, we may end up with states $s[\beta]$ and $s[\alpha][\beta]$ which differ in their top layers, and which therefore differ in their evaluation of factual sentences and epistemic modal sentences. This is what happens in the McGee example, as illustrated by the following diagram.



5.2.3 Interactions of conditionals and modals

As in flat information semantics, so also in GrIS a conditional antecedent restricts the state of evaluation, which provides the domain of quantification for modal expressions. This allows GrIS to inherit the success of flat information semantics in dealing with epistemic modals in the consequent of a conditional.

In particular, consider the marble scenario described in Section 2.3. The scenario can be modeled by a flat information state s containing two worlds:

$$s = \boxed{w, b}$$

Just as in flat information semantics, s supports $\neg w \Rightarrow \Box b$ and $\neg\Box b$, but does not support w, which shows that *modus tollens* can fail when the consequent contains a modal expression, in accordance with Desideratum 2.5.[22]

Next, consider the *or*-to-*if* inference. To see that $p \vee q \not\models \neg p \Rightarrow q$ in GrIS, consider a state which indicates the butler as the culprit, but takes no stance as to who, among the gardener and the chaffeur, did it if the butler is innocent:

$$s = \boxed{b}\boxed{g, c}$$

Since all the worlds in the top layer are b-worlds, this state supports $b \vee g$. But the state does not support $\neg b \Rightarrow g$, since supposing $\neg b$ results in the state $s[\neg b] = \boxed{g, c}$, which does not support the conclusion that the gardener did it.

Thus, GrIS accounts for the fact that inferring $\neg p \Rightarrow q$ from $p \vee q$ can be unwarranted in a context in which support for the disjunction $p \vee q$ is based on support for p. At the same time, GrIS predicts that the inference is justified if we have support for $p \vee q$ but *not* for p. Since disjunctions are typically used to describe the information available in cases where neither disjunct is individually

[22] A similar counterexample to *modus tollens* can be given, following Veltman (1985), for the case in which the consequent has the form of a factual conditional $q \Rightarrow s$.

supported, this can be taken to explain the strong intuitive appeal of the *or*-to-*if* inference, as proposed by Adams (1975).

Proposition 5.10 (Cautious or-to-if)**.**
$p \vee q, \Diamond\neg p \models \neg p \Rightarrow q$

Proof. Suppose $s \models \Diamond\neg p$ and $s \models p \vee q$. Since $s \models \Diamond\neg p$, some world in $\mathsf{Best}(s)$ is a $\neg p$-world; this means that updating s with $\neg p$ does not remove the top layer of s, so $\mathsf{Best}(s[\neg p]) \subseteq \mathsf{Best}(s)$. Since $s \models p \vee q$, we have $\mathsf{Best}(s) \subseteq |p \vee q|$, so also $\mathsf{Best}(s[\neg p]) \subseteq |p \vee q|$. Since no world in $\mathsf{Best}(s[\neg p])$ can be a p-world, all of them must be q-worlds. Thus, $s[\neg p] \models q$, which means that $s \models \neg p \Rightarrow q$. □

Thus, Desideratum 2.6 is satisfied in GrIS.

Finally, GrIS teases apart $p \Rightarrow \Diamond q$ and $\Diamond(p \wedge q)$, predicting that only the latter entails $\Diamond p$. To see this, consider the following information state:

$$s = \boxed{\emptyset}\boxed{p, pq}$$

Since the top layer of s contains no p-worlds, $s \not\models \Diamond p$. Updating with p leads to the state $s[p] = \boxed{p, pq}$, which supports $\Diamond q$: so, $s \models p \Rightarrow \Diamond q$. Hence, $p \Rightarrow \Diamond q$ does not entail $\Diamond p$, differing from $\Diamond(p \wedge q)$ as required by Desideratum 2.7.

Summing up, then, GrIS provides a relation of entailment that meets all the Desiderata laid out in Section 2.

Desideratum	2.1	2.2	2.3	2.4	2.5	2.6	2.7
Graded information semantics	✓	✓	✓	✓	✓	✓	✓

5.3 Conditionals and the universal quantifier

For simplicity, so far I have only considered propositional reasoning. In this section, I make a brief incursion into the connections with quantifiers to discuss an interesting desideratum pointed out by Yalcin (2015). For our purposes, it is not necessary to define a full-fledged predicate-logic version of GrIS; all we need is to suppose that our factual language $\mathcal{L}_0$ contains first-order formulas interpreted by means of standard truth-conditions. Following Yalcin (2015), as well as the syntax of natural languages, I will treat the universal quantifier $\forall x$ as a binary operator with a restrictor and a scope. Thus, a universal formula will have the form $\forall x(\varphi(x), \psi(x))$, and will be true at a world w iff all the individuals satisfying $\varphi(x)$ at w also satisfy $\psi(x)$ at w.

Consider the following pair of statements.

(17) a. Everyone who registered passed the test. $\forall x(Rx, Px)$
b. If Alice registered, she passed the test. $Ra \Rightarrow Pa$

Intuitively, (17-b) seems to follow from (17-a). However, as in the case of the *or*-to-*if* inference, there is a catch: if our acceptance of (17-a) is based crucially on the information that Alice did not register, then the inference is unjustified. Suppose that, according to our information, only Bob and Charlie registered, and they both passed. On this basis, we accept (17-a). Still, we have no reason to accept (17-b): rather than leading us to the conclusion that Alice passed the test, making the assumption that Alice registered results in a hypothetical state where we no longer accept the universal statement. This indicates that the inference from (17-a) from (17-b) is not valid in full generality.

On the other hand, the inference does seem correct in a context where our acceptance of (17-a) is not based on the information that Alice did not register—so that it is not undermined if we assume that she did. We can capture this by requiring as an extra premise the statement that Alice might have registered. This leads to the following desideratum.

Desideratum 5.11 (Cautious every-to-if)**.**

- $\forall x(Px, Qx) \not\models Pa \Rightarrow Qa$
- $\forall x(Px, Qx), \Diamond Pa \models Pa \Rightarrow Qa$

In GrIS, this desideratum is satisfied. To see that $\forall x(Px, Qx) \not\models Pa \Rightarrow Qa$, let w be a world where only Bob and Charlie registered, and where they both passed; and let w' be a world where, in addition, Alice also registered, but she did not pass. Then consider the following graded state:

$$s = \boxed{w \mid w'}$$

Clearly, this state supports $\forall x(Px, Qx)$ but does not support $Pa \Rightarrow Qa$.[23]

To see that the cautious version of the *every*-to-*if* inference is valid, notice that if $s \models \Diamond Pa$ then $\mathsf{Best}(s)$ contains some Pa-worlds; thus, an update with Pa will not remove the top layer of s, and we have $\mathsf{Best}(s[Pa]) \subseteq \mathsf{Best}(s)$. If we have $s \models \forall x(Px, Qx)$, this means that in any world in $\mathsf{Best}(s)$, the extension of P is included in the extension of Q. This holds, in particular, for the worlds in $\mathsf{Best}(s[Pa])$. Since in all these worlds a belongs to the extension of P, in these worlds a also belongs to the extension of Q. Thus, Qa is true in all worlds in $\mathsf{Best}(s[Pa])$. This means that $s[Pa] \models Qa$, which shows that $s \models Pa \Rightarrow Qa$.

Thus, as in the case of *or*-to-*if*, GrIS vindicates the *every*-to-*if* inference pattern in those circumstances where it seems sound, while also allowing us to recognize the kind of scenarios where this inference pattern can lead us astray.

6 Questioning the desiderata?

I argued for GrIS based on the desiderata laid out in Section 2. Whether these desiderata are achieved is a mathematical fact, but one may, of course, take issue with the desiderata themselves. In this section, I consider some natural objections that a flat information semanticist of the kind embracing antecedent compatibility may raise against desiderata 2.1, 2.4 and 2.7—the three desiderata that such an approach does not satisfy.

[23] I am assuming that first-order logic sentences have their usual truth-conditions, and that support for them still amounts to truth at all worlds in $\mathsf{Best}(s)$.

6.1 Non-triviality

The first problem that we identified for flat information semantics with the antecedent compatibility requirement is that, by requiring that $p \Rightarrow q$ can only be supported when $\Diamond p$ is the case, it makes $\neg p$ inconsistent with $p \Rightarrow q$. This contradicts our initial observation that we can be in a position to accept both $\neg p$ and $p \Rightarrow q$. However, the assumption of antecedent compatibility that is not without empirical motivation. Consider:

(18) Alice didn't go to Paris. ??If she did, she is staying with Bob.

This discourse strikes us as odd. Flat information semantics can explain the oddness as a case of inconsistency. By contrast, according to GrIS, (18) is a consistent discourse, so something else must be said about its infelicity.

I think the infelicity can be explained pragmatically, along the following lines: in normal circumstances, speakers are just concerned with drawing a distinction between the ways the world might be—the open possibilities—and the ways it is not—the possibilities which are ruled out. Normally, only the open possibilities are relevant to our purposes, and there is no point in drawing distinctions between possibilities which we do not take seriously anyway. Thus, the second sentence in (18) strikes us as a pointless conversational move.[24]

This, however, is not always the case. One case in which it becomes relevant to talk about possibilities that one does not take seriously is when one's interlocutor takes these possibilities seriously. For instance, suppose my interlocutor claims that Shakespeare never existed, and that Hamlet was written by T.S. Eliot. I might say:

[24]Given this explanation, it is natural to ask, with a reviewer, why the subjunctive counterpart of (18) is perfectly fine:

(i) Alice didn't go to Paris. If she had, she would have stayed with Bob.

On my view, the reason is that in (i), the conditional *is* used to draw a distinction between open possibilities. Among those worlds where Alice did not go to Paris, there are some where she would have stayed with Bob if she had gone, and others where she would not have. The subjunctive conditional conveys that the actual world is of the former kind.

The key difference between the two cases lies in the different hypothetical processes underlying indicative and subjunctive conditionals. Epistemic assumptions, involved in the interpretation of indicative conditionals, operate globally on an information state: given an information state s, an epistemic assumption of α leads us to the state $s[\alpha]$ obtained by restricting s to the α-worlds; if all the open possibilities in the conversation are $\neg\alpha$-worlds, all of them are eliminated when supposing α epistemically, so an indicative conditional $\alpha \Rightarrow \beta$ says nothing about them. Ontic assumptions, involved in the interpretation of subjunctive conditionals, operate primarily on worlds: given a world w, an ontic assumption of α takes us to another world (or set of worlds) $w\langle\alpha\rangle$; this is obtained, say, by an operation of intervention, as described in causal accounts (Pearl, 2000; Schulz, 2011; Briggs, 2012). The effect of an ontic assumption on a graded state s is then computed by applying this operation point-wise on the worlds in s, preserving the ranking. Even if all open possibilities w in the conversation are $\neg\alpha$ worlds, when supposing α ontically these worlds are not eliminated, but transformed into their α-counterparts, $w\langle\alpha\rangle$. A subjunctive conditional $\alpha \Rrightarrow \beta$ can then be used to draw a distinction between those open possibilities w such that $w\langle\alpha\rangle$ is a β-world, and the remaining ones. This explains why the discourse (i) is not odd in the way (18) is.

(19) Look, Hamlet was written by Shakespeare; but even if it wasn't, it was still written during the Elizabethan age. There is plenty of historical evidence for that.

Another reason to care about possibilities which we take to be non-actual is when the stakes for being wrong are high. For instance, imagine two engineers are discussing the safety of an elevator. One asks the other what will happen in the event of a blackout. The other replies:

(20) The emergency power generator will kick in. If that doesn't happen, the elevator will be blocked by a mechanical device.

Or consider two kids who just did something wrong. They discuss:

(21) a. What if mom finds out?
b. She won't. But if she does, I'll say it was your fault.

So, it seems reasonable to propose that a discourse of the form $\neg p; p \Rightarrow q$ is logically consistent, but sounds odd unless the speaker has some special motivation to care about the possibilities that they take to be non-actual.[25]

Now consider again an account implementing the antecedent compatibility requirement. This account will predict that a discourse of the form $\neg p; p \Rightarrow q$ is inconsistent. Therefore, a pragmatic explanation will be needed for why sequences of this kind are sometimes felicitous. Presumably the explanation will invoke some kind of context shift in these discourses, so that the first sentence and the second are not really accepted in the same state. I will not speculate in detail about the prospects of such an explanation here, but let me note two things that would have to be explained. First, such context shifts are usually assumed to broaden the relevant information state when we encounter an antecedent that cannot be consistently supposed. But this would not help with (22-b), where the conditional occurs first.

(22) a. What if mom finds out?
b. If she does, I'll say it was your fault. But she won't find out.

Second, if such context shifts are possible, one would have to explain why this does not rescue (18) from being odd in an ordinary context.

[25] Here are some examples of such discourses found on the internet (the last one from the novel *Frankenstorm*, by Ray Garton):

(i) a. It won't rain, but if it does rain it will be something soft and transparent.
b. She won't turn you down, but even if she does, you'll get over it.
c. Call Armie Hammer by his trackie, and he won't respond, but if he does respond, it will probably be in a burgundy velvet tuxedo.
d. And even if we get nailed, Clark—and that's *not* gonna happen, I'm telling ya—but if we do, the shit's gonna land on me, not you.

6.2 McGee's counterexample

Several proponents of flat information semantics (notably Gillies, 2004; Bledin, 2015) have argued that McGee's counterexample to *modus ponens* can be rejected once seen from the perspective of information semantics. In a nutshell, their argument goes as follows: the Republican candidates are Reagan and Anderson; therefore, if you *really* accept that a Republican will win, then you should accept that if Reagan does not will, Anderson will. In the given scenario, we do not accept the conditional only because we do not *really* accept the premise to begin with—we just take it to be likely. But likelihood is not, in this view, what logic is supposed to preserve: logic only preserves full acceptance, and it may well take us from something likely to something unlikely.

This argument sounds plausible, but I think it is wrong. The strong intuition that we should accept the conclusion rests on a hidden assumption about the situation: if *all* that you accept about the election is that a Republican will win, then you should indeed accept that if the winner is not Reagan, it will be Anderson. This is fully vindicated by GrIS, since as we saw we have that $r \vee a, \Diamond\neg r \models \neg r \Rightarrow a$. But now suppose what I accept is specifically that *Reagan* will win. Since Reagan is a Republican, I also accept that a Republican will win. But this does not commit me to accepting that if Reagan doesn't win, Anderson will. Taking the antecedent to be false just does not provide a reason to accept the conditional. And this is the case no matter how strong my conviction that Reagan will in fact win.

In fact, if the argument were right, by the same token I would be committed to accepting absurdities like "if Reagan doesn't win, Canada will declare war". For one could reason in the same way: I accept that Reagan will win; therefore, I accept that either Reagan will win or Canada will declare war; therefore, I must accept that if Reagan does not win, Canada will declare war. But surely it is clear that accepting that Reagan will win commits me to no such thing.

I believe the source of the problem here is the conflation of "full acceptance" with something like "irrevocable acceptance", acceptance that is so strong that one cannot consistently suppose to be wrong. While there may well be things that we accept in this irrevocable way, most of the things we accept do not fall in this category. So, if logic is to be useful, irrevocable acceptance had better not be the notion that logic is supposed to preserve.

6.3 Conditional might statements

Our Desideratum 2.7 was that $p \Rightarrow \Diamond q$ not come out equivalent to $\Diamond(p \wedge q)$: the latter entails $\Diamond p$, the former does not. Curiously, in one of seminal papers of flat information semantics, Gillies (2010) takes the opposite view, declaring the equivalence $p \Rightarrow \Diamond q \equiv \Diamond(p \wedge q)$ to be a desideratum. Gillies considers a pair of the following kind:

(23) a. If we win this match, we might win the World Cup. $w \Rightarrow \Diamond c$
b. We might win this match and the World Cup. $\Diamond(w \wedge c)$

According to Gillies, these sound equivalent. I disagree. Suppose we are at the knockout stages. I think our team is just average, and we are up against the world champions and current favorite. We hardly stand a chance. But, suppose we win: that would mean that we are much better than expected; the players' spirits will be high in the next matches; and all the remaining teams are not so strong. Based on these considerations, I wholeheartedly assert (23-a).

Now consider (23-b); since we are at the knockout stage, winning the world cup implies winning this match, so the conjunction $w \wedge c$ is equivalent to c. Thus (23-b) is just equivalent to (24):

(24) We might win the World Cup.

So, if Gillies is right, then when I asserted (23-a), my use of a conditional was redundant: I could have just asserted the consequent. But it is clear that in saying (23-a), the antecedent plays a crucial role and cannot be left out.

Granted, there is a sense in which any of the teams *might* win the Cup. But that is not what (23-a) conveys. What (23-a) conveys is that conditionally on winning this match, winning the World Cup becomes a *serious* possibility.

So, I think the very kind of examples that Gillies considers can in fact be used to see the difference between $p \Rightarrow \Diamond q$ and $\Diamond(p \wedge q)$ I want to bring out.[26]

7 Conditionals and entailment

Traditionally, the conditional operator bears a close relation to the meta-language relation of entailment, which is captured by the deduction theorem:

$$\overline{\varphi}, \psi \models \chi \quad \text{iff} \quad \overline{\varphi} \models \psi \Rightarrow \chi \tag{1}$$

In GrIS, however, neither direction of this bi-conditional holds. In one direction, we have $\neg p, p \models q$, although $\neg p \not\models p \Rightarrow q$. In the other direction, we have $p \Rightarrow (q \Rightarrow r) \models p \Rightarrow (q \Rightarrow r)$ but $p \Rightarrow (q \Rightarrow r), p \not\models q \Rightarrow r$.

As these examples show, the fact that GrIS breaks the deduction theorem is not accidental, but a direct consequence of the need to satisfy Desideratum 2.1 ($\neg p \not\models p \Rightarrow q$) and Desideratum 2.4 ($p \Rightarrow (q \Rightarrow r), p \not\models q \Rightarrow r$), while retaining classical logic for the factual fragment (which requires $\neg p, p \models \bot$) as well as the reflexivity of entailment (which requires $p \Rightarrow (q \Rightarrow r) \models p \Rightarrow (q \Rightarrow r)$).

However, the deduction theorem plays a key role in logic, as it establishes a direct connection between conditionals and logical inference—a connection that does seem to exist in some form. Thus, it would be nice if we could make the deduction theorem somehow compatible with our desiderata. Interestingly, in the context of GrIS this *can* be done, and in a way which, I believe, brings out interesting considerations about the significance of the entailment relation.

[26] Gillies also claims that $p \Rightarrow \Diamond q$ cannot felicitously be followed up by negating $\Diamond(p \wedge q)$. Negating a *might* statement directly is not easy, but in our example, it is natural to say:

(i) If we win this match, we might win the World Cup. But that's not gonna happen.

Let us start by asking: why does the deduction theorem fail in GrIS? The reason is that there is a difference between accepting and supposing a sentence. In our setting, having a sentence α as a premise corresponds to considering a context where it is accepted; on the other hand, having α as an antecedent involves the process of supposing it. Supposing and accepting are different things. Crucially, acceptance is defeasible: in a context where we accept p, we are still capable of consistently supposing $\neg p$ (as a genuine epistemic assumption, not just counterfactually). By contrast, suppositions are indefeasible: in a context where we suppose p, we are no longer capable of consistently supposing $\neg p$.

In order to restore the deduction theorem, what we seem to need is a more general notion of entailment with two kinds of premises: one kind corresponds to the sentences accepted in a context; the other corresponds to the sentences which are being supposed to be true. Given a sequence $\overline{\varphi} = \langle \varphi_1, \dots, \varphi_n \rangle$ of formulas, and a sequence $\overline{\alpha} = \langle \alpha_1, \dots, \alpha_m \rangle$ of factual formulas, we will write

$$\overline{\varphi};\ \overline{\alpha} \models \psi$$

to mean that, in a context where $\varphi_1, \dots, \varphi_n$ are accepted, supposing $\alpha_1, \dots, \alpha_n$ results in a state where ψ is accepted.

Formally, in GrIS a context will be a graded information state s, acceptance amounts to support, and supposition amounts to update. Thus, writing $s \models \overline{\varphi}$ to mean that $s \models \varphi_i$ for $i = 1, \dots, n$, and letting $s[\overline{\alpha}] := s[\alpha_1] \dots [\alpha_m]$, our more general notion of entailment will be defined as follows:[27]

Definition 7.1 (Generalized entailment)**.**
$\overline{\varphi};\ \overline{\alpha} \models \psi$ iff $\forall s : s \models \overline{\varphi}$ implies $s[\overline{\alpha}] \models \psi$

While one cannot consistently accept both p and $\neg p$, nor consistently suppose p and $\neg p$, one *can* consistently accept p and suppose $\neg p$, and conversely; in those cases, the supposition overrides the formula which is actually accepted.

Proposition 7.2 (Illustration of generalized entailment)**.**

- $p, \neg p; \models \bot$
- $; p, \neg p \models \bot$
- $p; \neg p \not\models \bot$
- $\neg p; p \not\models \bot$
- $p; \neg p \not\models p$
- $p; \neg p \models \neg p$
- $\neg p; p \not\models \neg p$
- $\neg p; p \models p$

More generally, a pair $\langle \overline{\varphi}; \overline{\alpha} \rangle$ is inconsistent if two contradictory premises occur within the same component (both within $\overline{\varphi}$ or both within $\overline{\alpha}$), but it can be consistent if they occur within different components. In that case, the conflict is resolved in favor of the second component: what is supposed overrides what is accepted.

[27] Note that, since we defined the notion of update only for factual formulas, only classical formulas can occur in the second component of an entailment relation.

This richer notion of entailment seems to be quite interesting both conceptually and formally. A detailed investigation must be left for future work, but—to conclude the discussion—let me point out that *modus ponens* and the deduction theorem can now be restored in full generality in the following form.

Proposition 7.3 (Modus ponens for GrIS)**.**

- $\alpha \Rightarrow \varphi;\ \alpha \models \varphi$

Proposition 7.4 (Deduction theorem for GrIS)**.**

- $\overline{\varphi};\ \overline{\alpha}, \beta \models \psi$ iff $\overline{\varphi};\ \overline{\alpha} \models \beta \Rightarrow \psi$

8 Relations with other accounts of conditionals

In this section I discuss the relations between GrIS and three prominent lines of work on conditionals: the line viewing if-clauses as restrictors of modals (Kratzer, 1981, 1986); the line connecting conditionals to belief revision (see, a.o., Gärdenfors, 1988; Levi, 1988; Rott, 1989; Arló-Costa and Levi, 1996; Arló-Costa, 1997; Leitgeb, 2010); and the line interpreting conditionals in terms of conditional probability (Adams, 1965, 1975).

8.1 Relations with the restrictor view of conditionals

One influential approach to conditionals which is related to GrIS is the restrictor theory of Kratzer (1981, 1986). This theory rejects the idea that *if* contributes a two-place connective $\Rightarrow$; instead, *if*-clauses can be seen on this view as modifiers of modal operators. Consider a sentence of the form "if A, then M C", where M is a modal expressing an operator O, and where A and C translate to φ and ψ, respectively. On the restrictor view, the *if*-clause acts as a modifier of the operator, turning O into a modified operator O_φ; this operator then combines with its argument to yield a sentential expression $O_\varphi(\psi)$.

Semantically, a modal is interpreted with respect to two parameters, a *modal base* f and a *ordering source* g, both of which are functions from worlds to sets of propositions (propositions, in turn, are sets of worlds). The modal base is used to determine a set of relevant worlds $s(w) := \bigcap f(w)$, while the ordering source is used to define a ranking (pre-ordering) of the worlds in $s(w)$: $v \leq_{g(w)} v'$ if every proposition from $g(w)$ which is true in v' is also true in v.

In the case of modals *must* and *might*, the clauses are as follows, where $\mathsf{Best}_{g(w)}(s(w))$ denotes the set of $\leq_{g(w)}$-minimal elements of $s(w)$:

- $w, f, g \models \Box\psi$ iff $\mathsf{Best}_{g(w)}(s(w)) \subseteq |\psi|^{f,g}$
- $w, f, g \models \Diamond\psi$ iff $\mathsf{Best}_{g(w)}(s(w)) \cap |\psi|^{f,g} \neq \emptyset$

The way in which an *if*-clause modifies an operator O is by adding the corresponding proposition to the modal base for O:

- $w, f, g \models O_\varphi(\psi)$ iff $w, f_\varphi, g \models O(\psi)$ where $f_\varphi(w) = f(w) \cup \{|\varphi|^{f,g}\}$

For those conditionals which do not involve any modal, the restrictor view postulates the presence of a silent occurrence of epistemic *must* at the level of logical form. Finally, entailment is defined as usual, as preservation of truth.

In some respects, the restrictor approach is similar to GrIS. As we saw, a modal base and ordering source are used to induce, respectively, a set of worlds and a pre-ordering of this set—thus yielding something very close to a graded state.[28] As in GrIS, modals are interpreted as quantifying over the top-ranked elements of the state, while if-clauses eliminate worlds from the state.

The restrictor approach shares most of the attractive features that we discussed for GrIS. For instance, if implemented in the way described here,[29] it gives an elegant account of iterated conditionals: a sequence of two if-clauses amounts to a double restriction of a modal, which has the same effect as a single restriction with a conjunction: $(O_\alpha)_\beta(\varphi) \equiv O_{\alpha\wedge\beta}(\varphi)$. Thus, *import-export* is accounted for. Similarly, the counterexamples to *modus ponens* and *modus tollens* can be accounted for much in the same way as in GrIS.

However, the restrictor approach also differs from GrIS in some important respects. In the restrictor approach, plain conditionals, whose truth depends on a relevant body of information, barely interact logically with non-modal sentences, whose truth depends only on the actual state of affairs.

This can be partly repaired by adding postulates like *factivity*—the requirement that $w \in \mathsf{Best}_{g(w)}(s(w))$—which, by tying the relevant information to the actual world, yields a form of *modus ponens* ($\Box_\alpha\beta, \alpha \models \beta$) and *modus tollens* ($\Box_\alpha\beta, \neg\beta \models \neg\alpha$) for bare conditionals (Schulz, 2018). However, factivity does not help with other interactions between conditionals and non-modal sentences. For example, consider the cautious version of the *or*-to-*if* inference, which on the restrictor view takes the form $p \vee q, \Diamond\neg p \models \Box_{\neg p} q$. The mere fact that $p \vee q$ is true at w, together with the fact that $\neg p$ holds in some top-ranked world, does not guarantee that all—or even some—of the top-ranked $\neg p$-worlds are q-worlds, which is what we need for the truth of $\Box_{\neg p} q$. Thus, Desideratum 2.6 is not satisfied by the restrictor approach.

The same problem arises, as pointed out by Yalcin (2015), with the cautious *every*-to-*if* inference discussed in Section 5.3. On the restrictor view, this takes the form $\forall x(Px, Qx), \Diamond Pa \models \Box_{Pa} Qa$. Again, this entailment is not validated: the fact that $\forall x(Px, Qx)$ is true at w, together with the fact that Pa holds in some top-ranked world, does not guarantee that all, or even some, of the

[28] A minor difference: the pre-order induced by an ordering source need not be total. But GrIS could also be generalized to let go of the totality requirement, if desired.

[29] This contrasts with the implementation which is most commonly found in the literature, which treats modals as two-places operators on propositions, which combine with the antecedent φ and the consequent ψ to yield a sentential expression $O(\varphi, \psi)$. It is not clear how this implementation can analyze iterated conditionals of the form "if A, then if B, C", given that the modal has only one antecedent slot. By analyzing an if-clause as a modifier of a modal, the above implementation avoids this problem.

top-ranked Pa-worlds are Qa-worlds. Thus, Desideratum 5.11 is not satisfied.[30]

From a theoretical point of view, GrIS is more parsimonious than the restrictor theory, since it obviates the need to stipulate silent epistemic modals in the LFs of bare conditionals—an assumption for which, as far as I know, no independent evidence has been provided.[31] On the other hand, the restrictor theory is more general, since it covers the way in which if-clauses can be used to restrict arbitrary modals, not just epistemic modals. While there does seem to be a straightforward way to extend GrIS to non-epistemic modals, I will leave the presentation of such an extension for another occasion.

8.2 Relation with theories based on belief revision

The main difference between flat and graded information semantics is that in flat information semantics a supposition always augments the available information, while in graded information semantics it might trigger a non-monotonic change. In this sense, GrIS fits within a family of approaches to conditionals developed by scholars working on belief revision (Gärdenfors, 1988; Levi, 1988; Rott, 1989; Arló-Costa and Levi, 1996; Arló-Costa, 1997; Leitgeb, 2010). These approaches are different from each other in the details, but they share the same central idea. Conditionals are to be interpreted by using tools for belief revision, in particular (i) a formal representation of belief states s, together with a specification of which sentences are accepted in a belief state s (to stay close to the notation in this paper, I will write this as $s \models \alpha$); and (ii) an update function $*$, mapping a belief state s and a sentence α to a new belief state $s * \alpha$. Conditionals are interpreted by specifying when they are accepted in a belief state. The acceptance condition is a natural formalization of the Ramsey test idea:

$$s \models \alpha \Rightarrow \varphi \text{ iff } s * \alpha \models \varphi$$

Notice that GrIS fits this scheme, when belief states are given as graded information states and $*$ is the operation $s * \alpha = s[\alpha]$. This fundamental convergence accounts for the many similarities existing between GrIS and these accounts. Notably, all these accounts make $p \Rightarrow q$ logically independent from $\neg p$, as required by Desideratum 2.1, since the fact that $\neg p$ is accepted in s says nothing about whether q will become accepted after s is revised to incorporate p.

However, there are also differences, especially when it comes to iterated conditionals, of the form $\alpha \Rightarrow (\beta \Rightarrow \varphi)$.[32] With the exception of Arló-Costa (1997), the accounts mentioned above do not validate the import-export principle, as required by Desideratum 2.3. This is not an accident: most of the above

[30] Related problems concerning the lack of logical interactions between epistemic modal sentences and non-modal sentences are discussed by Bledin and Lando (2018).

[31] See also Gillies (2010) on why specifying the distribution of such covert operators seems to be a non-trivial task.

[32] Some of these theories (e.g., Leitgeb, 2010), do not allow iterated conditionals in the syntax at all. In addition, unlike GrIS, none of the above theories treat the compositional interaction between conditionals and modals (Levi (1988) considers modals, but does not allow conditionals to embed them); I will therefore leave modals out of consideration here.

approaches interpret conditionals by means of an update function $*$ which is assumed to satisfy the so-called AGM postulates for belief revision (Alchourrón *et al.*, 1985). Now, recall that import-export requires that, for all α, β, φ:

$$\alpha \Rightarrow (\beta \Rightarrow \varphi) \equiv \alpha \wedge \beta \Rightarrow \varphi$$

Taking for granted that conditionals of the form $\gamma \Rightarrow \gamma$ are valid, import-export implies that the following must be valid for all α, β:

$$\alpha \Rightarrow (\beta \Rightarrow \alpha \wedge \beta)$$

As a special case, the following will have to be valid:

$$p \Rightarrow (\neg p \Rightarrow p \wedge \neg p)$$

This means that for every belief state s, $(s * p) * \neg p$ accepts a contradiction. This runs against one of the AGM postulates for belief revision, the *consistency* postulate, which demands that revising a consistent belief state with a consistent sentence should deliver a consistent belief state. Contrapositively, this shows that a theory of conditionals based on a belief revision operator satisfying the AGM postulates is necessarily bound to invalidate import-export.

In the face of this tension, what should we give up: the consistency postulate, or the cumulativity principle ($(s * \alpha) * \beta = s * (\alpha \wedge \beta)$) that underlies import-export? Neither. Rather, we should recognize that the two pertain two different processes: consistency is a requirement on a method for revising beliefs, while cumulativity is a feature of the process of making assumptions.

Consider the process of revising beliefs. When I revise my state to integrate p, I reach a state $s * p$ which accepts p, but only defeasibly so. If I later have reason to revise this state once more to incorporate $\neg p$, it ought to be possible to do so consistently, by giving up the previously acquired belief that p. Thus, belief revision is not a cumulative process: revising with α and then with β should not always result in a state which incorporates both α and β.

Making assumptions works differently. When I assume that p, I reach a hypothetical state which accepts p and no longer allows for the possibility that p is false (p is true "by assumption"). In this state, I can no longer consistently entertain the further assumption that $\neg p$. Assuming that if-clauses are pragmatically required to lead to a consistent hypothetical state, this provides an explanation of the fact that sentences like (25) are perceived as infelicitous:

(25) #If Alice is in Paris, then if she is not in Paris, we should invite her.

In constrast to belief revision, making assumptions *is* a cumulative process: assuming α and then β always leads to a hypothetical state which incorporates both. This is what underlies the validity of the import-export law.

The lesson to be drawn here is that, in a setting that allows iterated updates, it is important to distinguish carefully the process of actually revising beliefs from the process of making assumptions. The two work differently. Unlike the former, the latter is a cumulative process. Recognizing this difference is crucial in order to account for the logical behavior of conditionals, in particular for the validity of the import-export equivalence.[33]

This point has been stressed before by Arló-Costa (1997) and Segerberg (1998), who proposed systems which are closely related to GrIS.

Arló-Costa (1997) proposed a model of conditionals based on a map $*$ which is assumed to satisfy cumulativity. This account shares the core ideas of GrIS, but it is much more syntactic in nature. States and the operation $*$ are treated as primitives, and *import-export* is obtained as a consequence of an explicit stipulation that $(s * \alpha) * \beta = s * (\alpha \wedge \beta)$. By contrast, in GrIS we used an explicit model of states and of the update operation; the fact that $s[\alpha][\beta] = s[\alpha \wedge \beta]$ is not stipulated, but follows from the fact that making an assumption amounts to restricting the state to those worlds where the assumption is true.

The closest relative of GrIS is the system proposed by Segerberg (1998). Segerberg, too, interprets a formula relative to a world and a graded information state. Moreover, he also models the process of supposing α in terms of restricting the information state to the α-worlds; he calls this process the *irrevocable revision* of the state by α. However, Segerberg does not use this machinery to provide a semantics for conditionals, but rather to obtain a doxastic logic that describes, from an external perspective, what an agent believes under certain suppositions. Thus, for instance, in Segerberg's system the counterpart of our conditional $p \Rightarrow q$ would be the formula:

$$[*p]Bq$$

to be read as: *under the supposition of* p*, the agent believes* q. An important difference is that Segerberg does not have a semantic level corresponding to the notion of support in GrIS; entailment is simply defined as preservation of truth. Accordingly, not only the syntax of the system, but also the logic differs from the one of GrIS in a non-trivial way. For instance, although the truth-conditions for the formulas p, q and $p \Rightarrow q$ in GrIS coincide with those for p, q, and $[*p]Bq$ in Segerberg's system, in GrIS we have the entailment $p, p \Rightarrow q \models q$, while in Segerberg's system we don't have the corresponding entailment $p, [*p]Bq \models q$. Similarly, the GrIS formula $p \wedge \Diamond \neg p$ has the same truth-conditions as $p \wedge B \neg p$, but the former is a contradiction in GrIS, while the latter is consistent in Segerberg's system (and this is as it should be, given the respective intended interpretations).

Nevertheless, the relation is tight enough that it might be possible to give a translation from GrIS to Segerberg's system. It is easy to see, however, that such a translation cannot be a straightforward compositional one.[34]

[33] In fact, in GrIS, the consistency postulate can fail even when making a single assumption. Even when α is consistent, a state s may contain no α-worlds; in this case, $s[\alpha]$ will be the inconsistent state. Thus, GrIS invalidates the consistency postulate even in the non-iterated case. I think this is desirable for an account of indicative conditionals, since it allows us to model the case in which an assumption, while logically consistent, is too far fetched to be consistently entertained as genuine indicative assumptions, as in (i):

(i) ??If I am eating sushi right now, I am enjoying it.

[34] Note, for instance that in GrIS we have $p \equiv \Box p$ but $\neg p \not\equiv \neg \Box p$. Thus, p and $\Box p$ must be

8.3 Relation with the probabilistic theory

According to the view of conditionals defended by Adams (1965, 1975) (see also Edgington, 1986, 1995), conditionals do not have truth-values, but only probability values, given by a probability function P which maps formulas to values in the real interval $[0, 1]$. The central tenet of this view is the thesis that the probability of a conditional is given by the conditional probability of the consequent given the antecedent:

$$P(\alpha \Rightarrow \beta) = P(\beta|\alpha) := \frac{P(\alpha \wedge \beta)}{P(\alpha)} \tag{2}$$

Of course, if conditionals do not have truth-conditions, entailment for a language including conditionals cannot be defined in terms of truth-preservation. Instead, Adams proposes a probabilistic notion of entailment. The idea is that an entailment holds if the conclusion can be guaranteed to be highly probable provided the assumptions are sufficiently probable. More precisely:

$$\varphi_1, \ldots, \varphi_n \models \psi \quad \text{iff} \quad \begin{array}{l} \forall \varepsilon > 0\ \exists \delta > 0 \text{ s.t. for all probability assignments } P: \\ \text{if } P(\varphi_i) \geq 1 - \delta \text{ for } i = 1, \ldots, n, \text{ then } P(\psi) \geq 1 - \varepsilon \end{array}$$

It is easy to see that, with this notion of entailment, our desiderata 2.1 and 2.2 are satisfied. In general, probabilistic entailment has the merit of matching extremely well with intuitions about inference patterns involving conditionals. A disadvantage of this approach, though, is that only a restricted language $\mathcal{L}_A$ can be interpreted in this way (cf. Adams, 1975, §1.8): only conditionals of the form $\alpha \Rightarrow \beta$ with $\alpha, \beta \in \mathcal{L}_0$ are allowed in the language; moreover, conditionals are only allowed as main operators in a sentence, and cannot be embedded under other logical operators. Since our remaining desiderata (2.3-2.7) have to do with iterated conditionals and the connections between conditionals and modals, and since the language $\mathcal{L}_A$ contains neither iterated conditionals nor modals, the question of whether these desiderata are satisfied does not arise for the probabilistic theory.

In some important respects, GrIS is akin to the probabilistic theory. First, in both theories, conditionals differ from factual sentences in that they do not have truth-values relative to states of affairs, but depend for their interpretation on an epistemic component of the semantics (a probability function in one case, a graded information state in the other). Moreover, the two theories are similar in that the process of making assumptions is non-monotonic: making an assumption can in general take us from a state where something is accepted (i.e., supported/highly probable) to a new hypothetical state where it is no longer accepted; this is what is responsible for the non-entailment $\neg p \not\models p \Rightarrow q$ in these theories. In a sense, GrIS can be seen as a qualitative version of the probabilistic theory, where the quantitative probability assignment is replaced by a qualitative plausibility ordering among worlds.

translated to equivalent formulas, yet their negations must be translated to non-equivalent formulas. Since Segerberg's logic allows replacement of equivalents, this is impossible if the translation of $\neg\varphi$ is defined recursively from the translation of φ.

These similarities account for the fact that the two theories converge on many predictions. In fact, it is not hard to see, using results from Kraus *et al.* (1990) (see also Leitgeb, 2010), that GrIS entailment coincides with probabilistic entailment relative to the restricted language $\mathcal{L}_A$: with respect to this language, both systems characterize the logic known as system P. Thus, one could view GrIS as providing a way to extend the logic of the probabilistic theory beyond the fragment that can be provided with a probabilistic interpretation. This extension is interesting, since it allows us to consider inference patterns such as those that we have been concerned with in this paper, as well as patterns involving conditionals embedded under connectives.

9 Negating conditionals

It is not easy to assess directly how conditionals behave under negation; in English, we can only express such negations by means of circumlocutions like:

(26) It is not the case that if Alice comes she will come with Bob.

Here are two things (26) could mean:

1. that if Alice comes she will not come with Bob $\quad p \Rightarrow \neg q$
2. that if Alice comes, she will not *necessarily* come with Bob $\quad p \Rightarrow \neg\Box q.$

If we analyze (26) as expressing $\neg(p \Rightarrow q)$, then what GrIS predicts is the second reading: provided that $s[p]$ is consistent, $s \models \neg(p \Rightarrow q)$ iff $s \models p \Rightarrow \neg\Box q$.

Is this prediction correct? Some empirical work by Egré and Politzer (2013) suggests that it is—at least for some conditionals. But there are also arguments for the opposite conclusion. For instance, to avoid the unnatural circumlocution "it is not the case that" and get better intuitive predictions, one may look at expressions which involve negation, such as the attitude verb *doubt* (see, e.g., Santorio, 2017). The following two sentences seem intuitively equivalent:

(27) a. I doubt that Bob passed if he took the test.
b. I believe that Bob failed if he took the test.

How can we account for this? Let us translate *doubt* and *believe* by operators D and B, and let us assume that $D\varphi \equiv B\neg\varphi$. If $\neg(p \Rightarrow q)$ is semantically equivalent to $p \Rightarrow \neg q$, the perceived equivalence between (27-a) and (27-b) is predicted:

$$D(p \Rightarrow q) \equiv B\neg(p \Rightarrow q) \equiv B(p \Rightarrow \neg q)$$

Other items whose semantics involves negation, such as the quantifier *no* and the exhaustivity operator *only* have been argued to provide further evidence for the equivalence $\neg(p \Rightarrow q) \equiv p \Rightarrow \neg q$ (see Higginbotham, 1986; von Fintel, 1997; Klinedinst, 2010; Santorio, 2017).

This paper is not the place to settle this empirical issue. However, suppose that, contrary to the predictions of GrIS, the equivalence $\neg(p \Rightarrow q) \equiv p \Rightarrow \neg q$ should be rendered valid. How could we modify GrIS to get this result?

One option is to revise the system by defining two separate notions of *truth* ($\models^+$) and *falsity* ($\models^-$) at a world-state pair $\langle w, s\rangle$.[35] As derivative, we would have two notions of *support* and *rejection* at an information state, defined as:

- $s \models^+ \varphi$ iff $\forall w \in \mathsf{Best}(s) : \langle w, s\rangle \models^+ \varphi$
- $s \models^- \varphi$ iff $\forall w \in \mathsf{Best}(s) : \langle w, s\rangle \models^- \varphi$

The definition of truth and falsity would go as follows:

- $\langle w, s\rangle \models^+ p$ iff $w(p) = 1$
- $\langle w, s\rangle \models^- p$ iff $w(p) = 0$
- $\langle w, s\rangle \models^+ \neg\varphi$ iff $\langle w, s\rangle \models^- \varphi$
- $\langle w, s\rangle \models^- \neg\varphi$ iff $\langle w, s\rangle \models^+ \varphi$
- $\langle w, s\rangle \models^+ \varphi \wedge \psi$ iff $\langle w, s\rangle \models^+ \varphi$ and $\langle w, s\rangle \models^+ \varphi$
- $\langle w, s\rangle \models^- \varphi \wedge \psi$ iff $\langle w, s\rangle \models^- \varphi$ or $\langle w, s\rangle \models^- \varphi$
- $\langle w, s\rangle \models^+ \Box\alpha$ iff $s \models^+ \alpha$
- $\langle w, s\rangle \models^- \Box\alpha$ iff $s \not\models^+ \alpha$
- $\langle w, s\rangle \models^+ \alpha \Rightarrow \varphi$ iff $s[\alpha] \models^+ \varphi$ and $s[\alpha] \neq s_\bot$
- $\langle w, s\rangle \models^- \alpha \Rightarrow \varphi$ iff $s[\alpha] \models^- \varphi$ and $s[\alpha] \neq s_\bot$[36]

This system, call it $\mathsf{GrIS}^\pm$, preserves the basic insights of GrIS, and it makes the same predictions as GrIS with respect to the entailments that we have been concerned with in this paper. Moreover, it straightforwardly predicts the equivalence $\neg(\alpha \Rightarrow \varphi) \equiv \alpha \Rightarrow \neg\varphi$, since it is easy to check that these two formulas are assigned exactly the same truth and falsity conditions. Thus, $\mathsf{GrIS}^\pm$ looks like a plausible alternative to GrIS. However, a detailed assessment of this system must be left for another occasion.

[35] This follows the pattern of a number of bi-lateral semantics, defining a positive and a negative semantic relation side by side. Examples include data semantics (Veltman, 1981, 1985), and some implementations of truth-maker semantics (Fine, 2017), inquisitive semantics (Bledin, 2018) and dynamic semantics (Willer, 2018).

[36] The extra condition $s[\alpha] \neq s_\bot$ here is intended to prevent conditionals whose antecedents are not consistently supposable from counting as trivially supported. If that were the case, both $\bot \Rightarrow p$ and $\bot \Rightarrow \neg p$ would come out as theorems; given that in $\mathsf{GrIS}^\pm$ we have $\bot \Rightarrow \neg p \equiv \neg(\bot \Rightarrow p)$, this would mean that a pair of sentences of the form φ and $\neg\varphi$ could both be logically valid—a pretty strange prediction. This, however, should not worry us too much, since as we discussed in Footnotes 16 and 19, there are independent reasons to think that conditionals can only be accepted provided their antecedent is consistently supposable. Thus, the semantic modification above, which allows us to avoid the problematic predictions, is independently motivated.

10 Conclusion

Assessing a conditional involves supposing the antecedent and assessing the consequent in the resulting hypothetical state. I have proposed an implementation of this common idea, where the information states that are involved in the evaluation are *graded*, i.e., distinguish not only between worlds which are ruled in and worlds which are ruled out by the available information, but also between worlds which are ruled out to different degrees. I have shown that the resulting system, *graded information semantics*, gives rise to a logic that satisfies an array of desiderata which are not jointly satisfied by existing accounts: it allows us to deal with conditionals whose antecedent runs against the information accepted by the state; it yields the desired interpretation for iterated conditionals, involving multiple *if*-clauses; it predicts that *if*-clauses can restrict the domains of epistemic modals; it validates *modus ponens* and *modus tollens* for factual consequents, while also predicting the existence of counterexamples in just those configurations in which such counterexamples seem to arise; finally, it predicts in what circumstances an indicative conditional can be reliably inferred from a disjunction or a universal statement.

Obviously, much work remains to be done. Let me mention four important directions. First, the logic of GrIS should be investigated systematically, and related to the logics arising from other approaches to conditionals. Second, conditionals with disjunctive antecedents are widely taken to require a fine-grained representation of their antecedents which is not available in GrIS (Alonso-Ovalle, 2009; Ciardelli, 2016; Santorio, 2018; Khoo, 2018). Third, one should look beyond the restricted syntactic fragment considered here; in particular, it would be interesting to consider conditionals embedded under epistemic modals and attitude verbs, which present interesting challenges (Gillies, 2018). Finally, it seems interesting to determine to what extent the logical desiderata for subjunctive conditionals are the same as for indicative conditionals, and to ask whether a unified story can be told.

References

Adams, E. (1965). The logic of conditionals. *Inquiry*, **8**(1-4), 166–197.

Adams, E. (1970). Subjunctive and indicative conditionals. *Foundations of Language*, **6**(1), 89–94.

Adams, E. (1975). *The logic of conditionals: An application of probability to deductive logic*, volume 86. Springer Science & Business Media.

Alchourrón, C., Gärdenfors, P., and Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, **50**(02), 510–530.

Alonso-Ovalle, L. (2009). Counterfactuals, correlatives, and disjunction. *Linguistics and Philosophy*, **32**, 207–244.

Arló-Costa, H. (1997). Belief revision conditionals: models of suppositional reasoning. Technical Report, Carnegie Mellon University.

Arló-Costa, H. and Levi, I. (1996). Two notions of epistemic validity. *Synthese*, **109**(2), 217–262.

Baltag, A. and Smets, S. (2006). Dynamic belief revision over multi-agent plausibility models. In *Proceedings of LOFT*.

Bledin, J. (2014). Logic informed. *Mind*, **123**(490), 277–316.

Bledin, J. (2015). Modus ponens defended. *The Journal of Philosophy*, **112**(2), 57–83.

Bledin, J. (2018). Fatalism and the logic of unconditionals. Noûs.

Bledin, J. and Lando, T. (2018). Closure and epistemic modals. *Philosophy and Phenomenological Research*, **97**(1), 3–22.

Board, O. (2004). Dynamic interactive epistemology. *Games and Economic Behavior*, **49**(1), 49 – 80.

Briggs, R. (2012). Interventionist counterfactuals. *Philosophical studies*, **160**(1), 139–166.

Cantwell, J. (2008). Changing the modal context. *Theoria*, **74**(4), 331–351.

Ciardelli, I. (2016). Lifting conditionals to inquisitive semantics. In M. Moroney, C.-R. Little, J. Collard, and D. Burgdorf, editors, *Semantics and Linguistic Theory (SALT) 26*, Ithaca, NY. LSA and CLC Publications.

Dekker, P. (1993). *Transsentential Meditations. Ups and Downs in Dynamic Semantics.* Ph.D. thesis, ILLC, University of Amsterdam.

Edgington, D. (1986). Do conditionals have truth conditions? *Crítica: Revista Hispanoamericana de Filosofía*, pages 3–39.

Edgington, D. (1995). On conditionals. *Mind*, **104**, 235–329.

Egré, P. and Politzer, G. (2013). On the negation of indicative conditionals. In M. Aloni, M. Franke, and F. Roelofsen, editors, *Proceedings of the Nineteenth Amsterdam Colloquium*, pages 10–18.

Fine, K. (2017). Truthmaker semantics. *A companion to the philosophy of language*, pages 556–577.

von Fintel, K. (1997). Bare plurals, bare conditionals, and *only*. *Journal of Semantics*, **14**(1), 1–56.

von Fintel, K. and Gillies, A. (2010). Might made right. In A. Egan and B. Weatherson, editors, *Epistemic Modality*. Oxford University Press.

Gärdenfors, P. (1988). *Knowledge in Flux: Modeling the Dynamics of the Epistemic States.* MIT Press, Cambridge, MA.

Gibbard, A. (1980). Two recent theories of conditionals. In *Ifs: Conditionals, beliefs, decision, chance, time*, pages 211–247. Reidel, Dordrecht.

Gillies, A. (2004). Epistemic conditionals and conditional epistemics. *Noûs*, **38**(4), 585–616.

Gillies, A. (2009). On truth-conditions for if (but not quite only if). *Philosophical Review*, **118**(3), 325–349.

Gillies, A. S. (2007). Counterfactual scorekeeping. *Linguistics and Philosophy*, **30**(3), 329–360.

Gillies, A. S. (2010). Iffiness. *Semantics and Pragmatics*, **3**, 4–1.

Gillies, A. S. (2018). Updating Data Semantics. *Mind*.

Grove, A. (1988). Two modellings for theory change. *Journal of philosophical logic*, **17**(2), 157–170.

Herzberger, H. G. (1979). Counterfactuals and consistency. *The Journal of Philosophy*, **76**(2), 83–88.

Higginbotham, J. (1986). Linguistic theory and davidson's program in semantics. In E. Lepore, editor, *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson.*, pages 29–48. Basil Blackwell.

Khoo, J. (2018). Disjunctive antecedent conditionals. *Synthese*.

Klinedinst, N. (2010). Quantified conditionals and conditional excluded middle. *Journal of Semantics*, **28**(1), 149–170.

Kolodny, N. and MacFarlane, J. (2010). Ifs and oughts. *The Journal of philosophy*, **107**(3), 115–143.

Kratzer, A. (1981). The notional category of modality. In H. Eikmeyer and H. Rieser, editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. Berlin: Walter de Gruyter.

Kratzer, A. (1986). Conditionals. *Chicago Linguistics Society*, **22**(2), 1–15.

Kraus, S., Lehmann, D., and Magidor, M. (1990). Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial intelligence*, **44**(1-2), 167–207.

Leitgeb, H. (2010). On the ramsey test without triviality. *Notre Dame Journal of Formal Logic*, **51**(1), 21–54.

Levi, I. (1988). Iteration of conditionals and the Ramsey test. *Synthese*, **76**(1), 49–81.

Lewis, D. (1973). *Counterfactuals*. Blackwell.

McGee, V. (1985). A counterexample to modus ponens. *The Journal of Philosophy*, **82**(9), 462–471.

Pearl, J. (2000). *Causality: models, reasoning and inference*, volume 29. Springer.

Price, H. (1983). Does ?probably? modify sense? *Australasian Journal of Philosophy*, **61**(4), 396–408.

Ramsey, F. P. (1929). General propositions and causality. In H. Mellor, editor, *F. Ramsey, Philosophical Papers*. Cambridge University Press.

Rott, H. (1989). Conditionals and theory change: Revisions, expansions, and additions. *Synthese*, **81**(1), 91–113.

Santorio, P. (2017). Conditional excluded middle in informational semantics. In T. v. G. Alexandre Cremers and F. Roelofsen, editors, *Proceedings of the 21st Amsterdam Colloquium*, pages 385–394.

Santorio, P. (2018). Alternatives and truthmakers in conditional semantics. *The Journal of Philosophy*, **115**(10), 513–549.

Schulz, K. (2011). If you'd wiggled a, then b would've changed. *Synthese*, **179**(2), 239–251.

Schulz, M. (2010). Epistemic modals and informational consequence. *Synthese*, **174**(3), 385–395.

Schulz, M. (2018). Modus ponens under the restrictor view. *Journal of Philosophical Logic*, pages 1–28.

Segerberg, K. (1998). Irrevocable belief revision in dynamic doxastic logic. *Notre Dame journal of formal logic*, **39**(3), 287–306.

Stalnaker, R. (1968). A theory of conditionals. In N. Rescher, editor, *Studies in Logical Theory*. Blackwell, Oxford.

Stalnaker, R. (1976). Indicative conditionals. In *Ifs*, pages 193–210. Reidel.

Starr, W. (2014). A uniform theory of conditionals. *Journal of Philosophical Logic*, **43**(6), 1019–1064.

Veltman, F. (1981). Data semantics. In J. Groenendijk, T. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*. Mathematical Centre.

Veltman, F. (1985). *Logics for Conditionals*. Ph.D. thesis, University of Amsterdam.

Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic*, **25**(3), 221–261.

Von Fintel, K. (1998). The presupposition of subjunctive conditionals. *The interpretive tract*, **25**, 29–44.

Willer, M. (2014). Dynamic thoughts on ifs and oughts. *Philosophers' Imprint*, **14**(28), 1–30.

Willer, M. (2017). Lessons from sobel sequences. *Semantics and Pragmatics*, **10**.

Willer, M. (2018). Simplifying with free choice. *Topoi*, **37**(3), 379–392.

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

Yalcin, S. (2011). Nonfactualism about epistemic modality. In A. Egan and B. Weatherson, editors, *Epistemic Modality*. Oxford University Press.

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

Yalcin, S. (2015). Epistemic modality de re. *Ergo, an Open Access Journal of Philosophy*, **2**.