La logique des conditionnels

Rentrée Cogmaster

M. Cozic & P. Egré

IHPST/Paris 1, CNRS, DEC-ENS
IJN, CNRS, DEC-ENS

What are conditional sentences?

If P then Q

(1) If it's a square, then it's rectangle.

(2) If you strike the match, it will light.

(3) If you had struck the match, it would have lit.

Role of conditionals in mathematical, practical and causal reasoning.

Why care about conditionals?

- Also because they are a challenge for natural language semantics
- Natural language semantics: “semantics with no treatment of truth conditions is not semantics" (D. Lewis 1970)
- So what are the truth-conditions of conditionals?

Antecedent and consequent

(4) If P then Q

P: antecedent, protasis
Q: consequent, apodosis

Conditionals without "if...then..."

- **Imperative** (Bhatt and Pancheva 2005)

(5) a. Kiss my dog and you'll get fleas.
b. If p, q.

(6) a. Kiss my dog or you'll get fleas.
b. If $\neg p$, q.

- **No...No...** (Lewis 1972)

(7) a. No Hitler, no A-bomb
b. If there had been no Hitler, there would have been no A-bomb.

- **Unless**

(8) a. Unless you talk to Vito, you will be in trouble.
b. If you don't talk to Vito, you will be in trouble.

How to analyze conditional sentences?

Two main options we shall discuss today:

- Conditionals as truth-functional binary connectives: material conditional
- Conditionals as non-truth-functional, but truth-conditional binary connectives: Stalnaker-Lewis

Indicative vs. Subjunctive conditionals

- Another issue:

  (9) If Oswald did not kill Kennedy, someone else did.

  (10) If Oswald had not killed Kennedy, someone else would have.

The Material Conditional

Reminder: negation, conjunction and disjunction

$$\phi := \ p \mid \neg\phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid (\phi \rightarrow \phi) |$$

$[\![\phi]\!]$	$[\![(\neg\phi)]\!]$
1	0
0	1

$[\![\phi]\!]$	$[\![\psi]\!]$	$[\![(\phi \wedge \psi)]\!]$
1	1	1
1	0	0
0	1	0
0	0	0

$[\![\phi]\!]$	$[\![\psi]\!]$	$[\![(\phi \vee \psi)]\!]$
1	1	1
1	0	1
0	1	1
0	0	0

The material conditional

- Sextus Empiricus, *Adv. Math., VIII*: Philo used to say that the conditional is true when it does not start with the true to end with the false; therefore, there are for this conditional three ways of being true, and one of being false
- Frege to Husserl 1906: Let us suppose that the letters 'A' and 'B' denote proper propositions. Then there are not only cases in which A is true and cases in which A is false; but either A is true, or A is false; *tertium non datur*. The same holds of B. We therefore have four combinations:

  A is true and B is true
  A is true and B is false
  A is false and B is true
  A is false and B is false

  Of those the first, third and fourth are compatible with the proposition "if A then B", but not the second.

The truth-functional analysis

$[\![\phi]\!]$	$[\![\psi]\!]$	$[\![(\phi \to \psi)]\!]$
1	1	1
1	0	0
0	1	1
0	0	1

- $[\![\phi \to \psi]\!] = 0$ iff $[\![\phi]\!] = 1$ and $[\![\psi]\!] = 0$
- $[\![\to]\!] = cond : \{0,1\} \times \{0,1\} \to \{0,1\}$
  $cond(x,y) = 0$ iff $x = 1$ and $y = 0$
- $[\![\phi \to \psi]\!] = [\![\neg(\phi \wedge \neg\psi)]\!]$

Binary Boolean functions

		f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}
1	1	0	0	0	0	1	0	0	0	1	1	1	0	1	1	1	1
1	0	0	0	0	1	0	0	1	1	0	0	1	1	0	1	1	1
0	1	0	0	1	0	0	1	0	1	0	1	0	1	1	0	1	1
0	0	0	1	0	0	0	1	1	0	1	0	0	1	1	1	0	1
p	q	$\bot$	$\nmid$	$\nleftarrow$	$\nrightarrow$	$\wedge$	$\neg p$	$\neg q$	$\leftrightarrow$	$\nleftrightarrow$	q	p	$\mid$	$\rightarrow$	$\leftarrow$	$\vee$	$\top$

- Assuming a two-valued logic, and the conditional to be a binary connective: no other boolean function is a better candidate to capture the conditional's truth-conditions
- At least: the material conditional captures the falsity conditions of the indicative conditional of natural language.

Propositional validity

- ϕ is a tautology or logical truth iff $[\![\phi]\!]$=1 for all assignment of truth-value to the propositional atoms of ϕ. ($\models \phi$)
- ϕ is a logical consequence of a set Γ of formulae iff every assignment of truth-value that makes all the formulae of Γ true makes ϕ true. ($\Gamma \models \phi$)

"Good" validities

- $\phi \rightarrow \psi, \phi \models \psi$ (modus ponens)
- $\phi \rightarrow \psi, \neg\psi \models \neg\phi$ (modus tollens)
- $(\phi \vee \psi) \models \neg\phi \rightarrow \psi$ (Stalnaker's "direct argument"; aka disjunctive syllogism)
- $\models (((\phi \wedge \psi) \rightarrow \chi) \leftrightarrow (\phi \rightarrow (\psi \rightarrow \chi)))$ (import-export)
- $\models [(\phi \vee \psi) \rightarrow \chi] \leftrightarrow [(\phi \rightarrow \chi) \wedge (\psi \rightarrow \chi)]$ (simplification of disjunctive antecedents)

"Bad" validities

- $\neg\phi \models (\phi \rightarrow \psi)$ (falsity of the antecedent)
- $\phi \models (\psi \rightarrow \phi))$ (truth of the consequent)
- $(\phi \rightarrow \psi) \models (\neg\psi \rightarrow \neg\phi))$ (contraposition)
- $(\phi \rightarrow \psi), (\psi \rightarrow \chi) \models (\phi \rightarrow \chi))$ (transitivity)
- $(\phi \rightarrow \psi) \models ((\phi \wedge \chi) \rightarrow \psi)$ (antecedent strengthening)
- $\models \neg(\phi \rightarrow \psi) \leftrightarrow (\phi \wedge \neg\psi)$ (negation)

Why “bad" validities?

Undesirable validities w.r.t. natural language and ordinary reasoning:

- “Paradoxes of material implications" (Lewis). The paradox of the truth of the antecedent:

(11) a. John will teach his class at 10am.
b. ??Therefore, if John dies at 9am, John will teach his class at 10am.

(12) a. John missed the only train to Paris this morning and had to stay in London.
b. ??So, if John was in Paris this morning, John missed the only train to Paris this morning and had to stay in London.

Contraposition, Strengthening, Transitivity

(13) a. If Goethe had lived past 1832, he would not be alive today.
b. ??If Goethe was alive today, he would not have lived past 1832.

(14) a. If John adds sugar in his coffee, he will find it better.
b. ??If John adds sugar and salt in his coffee, he will find it better.

(15) a. If I quit my job, I won't be able to afford my apartment. If I win a million, I will quit my job.
b. ??If I win a million, I won't be able to afford my apartment. (Kaufmann 2005)

Qualms about non-monotonicity

- Does order matter? (von Fintel 1999)

(16) If I win a million, I will quit my job. ??If I quit my job, I won't be able to afford my apartment.

(17) If the US got rid of its nuclear weapons, there would be war. But if the US and all nuclear powers got rid of their weapons, there would be peace.

(18) If the US and all nuclear powers got rid of their nuclear weapons there would be peace; ?? but if the US got rid of its nuclear weapons, there would be war.

- Non-monotony seems less consistent when conjuncts are reversed.

Negation of a conditional

(19) a. It is not true that if God exists, criminals will go to heaven.
b. (??) Hence God exists, and criminals won't go to heaven.

The expected understanding of negation is rather:

(20) If God exists, criminals won't go to heaven.

(21) $\neg$(if p then q) = if p then $\neg q$

Several diagnoses

- The examples raise a problem for the pragmatics of conditionals, and do not call for a revision of the semantics. (Quine 1950 on indicative conditionals, Grice 1968, Lewis 1973).
- The examples call for a revision of the semantics of conditionals (Quine 1950 on counterfactual conditionals, Stalnaker 1968, Lewis 1973)

Limits of truth-functionality

Whatever the proper analysis of the contrafactual conditional may be, we may be sure in advance that it cannot be truth-functional; for, obviously ordinary usage demands that some contrafactual conditionals with false antecedents and false consequents be true and that other contrafactual conditionals with false antecedents and false consequents be false (Quine 1950)

(22) If I weighed more than 150 kg, I would weigh more than 100 kg.

(23) If I weighed more than 150 kg, I would weigh less than 25 kg.

Suppose I weigh 70 kg. Then the antecedent and consequent of both conditionals are presently false (put in present tense), yet the first is true, the second false.

Stalnaker's logic

Stalnaker's analysis: background

How do we evaluate a conditional statement?

- *First, add the antecedent hypothetically to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent; finally, consider whether or not the consequent is then true. (Stalnaker 1968)*
- *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)*

Central ideas

Let us review Stalnaker's semantics for "if ϕ then ψ", the core of all other non-monotonic conditional semantics:

- Either ϕ holds at the actual world: then the truth conditions of $\phi > \psi$ are those of the material conditional.
- Or ϕ does not hold at the actual world: $w \models \neg\phi$, and so $w \models \phi > \psi$ iff ψ holds at the closest ϕ-world to w.

Stalnaker's logic

- $\phi := p \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi \rightarrow \phi \mid \phi > \phi$
- Stalnaker-Thomason model: $M = \langle W, R, I, f, \lambda \rangle$, where $\langle W, R, I \rangle$ is reflexive Kripke model, λ absurd world (inaccessible from and with no access to any world), and $f : \wp(W) \times W \rightarrow W$ is a selection function satisfying:

(cl1) $f(\llbracket \phi \rrbracket, w) \in \llbracket \phi \rrbracket$

(cl2) $f(\llbracket \phi \rrbracket, w) = \lambda$ only if there is no w' s.t. wRw' and $w' \in \llbracket \phi \rrbracket$

(cl3) if $w \in \llbracket \phi \rrbracket$, then $f(\llbracket \phi \rrbracket, w) = w$

(cl4) if $f(\llbracket \phi_2 \rrbracket, w) \in \llbracket \phi_1 \rrbracket$ and $f(\llbracket \phi_1 \rrbracket, w) \in \llbracket \phi_2 \rrbracket$, then $f(\llbracket \phi_2 \rrbracket, w) = f(\llbracket \phi_1 \rrbracket, w)$

(cl5*) if $f(\llbracket \phi \rrbracket, w) \neq \lambda$, then $f(\llbracket \phi \rrbracket, w) \in R(w)$

Semantics

(i) $M, w \models p$ iff $w \in I(p)$
(ii) $M, w \models \neg\phi$ iff $M, w \not\models \phi$
(iii) $M, w \models (\phi \wedge \psi)$ iff $M, w \models \phi$ and $M, w \models \psi$
$M, w \models (\phi \vee \psi)$ iff $M, w \models \phi$ or $M, w \models \psi$
$M, w \models (\phi \rightarrow \psi)$ iff $M, w \not\models \phi$ or $M, w \models \psi$
(iv) $M, w \models (\phi > \psi)$ iff $M, f(\llbracket \phi \rrbracket, w) \models \psi$

- For every formula ϕ: $M, \lambda \models \phi$.

Looking at the clauses

- cl1 ensures that $\phi > \phi$, cl3 that no adjustment is necessary when the antecedent already holds at a world.
- cl2 and cl5*: selected world is absurd when antecedent is impossible.
- cl4: coherence on the ordering induced by the selection function.

Axiomatics

Stalnaker's C2

$\Box\phi =_{df} (\neg\phi > \phi)$
$\Diamond\phi =_{df} \neg(\phi > \neg\phi)$
$(\phi <> \psi) =_{df} ((\phi > \psi) \wedge (\psi > \phi))$
(PROP) All tautological validities
(K) $(\Box\phi \wedge \Box(\phi \rightarrow \psi)) \rightarrow \Box\psi$
(MP) From ϕ and $(\phi \rightarrow \psi)$ infer ψ
(RN) From ϕ infer $\Box\phi$
(a3) $\Box(\phi \rightarrow \psi) \rightarrow (\phi > \psi)$
(a4) $\Diamond\phi \rightarrow ((\phi > \psi) \rightarrow \neg(\phi > \neg\psi))$
(a5) $(\phi > (\psi \vee \chi)) \rightarrow ((\phi > \psi) \vee (\phi > \chi))$
(a6) $((\phi > \psi) \rightarrow (\phi \rightarrow \psi))$
(a7) $((\phi <> \psi) \rightarrow ((\phi > \chi) \rightarrow (\psi > \chi))$

Important consequence

- $\models (\phi > \psi) \rightarrow (\phi \rightarrow \psi)$
- Stalnaker's conditional is stronger than the material conditional

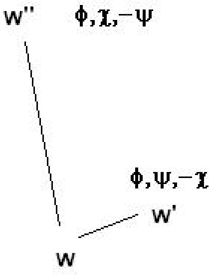
Invalidities

None of the “bad" validities comes out valid in Stalnaker's logic

- (FA) $\neg\phi \nvDash (\phi > \psi)$
- (TC) $\phi \nvDash (\psi > \phi))$
- (C) $(\phi > \psi) \nvDash (\neg\psi > \neg\phi)$
- (S) $(\phi > \psi) \nvDash ((\phi \wedge \chi) > \psi)$
- (T) $(\phi > \psi), (\psi > \chi) \nvDash (\phi > \chi)$

Example: monotonicity failure

- $(\phi > \psi) \not\models ((\phi \wedge \chi) > \psi)$.
- Take $w' = f(\llbracket \phi \rrbracket, w)$, such that $w' \models \psi$, and $w'' = f(\llbracket \phi \wedge \chi \rrbracket, w)$, such that $w'' \not\models \psi$.



Weak monotonicity

Monotonicity is lost, but a weakened form is preserved:

(CV) $(((\phi > \psi) \wedge \neg(\phi > \neg\chi)) \rightarrow ((\phi \wedge \chi) > \psi)))$

Positive properties

- Negation: $\Diamond\phi \models \neg(\phi > \psi) \leftrightarrow (\phi > \neg\psi)$
- Conditional excluded middle: $(\phi > \psi) \vee (\phi > \neg\psi)$
- Modus ponens: $\phi, (\phi > \psi) \models \psi$

Validities lost

Some of the “good" validities are lost in both Stalnaker's and Lewis's system:

- Import-Export: $\not\models (\phi > (\psi > \chi)) \leftrightarrow (\phi \wedge \psi > \chi)$ (both directions)
- Simplification of Disjunctive Antecedents: $\not\models (\phi \vee \psi > \chi) \rightarrow (\phi > \chi) \wedge (\psi > \chi)$
- Disjunctive Syllogism: $\phi \vee \psi \not\models \neg\phi > \psi$

Examples from Natural Language

IE
a. If Mary leaves, then if John arrives, it won't be a disaster.
b. If Mary leaves and John arrives, it won't be a disaster.

SDA
a. If Mary or John leaves, it will be a disaster.
b. If Mary leaves, it will be a disaster, and if John leaves, it will be a disaster.

DS The car took left, or it took right. Hence, if it did not take left, it took right.

What can be done?

- All such failures have been discussed: IE (McGee 1989), SDA (Alonso-Ovalle 2004, Klinedinst 2006), DS (Stalnaker 1975, see Lecture 5).
- The problem with all schemata: they all drive monotonicity back
- SDA: suppose $[\![\phi']\!] \subset [\![\phi]\!]$. Then $\phi \vee \phi' \equiv \phi$. If $\phi \vee \phi' > \chi \rightarrow \phi > \chi \wedge \phi' > \chi$, then $\phi > \chi \rightarrow \phi' > \chi$. (Klinedinst 2006).

The case of SDA

Klinedinst (2006: 127): *the problem is that what seems to be wanted is a semantics for conditionals that is both downward monotonic for disjunctive antecedents (at least in the normal case), but non-monotonic for antecedents in general.*

(24) If John had married Susan or Alice, he would have married Alice.

(25) If John had taken the green pill or the red pill – I don't remember which, maybe even both –, he wouldn't have gotten sick.

Klinedinst's proposal: the problem is pragmatic, and concerns our use of disjunction.

What should you remember?

- You saw two distinct semantic analyses of the conditional of natural language: truth-functional (=truth-tables) and non truth-functional (= here, possible world semantics)
- Key element of Stalnaker's analysis: non-monotonicity of conditionals
- Tradeoff when going from one analysis to the other: bad validities are lost, but some good ones too.