Intro to Conditionals

Day 1 // Conditionals and Information-Sensitivity, ESSLLI // David Boylan and Matt Mandelkern

1 Basics of Indicatives

Indicative conditionals:

- Examples:

  (1) If the coin was flipped it landed heads.

  (2) If DB doesn't teach tomorrow, MM will.

- We assume that, in very many cases, "if ... then..." can be treated as a constituent whose semantic value is an operator on propositions.
  - Call this operator $>$.

Argument that $> \neq \supset$: Only four of the 16 truth-functions from pairs of truth-values to a truth-value could possibly be the indicative.

At least if we restrict ourselves to a bivalent setting. See Egre, Rossi and Sprenger 2021 on the de Finetti/Reichenbach trivalent approach.

- In the case of $\ulcorner T > T \urcorner$ it seems clear we want T.
- In the case of $\ulcorner T > F \urcorner$ it seems clear we want F.
- In the case of $\ulcorner F > T \urcorner$ and $\ulcorner F > F \urcorner$, it's harder to say.

But in fact, none of these could be the indicative:

- Suppose both $\ulcorner F > T \urcorner$ and $\ulcorner F > F \urcorner$ are T.
  - Then we have the material conditional, together with its familiar 'paradoxes'.
    E.g. $\neg(p > q)$ entails $p \wedge \neg q$:

    I.e. $p > q \equiv p \supset q$, true iff p is false or q true.

    (3) Not (if Patch is a rabbit, she is a rodent) $\not\models$ Patch is a rabbit and she is not a rodent.

    (4) Nothing is a rodent if it is a rabbit $\not\models$ Everything is a rabbit.

- Suppose both are F.

  Though curiously some experimental studies show that people often interpret these as equivalent.

  - Then $p > q \equiv p \wedge q$. But

    (5) If Patch is a rabbit, she is a rodent $\neq$ Patch is a rabbit and a rodent.

- Suppose $\ulcorner F > T \urcorner$ is T and $\ulcorner F > F \urcorner$ is F.
  - Then $p > q$ entails q.

- Suppose $\ulcorner F > T \urcorner$ is F and $\ulcorner F > F \urcorner$ is T.
  - Then $(p > q) \wedge \neg p$ entails $\neg q$.

2 A baseline semantics for indicatives: Stalnaker 1968

Ramsey proposes that, in assessing $p > q$, you add p

> hypothetically to your stock of knowledge (or beliefs), and then consider whether or not q is true. Your belief about the conditional should be the same as your hypothetical belief, under this condition, about the consequent.

Needs refinement when you already believe the antecedent to be false: then you must 'make whatever adjustments are required to maintain consistency'

How to map this intuition about how we *evaluate* a conditional's truth-value to its *truth-conditions*?

> Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. ⌜If A, then B⌝ is true (false) just in case B is true (false) in that possible world. Stalnaker 1968

2.1 Multi-modal logic

Review: start with a *frame*, a sequence $\langle W, R_1, R_2, \ldots R_n \rangle$

- W is a set of "possible worlds";
- the R_i are binary accessibility relations between worlds.

A model $\mathcal{M}$ adds an atomic valuation $\mathcal{I}_\mathcal{M}$ which takes a world and atomic sentence to a truth value. Then we specify the interpretation $[\![]\!]_\mathcal{M}$ as follows, for $w \in K$:

Relativization to models is suppressed for readability.

- where A is atomic, $[\![A]\!]^w = 1$ iff $\mathcal{I}_\mathcal{M}(A, w) = 1$
- $[\![A \wedge B]\!]^w = 1$ iff $[\![A]\!]^w = 1$ and $[\![B]\!]^w = 1$
- $[\![\neg A]\!]^w = 1$ iff $[\![A]\!]^w = 0$
- For $i \in [1, n]$: $[\![\Box_i A]\!]^w = 1$ iff $\forall w' \in R_i(w) : [\![A]\!]^{w'} = 1$
- $[\![\Diamond_i A]\!]^w = [\![\neg\Box_i\neg A]\!]^w = 1$ iff $\exists w' \in R_i(w) : [\![A]\!]^{w'} = 1$

We write $R(w)$ for $\{w' : wRw'\}$.

The result is sound and complete with respect to K, the smallest normal modal logic, comprising PC, closed under detachment for $\supset$ and:

- *Necessitation*: from $\vdash A$ conclude $\vdash \Box_i A : i \in [1, n]$
- K: $\vdash \Box(A \supset B) \supset (\Box A \supset \Box B)$

2.2 Variably strict semantics

Variably strict semantics treat conditionals as necessity operator, with the flavor of necessity varying by antecedent.

- So $A > C$ can be treated as abbreviating $\Box_A C$, which corresponds to an $[\![A]\!]$-accessibility relation R_A.

Compress these indexed accessibility relations into a *selection function* f.

Intuitively, the closest $[\![A]\!]$-worlds to w.

- Given a proposition $[\![A]\!]$ and a world w, $f([\![A]\!], w)$ yields $R_A(w)$
- $[\![A > C]\!]^w = 1$ iff $\forall w' \in f([\![A]\!], w) : [\![C]\!]^{w'} = 1$

Recall CK, which comprises *PC* plus detachment for $\supset$ and

- *CN*: from $\vdash C$ conclude $\vdash A > C$
- *CK*: $\vdash (A > (B \supset C)) \supset ((A > B) \supset (A > C))$
- *LLE*: $\vdash A > C \equiv B > C$ whenever $\vdash A \equiv B$

The model so far is sound and complete with respect to CK. This is easy to see from the corresponding result for K.

2.3 Constraining the selection function

Some basic constraints on the selection function:

- *Success*: $f(\varphi, w) \subseteq \varphi$

this makes f a world-relative 'choice function'

  ⇝ *Identity*: $\vdash A > A$
- *Weak Centering*: $w \in \varphi \rightarrow w \in f(\varphi, w)$

  ⇝ *Modus Ponens*: $\{A > C, A\} \vdash C$
- *Strong Centering*: $w \in \varphi \rightarrow f(\varphi, w) \subseteq \{w\}$

  ⇝ *Conjunctive Suffiency:* $A \wedge C \vdash A > C$

A further important constraint: CSO.

- *cso*: $(f(\varphi, w) \subseteq \psi \wedge f(\psi, w) \subseteq \varphi) \rightarrow f(\varphi, w) = f(\psi, w)$

  ⇝ *CSO*: $\vdash ((A > B) \wedge (B > A) \wedge (A > C)) \supset B > C$
  - if the closest A's are B's, and the closest B's are A's, they are the same.

This is plausibly intrinsic to the notion of closeness/choice.

- As P. Schlenker points out, a weaker property suffices:

  *cso**: $f(\varphi, w) \subseteq \psi \rightarrow f(\varphi, w) = f(\varphi \cap \psi, w)$

Equivalent to the conjunction of *Cautious Monotonicity* and *Cautious Transitivity*, i.e. $A > B \supset (A > C \equiv (A \wedge B) > C)$

- This is entailed by *cso* given *Success*.
- Conversely assume *cso**
- assume $f(\varphi, w) \subseteq \psi \wedge f(\psi, w) \subseteq \varphi$.
- By *cso** and $f(\psi, w) \subseteq \varphi$, we have $f(\psi, w) = f(\varphi \cap \psi, w)$
- By *cso** and $f(\varphi, w) \subseteq \psi$, we have $f(\varphi, w) = f(\varphi \cap \psi, w)$
- Hence $f(\varphi, w) = f(\psi, w)$

This simpler formulation brings out the closeness to Sen's condition α.

A final, very controversial condition:

◦ *Uniqueness*: $|f(\varphi, w)| \leq 1$

⇝ *CEM*: $\vdash (A > B) \vee (A > \neg B)$

◦ Putative counterexample:

(8) If the coin is flipped it will land heads.

(9) If the coin is flipped it will land tails.

Some think neither conditional is true here.

Many find these kinds of examples more convincing for counterfactuals:

(6) If Verdi hadn't been Italian, he would have been French.

(7) If Verdi hadn't been Italian, he would have been Spanish.

Which is true??

◦ Possible response: it's simply vague which one is true.

- Many mainstream approaches to vagueness (e.g. supervaluationism, epistemicism) are consistent with the validity of excluded middle principles.

2.4 Orderings

Alternative formulation of the baseline semantics: *orderings over worlds*.

◦ For any w, let $\leq_w$ be a reflexive, transitive and symmetric relation between worlds.

- Intuitively, $w_1 \leq_w w_2$ iff w_1 is more similar, or "closer", to w than w_2.

There is much debate about how to analyse this notion of closeness and indeed whether it can be analysed at all. We will not get into this though.

◦ On the ordering semantics, conditionals consider similar *accessible* worlds.

- Assume we are given some accessibility relation R.
    - · We assume that accessible worlds are always more similar to w than non-accessible worlds

◦ So $[\![p > q]\!]^w = 1$ iff either

1. there are no p-worlds R-accessible from w;
2. for any p-world w_1 R-accessible from w there is some other p-world w_2 such that $w_2 \leq_w w_1$ and for any p-world w_3 where $w_3 \leq_w w_2$, q is also true at w_3.

- Given well-foundedness, this amounts to $[\![p > q]\!]^w = 1$ iff $min([\![p]\!], w) \subseteq [\![q]\!]$

Where $min(A, w) = \{w_1 \in A \mid \neg\exists w_2 : w_2 \in A \text{ and } w_2 <_w w_1\}$

Any selection function semantics can be translated into the ordering semantics:

- Define an ordering from the selection function as follows:
  - wRw' iff $f(\{w'\}, w) = \{w'\}$
  - $w, x \in R(z) \rightarrow (w <_z x \equiv \forall\varphi : w, x \in \varphi \rightarrow x \notin f(\varphi, z))$.
  - $w \in R(z), x \notin R(z) \rightarrow w <_z x$
- Given all the above, f corresponds to a function $<$ from a world w to a strict well-order on accessible worlds.
- W/o *Uniqueness*, $<$ is a well-founded, strict partial order.

I.e. irreflexive, transitive, asymmetric.

But the ordering semantics also allows for non-well-founded orderings.

reject the Limit Assumption

- Is this extra expressivity ever needed? Lewis said *yes*.
  - Consider a one-inch line. What world is most similar to the actual one (the one inch line) but where the line is longer?
    - Well-foundedness says that there must be a most similar world!
  - Lewis claims instead there is a continuous series of closer and closer worlds.
- The logical consequences of denying the Limit Assumption are iffy:
  - In Lewis's model, (10) is true for every $\epsilon \in \mathbb{R}$:

    (10) If the line were longer than one inch, it would be no more than ϵ longer than one inch.

  - But (10) is weird. If you don't have well-foundedness, we lose the principle that, if $\Gamma \models q$, then $\bigwedge_{r \in \Gamma} p > r$ entails $p > q$.

Herzberger showed these are equivalent, I think.

3 Logic

$>$ is intermediate between the strict conditional $\Box(A \supset B)$ and material.

- To see stronger than material, suffices to note that $\ulcorner F > F \urcorner$ is not truth-functional
  - Suppose D and M both false at w. But $f(D, w) \in T$ and $f(M, w) \in \neg T$.
- To see weaker than strict, think through failures:

- Transitivity: $\{A > B, B > C\} \not\models A > C$.
- Strengthening the Antecedent: $\{A > B\} \not\models (A \wedge C) > B$:
- Contraposition: $A > B \not\models \neg B > \neg A$.

(i) 'If Edgar Hoover were today a communist, then he would be a traitor.' (ii) 'If Hoover had been born a Russian, he would be a communist.' (iii) 'If Hoover had been a Russian, he would be a traitor.'
'If the match were struck, it would light' doesn't seem to entail 'If the match had been soaked in water overnight and it were struck, it would light.'

◦ Note MT still valid: $\{A > B, \neg B\} \models \neg A\}$.

Relationship with 'might' conditionals:

◦ Lewis regiments his language with *two* conditional operators:

Officially just a theory of counterfactuals.

(11) $\ulcorner p \Diamond\!\!\rightarrow q \urcorner =_{df} \ulcorner$If p, then it might be that q$\urcorner$

◦ Lewis assumes the two conditionals are interdefinable:

On analogy with $\Box$ and $\Diamond$ on their standard treatment.

Duality: $\neg(p > q) =_{df} p \Diamond\!\!\rightarrow \neg q$

(12) If it rains, the picnic will be canceled ≡ Not: if it rains, the picnic might be canceled.

We can't have CEM and Duality. Given CEM, $\neg(p > q) \models p > \neg q$, so this would make $p \Diamond\!\!\rightarrow \neg q$ entail $p > \neg q$.

4 Indicatives and Subjunctives

Return to minimal pairs of indicatives and subjunctives:

We regiment the subjunctive conditional with an operator $\Box\!\!\rightarrow$.

◦ (13) a. If Oswald doesn't shoot Kennedy, then nobody else will.
b. If Oswald hadn't shot Kennedy, nobody else would have.

Similarities between indicatives and subjunctives:

◦ Across many languages, both kinds of conditional get expressed use the same kind of construction (e.g "if... then...")

◦ Similar logic:

Not necessarily identical, as we will discuss on days 2 and 3.

- *Prima facie*, good case for Identity, Modus Ponens, Conjunctive Sufficiency, CSO and CEM.
- AS, Transitivity, Contraposition all seem invalid for $\Box\!\!\rightarrow$.

Although some deny CS for subjunctives; and some find CEM (even) more implausible for subjunctives.

Differences:

◦ Extra "subjunctive" morphology. Specific morphology varies across languages.

In many languages, this morphology is not. the subjunctive mood.

- An extra layer of (past) tense and/or (imperfective) aspect, in English and other languages.

Imperfective aspect is required in Greek.

  - In some languages, e.g. Hungarian, a dedicated particle, added to both antecedent and consequent.

- Different presuppositions:

Will be discussed further on day 2.

  - Indicatives seem to presuppose the antecedent is a live possibility, in some sense.
  - Subjunctives clearly do not; but hard to exactly pin down what if anything they suggest.
    - Do not presuppose counterfactuality:

(14) If Alice had taken arsenic, she would display exactly *those* symptoms.

From Anderson.

- Oswald pair above shows indicatives and subjunctives are not equivalent.

(Surely oversimplified) hypothesis about the relationship:

- Meaning of subjunctives is composed out of applying some operator, call it $*$ to the indicative conditional.
  - So $[\![p \mathrel{\Box\!\!\rightarrow} q]\!] = *([\![p > q]\!])$

Some basic ideas about what $*$ could mean:

- Could be a device for altering the modal flavour of the conditional:
  - Indicatives seem related to the modality expressed by “may" or “might".
  - Subjunctives seem related to the modality expressed by “could have".
- Could just be the past tense.
  - Prima facie difficulties:

(15) *Uttered before the murderer is revealed.*

If the gardener didn’t murder the Count, then the butler murdered the Count.

(16) *Uttered after the murderer is revealed to be the butler.*

If the gardener hadn’t murdered the Count, then the butler would have murdered the Count.

  - Can’t *just* be the past of what ordinary indicatives express. Some extra story about flavour is needed too.

5 Three Grades of Information-Sensitivity

General idea of info-sensitivity:

- what an indicative sentence means and whether it is true partially depends on an information state.

First grade of info-sensitivity: at the level of the context.

- Basic contextualism: the semantic value of an indicative conditional sentence depends on the context of utterance.
- Simple way to achieve this: when $\ulcorner p > q \urcorner$ is asserted by Alice, the *p*-worlds must be consistent with Alice's knowledge/evidence.

To state other grades, helpful to distinguish between semantic values and. assertoric contents:

- Semantic values are the kinds of things that can compose, particularly with various operators in the language.
- Assertoric contents are the kinds of things that get asserted, that are known or believed, that are added to the common ground...

Prima facie case that these are not the same:

- In the grand scheme of things, the semantic interpretation function takes many parameters as inputs:
  - worlds, times, variable assignments, ...
- Not clear that the objects of knowledge or belief are sensitive to these kinds of parameters.

Second grade of information-sensitivity: info-sensitivity at the level of semantic values.

Days 2, 3 and 5

- Conditionals require some kind of information-state parameter to do non-trivial compositional work.
  - Various implementations: domain semantics, local context semantics, dynamic semantics...

Third grade: at the level of assertoric contents.

Day 5

- Conditional propositions, the assertoric contents of conditional sentences, are more complex than just sets of possible worlds.

  Rather they are sets of objects like world-information state pairs.
  - Most familiar from certain relativist and/or self-locating views of epistemic modals and conditionals.