Introduction to the Semantics of Conditionals

Conditionals in Decision Theory
Lecture 2

Calum McNamara and Paolo Santorio
IU and UMD

ESSLLI 2026, Prague

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

- Day 1: Introduction to Decision Theory.

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

- Day 1: Introduction to Decision Theory.
- Day 2: Introduction to the Semantics of Conditionals.

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

- Day 1: Introduction to Decision Theory.
- Day 2: Introduction to the Semantics of Conditionals.
- Day 3: Decision-making under Determinism.

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

- Day 1: Introduction to Decision Theory.
- Day 2: Introduction to the Semantics of Conditionals.
- Day 3: Decision-making under Determinism.
- Day 4: Triviality and Tenability.

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

- Day 1: Introduction to Decision Theory.
- Day 2: Introduction to the Semantics of Conditionals.
- Day 3: Decision-making under Determinism.
- Day 4: Triviality and Tenability.
- Day 5: Bonus Topics (time dependent)

Oops...

One thing we omitted to do yesterday was provide a detailed roadmap for the course.

Here it is:

- Day 1: Introduction to Decision Theory.
- Day 2: Introduction to the Semantics of Conditionals.
- Day 3: Decision-making under Determinism.
- Day 4: Triviality and Tenability.
- Day 5: Bonus Topics (time dependent)

Today (Day 2), we'll start with a brief recap of yesterday's material.

A plea

Before that, however...

A plea

Before that, however...

Please ask questions!

A "realistic" Newcomb case

Yesterday, we talked about the Newcomb problem. Here's a slightly more realistic case, which brings out many of the same features.

> *Imagine you're wondering whether you have a specific heart condition. If you have the condition, then, unfortunately, you'll die prematurely. (Sorry!) You know that this condition is wide-spread among philosophers, although well-confirmed studies show there's no causal connection. (Perhaps, instead, the heart condition and a predilection for academia are symptoms of a common cause.) You must choose between a career as an academic and some other career—say, a career as an athlete. Additionally, you would prefer the life of an academic to the life of an athlete, both in the case in which you have the disease and the case in which you don't. But at the same time, you would* **strongly** *prefer to live a long life than a short one. What should you choose?*

The decision table

Here's the corresponding decision table:

	Disease	*No disease*
Life as an academic	*A fulfilling but short life*	*A fulfilling, long life*
Life as an athlete	*An unfulfilling and short life*	*An unfulfilling, long life*

The decision table

Here's the corresponding decision table:

	Disease	*No disease*
Life as an academic	*A fulfilling but short life*	*A fulfilling, long life*
Life as an athlete	*An unfulfilling and short life*	*An unfulfilling, long life*

Question 1: Which option should you choose, intuitively?

The decision table

Here's the corresponding decision table:

	Disease	*No disease*
Life as an academic	*A fulfilling but short life*	*A fulfilling, long life*
Life as an athlete	*An unfulfilling and short life*	*An unfulfilling, long life*

Question 1: Which option should you choose, intuitively?

Question 2: Which option does EDT recommend? How about CDT?

Evidential decision theory

Recall that evidential decision theory says that you should choose the option that's *correlated* with the best outcome.

Evidential decision theory

Recall that evidential decision theory says that you should choose the option that's *correlated* with the best outcome.

Formally, it says you should maximize:

$$\mathrm{EEU}(O) = \sum_i p(S_i \mid O) \cdot u(O \cap S_i)$$

Evidential decision theory

Recall that evidential decision theory says that you should choose the option that's *correlated* with the best outcome.

Formally, it says you should maximize:

$$\mathrm{EEU}(O) = \sum_i p(S_i \mid O) \cdot u(O \cap S_i)$$

Now let's assume that:

$$\begin{aligned} u(A \textit{ long, fulfilling life}) &> u(A \textit{ long, unfulfilling life}) \\ > u(A \textit{ short, fulfilling life}) &> u(A \textit{ short, unfulfilling life}) \end{aligned}$$

Evidential decision theory

Recall that evidential decision theory says that you should choose the option that's *correlated* with the best outcome.

Formally, it says you should maximize:

$$\mathrm{EEU}(O) = \sum_i p(S_i \mid O) \cdot u(O \cap S_i)$$

Now let's assume that:

$$\begin{aligned} u(A \textit{ long, fulfilling life}) &> u(A \textit{ long, unfulfilling life}) \\ > u(A \textit{ short, fulfilling life}) &> u(A \textit{ short, unfulfilling life}) \end{aligned}$$

Let's also assume that:

$$\begin{aligned} p(A \textit{ short, fulfilling life} \mid \textit{Life as an academic}) &\approx 1 \\ p(A \textit{ long, unfulfilling life} \mid \textit{Life as an athlete}) &\approx 1 \end{aligned}$$

Question: With this information, can you guess the answer?

Causal decision theory

So, EDT recommends you choose the life of athelete, because this *correlated* with a better outcome.

Causal decision theory

So, EDT recommends you choose the life of athelete, because this *correlated* with a better outcome.

What does causal decision theory (CDT) say instead?

Causal decision theory

So, EDT recommends you choose the life of athelete, because this *correlated* with a better outcome.

What does causal decision theory (CDT) say instead?

Recall that it says that you should choose the option that *causes* the best outcome.

Causal decision theory

So, EDT recommends you choose the life of athelete, because this *correlated* with a better outcome.

What does causal decision theory (CDT) say instead?

Recall that it says that you should choose the option that *causes* the best outcome.

Formally, it says you should maximize:

$$\mathrm{CEU}(O) = \sum_i p(O \mathrel{\square\!\!\rightarrow} S_i) \cdot u(O \cap S_i)$$

Causal decision theory's answer

Since you believe nothing you do can *affect* whether or not you have the heart condition, your credences satisfy the following:

Causal decision theory's answer

Since you believe nothing you do can *affect* whether or not you have the heart condition, your credences satisfy the following:

- $p(\textit{Life as an academic} \mathrel{\square\!\!\rightarrow} \textit{Disease}) = p(\textit{Life as an athelete} \mathrel{\square\!\!\rightarrow} \textit{Disease}) = p(\textit{Disease})$.
- $p(\textit{Life as an academic} \mathrel{\square\!\!\rightarrow} \textit{No disease}) = p(\textit{Life as an athelete} \mathrel{\square\!\!\rightarrow} \textit{No disease}) = p(\textit{No disease})$.

Causal decision theory's answer

Since you believe nothing you do can *affect* whether or not you have the heart condition, your credences satisfy the following:

- $p(\textit{Life as an academic} \mathrel{\square\!\!\rightarrow} \textit{Disease}) = p(\textit{Life as an athelete} \mathrel{\square\!\!\rightarrow} \textit{Disease}) = p(\textit{Disease})$.
- $p(\textit{Life as an academic} \mathrel{\square\!\!\rightarrow} \textit{No disease}) = p(\textit{Life as an athelete} \mathrel{\square\!\!\rightarrow} \textit{No disease}) = p(\textit{No disease})$.

These credences represent the fact that you believe your choice is *causally independent* of the state.

Causal decision theory's answer

Since you believe nothing you do can *affect* whether or not you have the heart condition, your credences satisfy the following:

- $p(\textit{Life as an academic} \boxdot\!\!\rightarrow \textit{Disease}) = p(\textit{Life as an athelete} \boxdot\!\!\rightarrow \textit{Disease}) = p(\textit{Disease})$.
- $p(\textit{Life as an academic} \boxdot\!\!\rightarrow \textit{No disease}) = p(\textit{Life as an athelete} \boxdot\!\!\rightarrow \textit{No disease}) = p(\textit{No disease})$.

These credences represent the fact that you believe your choice is *causally independent* of the state.

Question: With that in mind, then, what does CDT recommend?

(*Pssst!* We don't even need to do any calculations here. Why?)

Causal counterfactuals

A caveat: In order to get the foregoing answer in the decision problem (and Newcomb's problem), it's important that our semantics for counterfactuals track relations of causal dependence and independence.

Causal counterfactuals

A caveat: In order to get the foregoing answer in the decision problem (and Newcomb's problem), it's important that our semantics for counterfactuals track relations of causal dependence and independence.

If we don't have that, then CDT might recommend the wrong answers in certain cases.

Causal counterfactuals

A caveat: In order to get the foregoing answer in the decision problem (and Newcomb's problem), it's important that our semantics for counterfactuals track relations of causal dependence and independence.

If we don't have that, then CDT might recommend the wrong answers in certain cases.

Today, we're going to be looking at the semantics of conditionals in general, and seeing, in particular, whether the standard semantics for counterfactual is up to this task.

Today's lecture

Today's lecture:

i. Semantics for conditionals
ii. Indicative vs subjunctive
iii. Similarity for subjunctives
iv. Which conditionals for deliberation? DeRose's argument

Semantics for conditionals

Conditionals

The goal: a general semantics for *if-then* sentences in natural language.

(1) a. If Frida danced, Maria danced.

b. If you choose one box, you will win a million dollar.

c. If you had not paid the extortionist, your windshield would have been broken.

Two problems about conditionals

A useful distinction, dating back to Stalnaker (1968):

Two problems about conditionals

A useful distinction, dating back to Stalnaker (1968):

- The logical problem of conditionals:

  the problem of specifying a schematic semantics for conditionals, which fixes a logic for conditionals.

Two problems about conditionals

A useful distinction, dating back to Stalnaker (1968):

- The logical problem of conditionals:

  the problem of specifying a schematic semantics for conditionals, which fixes a logic for conditionals.

- The 'pragmatic' problem of conditionals:

  the problem of specifying values for all schematic parameters, in a way that we get truth conditions.

Two problems about conditionals

A useful distinction, dating back to Stalnaker (1968):

- The logical problem of conditionals:

  the problem of specifying a schematic semantics for conditionals, which fixes a logic for conditionals.

- The 'pragmatic' problem of conditionals:

  the problem of specifying values for all schematic parameters, in a way that we get truth conditions.

We'll start from the logical problem.

Our question, then: what kind of logical operators are conditionals?

Basis of modal semantics

All semantics for conditionals descend from the semantics of modal logics.

Basis of modal semantics

All semantics for conditionals descend from the semantics of modal logics.

The basic elements of the model theory of a modal semantics are two:

- A domain $\mathcal{W}$ of possible worlds;
- an accessibility relation $\mathcal{R}$, that relates two worlds w and w' just in case w can 'see' w' for the purposes of evaluating modal claims.

Basis of modal semantics

All semantics for conditionals descend from the semantics of modal logics.

The basic elements of the model theory of a modal semantics are two:

- A domain $\mathcal{W}$ of possible worlds;
- an accessibility relation $\mathcal{R}$, that relates two worlds w and w' just in case w can 'see' w' for the purposes of evaluating modal claims.

Some basic semantic clauses for modal operators:

(2) a. $[\![\Box A]\!]^w = 1$ iff $\forall w' \in \mathcal{W}: \ w'\mathcal{R}w, [\![A]\!]^{w'} = 1$

b. $[\![\Diamond A]\!]^w = 1$ iff $\exists w' \in \mathcal{W}: \ w'\mathcal{R}w, [\![A]\!]^{w'} = 1$

Basis of modal semantics

Rather than using accessibility relations, we will follow custom in the semantics literature and relativize interpretation to a modal base (Kratzer).

A modal base is just a function from a world to a *set* of accessible worlds, which we represent as a subscript 'f' on the modal.

(3) a. $[\![\Box_f A]\!]^w = 1$ iff $\forall w' \in f(w)$: $[\![A]\!]^{w'} = 1$

b. $[\![\Diamond_f A]\!]^w = 1$ iff $\exists w' \in f(w)$: $[\![A]\!]^{w'} = 1$

Basis of modal semantics

Via these clauses, we can give a semantics for modals irrespective of flavor.

(4) a. Frida must be in Chicago. $\rightsquigarrow$ epistemic

b. Frida must go to Chicago. $\rightsquigarrow$ deontic

Basis of modal semantics

Via these clauses, we can give a semantics for modals irrespective of flavor.

(4) a. Frida must be in Chicago. $\rightsquigarrow$ epistemic

b. Frida must go to Chicago. $\rightsquigarrow$ deontic

Both the occurrences of *must* in (4) can be given a unitary analysis:

(5) $[\![\text{must}_f\, A]\!]^{w,f} = 1$ iff $\forall w' \in f(w)$: $[\![A]\!]^{w',f} = 1$

Basis of modal semantics

Via these clauses, we can give a semantics for modals irrespective of flavor.

(4) a. Frida must be in Chicago. $\rightsquigarrow$ epistemic

b. Frida must go to Chicago. $\rightsquigarrow$ deontic

Both the occurrences of *must* in (4) can be given a unitary analysis:

(5) $[\![\text{must}_f\, A]\!]^{w,f} = 1$ iff $\forall w' \in f(w)$: $[\![A]\!]^{w',f} = 1$

- The difference in flavor (epistemic vs deontic) is taken to be due to a difference in contextual input
- Context selects a set of epistemically accessible worlds in one case, and deontically accessible worlds in another.

Basis of modal semantics

Via these clauses, we can give a semantics for modals irrespective of flavor.

(4) a. Frida must be in Chicago. $\rightsquigarrow$ epistemic

b. Frida must go to Chicago. $\rightsquigarrow$ deontic

Both the occurrences of *must* in (4) can be given a unitary analysis:

(5) $[\![\text{must}_f\, A]\!]^{w,f} = 1$ iff $\forall w' \in f(w)$: $[\![A]\!]^{w',f} = 1$

- The difference in flavor (epistemic vs deontic) is taken to be due to a difference in contextual input
- Context selects a set of epistemically accessible worlds in one case, and deontically accessible worlds in another.

Via a similar maneuver, we will use the same logical structure to analyze both indicative and subjunctive conditionals.

Strict conditional semantics

How to give a semantics for conditionals in this context?

A first stab: strict conditional semantics.

Strict conditional semantics

How to give a semantics for conditionals in this context?

A first stab: strict conditional semantics.

A strict conditional is a necessitated material conditional.

$$\Box_f(A \supset C)$$

A first-pass analysis treats natural language conditionals as strict conditionals:

(6) $[\![A \rightarrow C]\!]^{w,f} = 1$ iff $\forall w' \in f(w)$ s.t. $[\![A]\!]^{w',f} = 1$, $[\![C]\!]^{w',f} = 1$

Strict conditional semantics

How to give a semantics for conditionals in this context?

A first stab: strict conditional semantics.

A strict conditional is a necessitated material conditional.

$$\Box_f(A \supset C)$$

A first-pass analysis treats natural language conditionals as strict conditionals:

(6) $[\![A \rightarrow C]\!]^{w,f} = 1$ iff $\forall w' \in f(w)$ s.t. $[\![A]\!]^{w',f} = 1$, $[\![C]\!]^{w',f} = 1$

Assuming that bare indicative conditionals default to epistemic flavor, and that epistemic flavor tracks knowledge of some relevant agents, we get:

(7) a. If Frida danced, Maria danced.

b. $[\![(24a)]\!]^{w',f} = 1$ iff all worlds w' compatible with what is known in w such that Frida danced in w' are such that Maria danced in w'

Strict conditional semantics

But: strict conditional semantics is empirically incorrect.

Strict conditional semantics

But: strict conditional semantics is empirically incorrect.

This is (still) a controversial claim, but here are the two basic arguments for it:

Strict conditional semantics

But: strict conditional semantics is empirically incorrect.

This is (still) a controversial claim, but here are the two basic arguments for it:

(i) nonmonotonic inference patterns;

(ii) the link between conditionals and probability.

Nonmonotonicity

Argument #1: nonmonotonicity. Strict conditional semantics entails:

Antecedent strengthening (AS) $A \to C \vDash A^+ \to C$ (where $A^+ \vDash A$)

Nonmonotonicity

Argument #1: nonmonotonicity. Strict conditional semantics entails:

Antecedent strengthening (AS) $A \to C \vDash A^+ \to C$ (where $A^+ \vDash A$)

But AS fails for at least some conditionals, in particular subjunctives.

Nonmonotonicity

Argument #1: nonmonotonicity. Strict conditional semantics entails:

Antecedent strengthening (AS) $A \rightarrow C \vDash A^{+} \rightarrow C$ (where $A^{+} \vDash A$)

But AS fails for at least some conditionals, in particular subjunctives.

(8) a. If Alice had struck this match, the match would have lit.

b. If Alice had struck this match and the match had been wet, the match would not have lit.

Nonmonotonicity

Argument #1: nonmonotonicity. Strict conditional semantics entails:

Antecedent strengthening (AS) $A \rightarrow C \vDash A^{+} \rightarrow C$ (where $A^{+} \vDash A$)

But AS fails for at least some conditionals, in particular subjunctives.

(8) a. If Alice had struck this match, the match would have lit.

b. If Alice had struck this match and the match had been wet, the match would not have lit.

(8) is a Sobel sequence, i.e. a discourse of the apparent form

$$A \rightarrow C,\ A^{+} \rightarrow \neg C$$

Nonmonotonicity

Argument #1: nonmonotonicity. Strict conditional semantics entails:

Antecedent strengthening (AS) $A \rightarrow C \vDash A^{+} \rightarrow C$ (where $A^{+} \vDash A$)

But AS fails for at least some conditionals, in particular subjunctives.

(8) a. If Alice had struck this match, the match would have lit.

b. If Alice had struck this match and the match had been wet, the match would not have lit.

(8) is a Sobel sequence, i.e. a discourse of the apparent form

$$A \rightarrow C,\ A^{+} \rightarrow \neg C$$

- Sobel sequences are predicted to be inconsistent by a strict semantics.
- Yet discourses like (8) are perfectly consistent.
- So strict conditional semantics makes incorrect predictions about conditional logic.

Probability

Argument #2: probability.

Plausibly, the probabilities of conditionals conform to some bridge principles.

Stalnaker's Thesis
For all Pr that model rational credence, and for all A, B:

$$Pr(A \rightarrow_{\text{ind}} B) = Pr(B \mid A) \quad (\text{if } Pr(A) > 0)$$

Probability

Argument #2: probability.

Plausibly, the probabilities of conditionals conform to some bridge principles.

> **Stalnaker's Thesis**
> For all Pr that model rational credence, and for all A, B:
>
> $$Pr(A \rightarrow_{\text{ind}} B) = Pr(B \mid A) \quad (\text{if } Pr(A) > 0)$$

- Stalnaker's Thesis is hard to vindicate as stated.
- But versions of it can be vindicated (Van Fraassen 1976, Kaufmann 2019, Bacon 2015, Goldstein and Santorio 2021, Khoo 2022, a.o.)

Probability

But: conditional probabilities are nonmonotonic!

But: conditional probabilities are nonmonotonic!

- We can have both $Pr(C \mid A) < Pr(C \mid (A \wedge B))$ and $Pr(C \mid A) > Pr(C \mid (A \wedge B))$, depending on our choice of A, B, and C.

But: conditional probabilities are nonmonotonic!

- We can have both $Pr(C \mid A) < Pr(C \mid (A \wedge B))$ and $Pr(C \mid A) > Pr(C \mid (A \wedge B))$, depending on our choice of A, B, and C.
- Example: suppose that we are throwing a fair die. We have:

$$\begin{array}{lcl} Pr(5 \vee 6 \mid 3 \vee 4 \vee 5 \vee 6) & < & Pr(5 \vee 6 \mid 4 \vee 5 \vee 6) \\ Pr(5 \vee 6 \mid 3 \vee 4 \vee 5 \vee 6) & > & Pr(5 \vee 6 \mid 3 \vee 4 \vee 5) \end{array}$$

But: conditional probabilities are nonmonotonic!

- We can have both $Pr(C \mid A) < Pr(C \mid (A \wedge B))$ and $Pr(C \mid A) > Pr(C \mid (A \wedge B))$, depending on our choice of A, B, and C.
- Example: suppose that we are throwing a fair die. We have:

$$\begin{array}{lcl} Pr(5 \vee 6 \mid 3 \vee 4 \vee 5 \vee 6) & < & Pr(5 \vee 6 \mid 4 \vee 5 \vee 6) \\ Pr(5 \vee 6 \mid 3 \vee 4 \vee 5 \vee 6) & > & Pr(5 \vee 6 \mid 3 \vee 4 \vee 5) \end{array}$$

So: if some version of Stalnaker's Thesis holds, and entailment preserves probability, conditionals cannot be monotonic.

Selection semantics

What is the alternative to strict semantics?

- A number of ways of framing what is essentially the same semantic idea (Stalnaker 1968, Lewis 1973, Kratzer 1977, 1981, 1986, 2012 a.o.)
  i. ordering semantics;
  ii. premise semantics;
  iii. selection semantics;
  ...
- Here we use selection semantics, as defined and developed by Stalnaker.

As usual, the Ramsey test

Stalnaker's intuitive motivation behind selection semantics: the Ramsey test.

As usual, the Ramsey test

Stalnaker's intuitive motivation behind selection semantics: the Ramsey test.

Stalnaker asks: how do we decide whether to believe a conditional?

As usual, the Ramsey test

Stalnaker's intuitive motivation behind selection semantics: the Ramsey test.

Stalnaker asks: how do we decide whether to believe a conditional?

An answer suggested by Ramsey in a famous footnote:

> *First, add the antecedent (hypothetically) to your stock of beliefs; second, [if the conditional is a counterfactual] 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.*

As usual, the Ramsey test

Stalnaker's intuitive motivation behind selection semantics: the Ramsey test.

Stalnaker asks: how do we decide whether to believe a conditional?

An answer suggested by Ramsey in a famous footnote:

> *First, add the antecedent (hypothetically) to your stock of beliefs; second, [if the conditional is a counterfactual] 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.*

- These are 'belief conditions' for conditionals.
- We can import this idea in a statement of truth conditions, by letting the semantics select a 'minimally different' world.

Selection semantics

An informal statement of truth conditions on selection semantics:

> $A \rightarrow C$ is true at w iff C is true at the most similar A-world to w

Formally: we build into the semantics a mechanism to select the most similar A-world to the world of evaluation.

Selection semantics

An informal statement of truth conditions on selection semantics:

> $A \rightarrow C$ is true at w iff C is true at the most similar A-world to w

Formally: we build into the semantics a mechanism to select the most similar A-world to the world of evaluation.

A selection function is a function $s : W \times \mathcal{P}(W) \mapsto \mathcal{P}(W)$ that maps a pair of a world and a set of worlds to a set of worlds.

- In particular: we use selection functions that select a singleton (or, equivalently, just a world).
- A Stalnaker selection function is a function that (i) selects a single world and (ii) is subject to certain conditions.

Selection semantics

A Stalnaker model includes a set of possible worlds $\mathcal{W}^-$ as well as an *absurd world* λ, which makes true every sentence. Let $\mathcal{W} = \mathcal{W}^- \cup \{\lambda\}$ $\mathcal{W}$.

A function $s : \mathcal{W} \times \mathcal{P}(\mathcal{W}) \mapsto \mathcal{W}$ is a Stalnaker selection function iff

i. If $[\![A]\!]$ is non-empty, $s(w, A) \in [\![A]\!]$

(Inclusion: the selected world makes the antecedent true, if at all possible.)

Selection semantics

A Stalnaker model includes a set of possible worlds $\mathcal{W}^-$ as well as an *absurd world* λ, which makes true every sentence. Let $\mathcal{W} = \mathcal{W}^- \cup \{\lambda\}$ $\mathcal{W}$.

A function $s : \mathcal{W} \times \mathcal{P}(\mathcal{W}) \mapsto \mathcal{W}$ is a Stalnaker selection function iff

i. If $[\![A]\!]$ is non-empty, $s(w, A) \in [\![A]\!]$

(Inclusion: the selected world makes the antecedent true, if at all possible.)

ii. If $s(w, A) = \lambda$, then $[\![A]\!] = \varnothing$

(Absurdity-as-last-resort: the impossible world is selected only if no possible world can be selected.)

Selection semantics

A Stalnaker model includes a set of possible worlds $\mathcal{W}^-$ as well as an *absurd world* λ, which makes true every sentence. Let $\mathcal{W} = \mathcal{W}^- \cup \{\lambda\}$ $\mathcal{W}$.

A function $s : \mathcal{W} \times \mathcal{P}(\mathcal{W}) \mapsto \mathcal{W}$ is a Stalnaker selection function iff

i. If $[\![A]\!]$ is non-empty, $s(w, A) \in [\![A]\!]$

(Inclusion: the selected world makes the antecedent true, if at all possible.)

ii. If $s(w, A) = \lambda$, then $[\![A]\!] = \varnothing$

(Absurdity-as-last-resort: the impossible world is selected only if no possible world can be selected.)

iii. If $w \in [\![A]\!]$, then $s(w, A) = w$

(Centering: if the world of eval. makes the antecedent true, it is selected.)

Selection semantics

A Stalnaker model includes a set of possible worlds $\mathcal{W}^-$ as well as an *absurd world* λ, which makes true every sentence. Let $\mathcal{W} = \mathcal{W}^- \cup \{\lambda\}$ $\mathcal{W}$.

A function $s : \mathcal{W} \times \mathcal{P}(\mathcal{W}) \mapsto \mathcal{W}$ is a Stalnaker selection function iff

i. If $[\![A]\!]$ is non-empty, $s(w, A) \in [\![A]\!]$

(Inclusion: the selected world makes the antecedent true, if at all possible.)

ii. If $s(w, A) = \lambda$, then $[\![A]\!] = \varnothing$

(Absurdity-as-last-resort: the impossible world is selected only if no possible world can be selected.)

iii. If $w \in [\![A]\!]$, then $s(w, A) = w$

(Centering: if the world of eval. makes the antecedent true, it is selected.)

iv. For all A, A': if $s(w, A) \in [\![A']\!]$ and $s(w, A') \in [\![A]\!]$, then:
$s(w, A) \in [\![A']\!] = s(w, A') \in [\![A]\!]$

(Consistency of selection: the selection must be consistent for all choice of antecedents.)

Selection semantics

Below is a semantics of conditional based on Stalnaker selection functions.

We relativize interpretation to three parameters: a world w, a modal base f, and a Stalnaker selection function s.

(9) $[\![A \rightarrow C]\!]^{w,f,s} = \text{true iff } [\![C]\!]^{s(w,f(w)\cap[\![A]\!]),f,s} = \text{true}$

In words: $A \rightarrow C$ is true relative to w, f, and s iff the selected A-world in $f(w)$ is a C-world.

Selection semantics

- Selection semantics can be recast in terms of comparative similarity and orderings.
- In particular: a single-world selection semantics is equivalent to a semantics where worlds form a linear ordering (no ties and no incomparabilities).

Indicatives vs subjunctives

Indicatives vs subjunctives

So we have a semantics for conditionals:

(9) $[\![A \rightarrow C]\!]^{w,f,s} = \text{true iff } [\![C]\!]^{s(w,f(w)\cap[\![A]\!]),f,s} = \text{true}$

Indicatives vs subjunctives

So we have a semantics for conditionals:

(9) $[\![A \rightarrow C]\!]^{w,f,s} = \text{true iff } [\![C]\!]^{s(w,f(w)\cap[\![A]\!]),f,s} = \text{true}$

- This gives us the general form of the truth conditions for conditionals (and fixes their logic).

Indicatives vs subjunctives

So we have a semantics for conditionals:

(9) $[\![A \rightarrow C]\!]^{w,f,s} = \text{true iff } [\![C]\!]^{s(w,f(w)\cap[\![A]\!]),f,s} = \text{true}$

- This gives us the general form of the truth conditions for conditionals (and fixes their logic).
- So: it solves Stalnaker's logical problem.

Indicatives vs subjunctives

So we have a semantics for conditionals:

(9) $[\![A \rightarrow C]\!]^{w,f,s} = \text{true iff } [\![C]\!]^{s(w,f(w)\cap[\![A]\!]),f,s} = \text{true}$

- This gives us the general form of the truth conditions for conditionals (and fixes their logic).
- So: it solves Stalnaker's logical problem.
- But: we still need to say what world is selected each time, i.e. what world counts as 'the most similar antecedent world'.

Indicatives vs subjunctives

So we have a semantics for conditionals:

(9) $[\![A \rightarrow C]\!]^{w,f,s} = \text{true iff } [\![C]\!]^{s(w,f(w)\cap[\![A]\!]),f,s} = \text{true}$

- This gives us the general form of the truth conditions for conditionals (and fixes their logic).
- So: it solves Stalnaker's logical problem.
- But: we still need to say what world is selected each time, i.e. what world counts as 'the most similar antecedent world'.
- This is what Stalnaker called 'pragmatic problem' (and that today would be called 'the problem of specifying contextual parameters', or st like that).

Indicatives vs subjunctives

A general point:

- There is no 'unique' notion of similarity;
- conditionals of different flavors will exploit different selection functions.

In particular, there is agreement on the following:

- The similarity relation exploited by indicatives is epistemic—it tracks facts determined by knowledge and evidence.
- The similarity relation exploited by subjunctives is metaphysical—it tracks facts determined by history and laws of nature.

Indicatives vs subjunctives

Why think we need two different flavors?

Indicatives vs subjunctives

Why think we need two different flavors?

We already saw yesterday a key minimal pair:

(10) a. If Shakespeare didn't write *Hamlet*, someone else did.
⇝ true/high probability

b. If Shakespeare hadn't written *Hamlet*, someone else would have.
⇝ false/low probability

Indicatives vs subjunctives

Why think we need two different flavors?

We already saw yesterday a key minimal pair:

(10) a. If Shakespeare didn't write *Hamlet*, someone else did.
$\rightsquigarrow$ true/high probability

b. If Shakespeare hadn't written *Hamlet*, someone else would have.
$\rightsquigarrow$ false/low probability

So we know: there needs to be (at least) two types of similarity orderings.

- Selection semantics gives the skeleton for the semantics in both cases.
- But conditionals of different flavors exploit different selection functions.

Indicatives vs subjunctives

Before we proceed: what do we actually mean by ‘indicative’ and ‘subjunctive’?

Indicatives vs subjunctives

Before we proceed: what do we actually mean by 'indicative' and 'subjunctive'?

A rough and ready characterization:

- Indicative conditionals are conditionals with no overt modal.

(10a) If Shakespeare didn't write *Hamlet*, someone else did.

Indicatives vs subjunctives

Before we proceed: what do we actually mean by 'indicative' and 'subjunctive'?

A rough and ready characterization:

- Indicative conditionals are conditionals with no overt modal.

(10a) If Shakespeare didn't write *Hamlet*, someone else did.

- Subjunctive conditionals are conditionals with certain morphological markings, e.g. past tense (interestingly, a 'fake' one) in the antecedent, and *would* in the consequent. (What has been called 'X-marking'.)

(10a) If Shakespeare hadn't written *Hamlet*, someone else would have.

Indicatives vs subjunctives

But: this rough and ready characterization runs into trouble.

(11) If Calum doesn't write the next great book on decision theory, someone else will.

Indicatives vs subjunctives

But: this rough and ready characterization runs into trouble.

(11) If Calum doesn't write the next great book on decision theory, someone else will.

- On standard theories, *will* is a modal, and moreover it is morphologically and semantically related to *would*.
  (Klecha 2015, Cariani and Santorio 2018 a.o.).

Indicatives vs subjunctives

But: this rough and ready characterization runs into trouble.

(11) If Calum doesn't write the next great book on decision theory, someone else will.

- On standard theories, *will* is a modal, and moreover it is morphologically and semantically related to *would*.
  (Klecha 2015, Cariani and Santorio 2018 a.o.).
- Moreover: we already noticed that *will*- and *would*-conditionals about the future can seem very similar in meaning.

(12) a. If I take my umbrella, I will stay dry.

b. If I were to take my umbrella, I would stay dry.

(For relevant work: Ippolito 2013, Mendes 2025 a.o.)

Indicatives vs subjunctives

But: this rough and ready characterization runs into trouble.

(11) If Calum doesn't write the next great book on decision theory, someone else will.

- On standard theories, *will* is a modal, and moreover it is morphologically and semantically related to *would*.
  (Klecha 2015, Cariani and Santorio 2018 a.o.).
- Moreover: we already noticed that *will*- and *would*-conditionals about the future can seem very similar in meaning.

(12) a. If I take my umbrella, I will stay dry.

b. If I were to take my umbrella, I would stay dry.

(For relevant work: Ippolito 2013, Mendes 2025 a.o.)

- So how should we think of conditionals like (11)? This issue will come back when we consider DeRose on conditionals of deliberation.

Similarity for subjunctives

The nature of similarity

So how is similarity determined, for both indicatives and subjunctives?

The nature of similarity

So how is similarity determined, for both indicatives and subjunctives?

- For indicatives: dramatically undertheorized. (Though see Krater 1981.)
- For subjunctives: plenty of theories, and this will really matter for the rest of our course.

The nature of similarity

So how is similarity determined, for both indicatives and subjunctives?

- For indicatives: dramatically undertheorized. (Though see Krater 1981.)
- For subjunctives: plenty of theories, and this will really matter for the rest of our course.

Why similarity for subjunctives matters:

- Causal decision theory crucially relies on truth-conditions of counterfactuals.
- Some problems for CDT will hinge on theories of similarity for counterfactuals.

Similarity for subjunctives

A preliminary point: backtrackers.

Suppose you stand on the top of a building with no safety net below.

Me: *We're high enough that, if you were to jump, you would die.*
You: *I don't have a death wish! If I were to jump, there would be a safety net, and I wouldn't die.* *(Gallow 2022)*

Similarity for subjunctives

A preliminary point: backtrackers.

Suppose you stand on the top of a building with no safety net below.

Me: We're high enough that, if you were to jump, you would die.
You: I don't have a death wish! If I were to jump, there would be a safety net, and I wouldn't die. *(Gallow 2022)*

- We have two incompatible counterfactuals:

$$\textit{jump} \boxdot\!\!\rightarrow \textit{die}, \quad \textit{jump} \boxdot\!\!\rightarrow \neg \textit{die}$$

Similarity for subjunctives

A preliminary point: backtrackers.

Suppose you stand on the top of a building with no safety net below.

Me: We're high enough that, if you were to jump, you would die.
You: I don't have a death wish! If I were to jump, there would be a safety net, and I wouldn't die. *(Gallow 2022)*

- We have two incompatible counterfactuals:

$$\textit{jump} \boxrightarrow \textit{die}, \quad \textit{jump} \boxrightarrow \neg \textit{die}$$

- Both seem to be uttered truly!

Similarity for subjunctives

A preliminary point: backtrackers.

Suppose you stand on the top of a building with no safety net below.

Me: We're high enough that, if you were to jump, you would die.
You: I don't have a death wish! If I were to jump, there would be a safety net, and I wouldn't die. *(Gallow 2022)*

- We have two incompatible counterfactuals:

$$\textit{jump} \boxrightarrow \textit{die}, \quad \textit{jump} \boxrightarrow \neg \textit{die}$$

- Both seem to be uttered truly!
- So: we need to invoke context dependence. The two counterfactuals are evaluated wrt different orderings.

Similarity for subjunctives

A preliminary point: backtrackers.

Suppose you stand on the top of a building with no safety net below.

Me: We're high enough that, if you were to jump, you would die.
You: I don't have a death wish! If I were to jump, there would be a safety net, and I wouldn't die. *(Gallow 2022)*

- We have two incompatible counterfactuals:

$$\textit{jump} \boxrightarrow \textit{die}, \quad \textit{jump} \boxrightarrow \neg \textit{die}$$

- Both seem to be uttered truly!
- So: we need to invoke context dependence. The two counterfactuals are evaluated wrt different orderings.
- In particular, the second one is a paradigm 'backtracker': it involves an inference to what conditions would have had to obtain to bring about a certain outcome.

Similarity for subjunctives

A preliminary point: backtrackers.

Suppose you stand on the top of a building with no safety net below.

Me: We're high enough that, if you were to jump, you would die.
You: I don't have a death wish! If I were to jump, there would be a safety net, and I wouldn't die. *(Gallow 2022)*

- We have two incompatible counterfactuals:

$$jump \boxrightarrow die, \quad jump \boxrightarrow \neg die$$

- Both seem to be uttered truly!
- So: we need to invoke context dependence. The two counterfactuals are evaluated wrt different orderings.
- In particular, the second one is a paradigm 'backtracker': it involves an inference to what conditions would have had to obtain to bring about a certain outcome.
- This is not the kind of causal conditional we need in decision theory. So (following pretty much everyone else!) we set backtrackers aside.

Similarity for subjunctives

Let's focus then on characterizing similarity for ordinary subjunctives, like (13).

(13) If I had struck this match, it would have lit.

Similarity for subjunctives

Let's focus then on characterizing similarity for ordinary subjunctives, like (13).

(13) If I had struck this match, it would have lit.

Question: won't an intuitive notion of closeness do?

Similarity for subjunctives

Let's focus then on characterizing similarity for ordinary subjunctives, like (13).

(13) If I had struck this match, it would have lit.

Question: won't an intuitive notion of closeness do?

Answer: no.

The nature of similarity

The counterfactual "If Nixon had pressed the button there would have been a nuclear holocaust" is true or can be imagined to be so. Now suppose that there never will be a nuclear holocaust. Then that counterfactual is, [on an analysis that employs a naïve similarity relation], very likely false. For given any world in which antecedent and consequent are both true it will be easy to imagine a closer world in which the antecedent is true and the consequent false. For we need only imagine a change that prevents the holocaust but that does not require such a great divergence from reality. (Fine 1975)

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

Intuitively true

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

Intuitively true, and (if similarity is naïve similarity) predicted to be

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

Intuitively true, and (if similarity is naïve similarity) predicted to be false.

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

Intuitively true, and (if similarity is naïve similarity) predicted to be false.

- Worlds where no nuclear holocaust happens are overall more similar to the actual world where worlds where there is indeed a nuclear holocaust.

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

Intuitively true, and (if similarity is naïve similarity) predicted to be false.

- Worlds where no nuclear holocaust happens are overall more similar to the actual world where worlds where there is indeed a nuclear holocaust.
- Hence the most similar worlds where Nixon pushes the nuclear button are worlds where some malfunction happens and no nuclear bombs are launched, making the counterfactual false.

The nature of similarity

(14) If Nixon had pushed the button [that launches nuclear missiles], there would have been a nuclear holocaust.

Intuitively: true? false?

Intuitively true, and (if similarity is naïve similarity) predicted to be false.

- Worlds where no nuclear holocaust happens are overall more similar to the actual world where worlds where there is indeed a nuclear holocaust.
- Hence the most similar worlds where Nixon pushes the nuclear button are worlds where some malfunction happens and no nuclear bombs are launched, making the counterfactual false.

Moral: we cannot just rely on a naïve notion of similarity.

The nature of similarity

In "Counterfactual Dependence and Time's Arrow", Lewis notices a fundamental asymmetry:

> *The way the future is depends counterfactually on the way the present is.* *If the present were different, the future would be different; and there are counterfactual conditionals, many of them as unquestionably true as counterfactuals ever get, that tell us a good deal about how the future would be different if the present were different in various ways. Likewise the present depends counterfactually on the past, and in general the way things are later depends on the way things were earlier.*
>
> *Not so in reverse.* *Seldom, if ever, can we find a clearly true counterfactual about how the past would be different if the present were somehow different. Such a counterfactual, unless clearly false, normally is not clear one way or the other. It is at best doubtful whether the past depends counterfactually on the present, whether the present depends on the future, and in general whether the way things are earlier depends on the way things will be later.*

- the present counterfactually depends on the past, but not *vice versa*;
- an account of similarity should build on this.

The nature of similarity

So: similarity exploits a past-future asymmetry.

The nature of similarity

So: similarity exploits a past-future asymmetry.

Two main criteria for similarity:

- match in history up to the reference time of the antecedent;
- match with actual laws of nature.

The nature of similarity

So: similarity exploits a past-future asymmetry.

Two main criteria for similarity:

- match in history up to the reference time of the antecedent;
- match with actual laws of nature.

Unfortunately: given minimal assumptions, we cannot assume that similarity requires both perfect match in history and perfect match with actual laws.

The nature of similarity

- Let L_w be the proposition that encodes all the laws of nature in w;
- Let $H_{w,t}$ be the proposition that encodes all particular historical events in history in w, up to time t.
- Let H_w be the proposition that encodes all of the history of w.

The nature of similarity

- Let L_w be the proposition that encodes all the laws of nature in w;
- Let $H_{w,t}$ be the proposition that encodes all particular historical events in history in w, up to time t.
- Let H_w be the proposition that encodes all of the history of w.

Let us consider the hypothesis of determinism, defined as follows:

> Determinism. L_w and $H_{w,t}$ jointly entail H_w (for any t).

The nature of similarity

- Let L_w be the proposition that encodes all the laws of nature in w;
- Let $H_{w,t}$ be the proposition that encodes all particular historical events in history in w, up to time t.
- Let H_w be the proposition that encodes all of the history of w.

Let us consider the hypothesis of determinism, defined as follows:

> Determinism. L_w and $H_{w,t}$ jointly entail H_w (for any t).

- If Determinism is true, then worlds that we reach via counterfactual supposition cannot verify both actual history and actual laws.
  (At least, unless we want to make all counterfactuals trivially true.)

The nature of similarity

- Let L_w be the proposition that encodes all the laws of nature in w;
- Let $H_{w,t}$ be the proposition that encodes all particular historical events in history in w, up to time t.
- Let H_w be the proposition that encodes all of the history of w.

Let us consider the hypothesis of determinism, defined as follows:

> Determinism. L_w and $H_{w,t}$ jointly entail H_w (for any t).

- If Determinism is true, then worlds that we reach via counterfactual supposition cannot verify both actual history and actual laws.
  (At least, unless we want to make all counterfactuals trivially true.)
- Perhaps determinism is false, but we want our semantics for counterfactuals to be general enough to apply to both the deterministic and the indeterministic case.

The nature of similarity

- Let L_w be the proposition that encodes all the laws of nature in w;
- Let $H_{w,t}$ be the proposition that encodes all particular historical events in history in w, up to time t.
- Let H_w be the proposition that encodes all of the history of w.

Let us consider the hypothesis of determinism, defined as follows:

> Determinism. L_w and $H_{w,t}$ jointly entail H_w (for any t).

- If Determinism is true, then worlds that we reach via counterfactual supposition cannot verify both actual history and actual laws.
  (At least, unless we want to make all counterfactuals trivially true.)
- Perhaps determinism is false, but we want our semantics for counterfactuals to be general enough to apply to both the deterministic and the indeterministic case.

Given this: we have two options.

The nature of similarity

Option #1: miracles. This is the option adopted by Lewis.

The nature of similarity

Option #1: miracles. This is the option adopted by Lewis.

Lewis's criteria for similarity:

(1) It is of the first importance to avoid big, widespread, diverse violations of law.

(2) It is of the second importance to maximize the spatiotemporal region throughout which perfect match of particular fact prevails.

(3) It is of the third importance to avoid even small, localized, simple violations of law.

(4) It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.

The nature of similarity

Option #1: miracles. This is the option adopted by Lewis.

Lewis's criteria for similarity:

(1) It is of the first importance to avoid big, widespread, diverse violations of law.

(2) It is of the second importance to maximize the spatiotemporal region throughout which perfect match of particular fact prevails.

(3) It is of the third importance to avoid even small, localized, simple violations of law.

(4) It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.

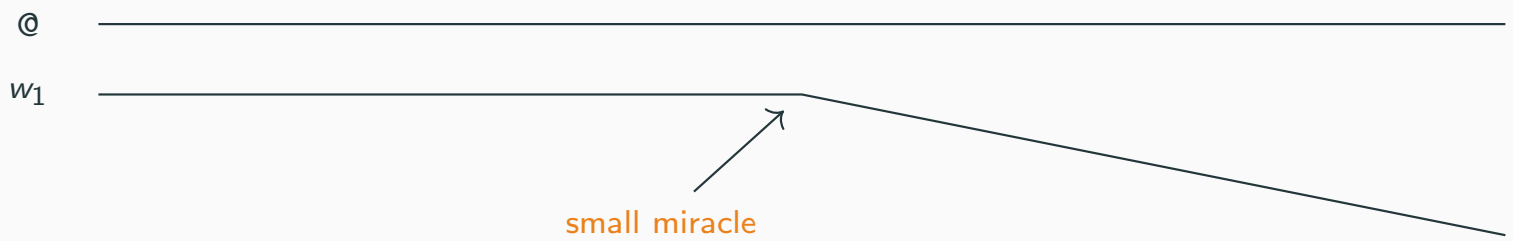
So Lewis chooses:

(i) perfect match in (past) history;

(ii) (possibly) imperfect match in laws.

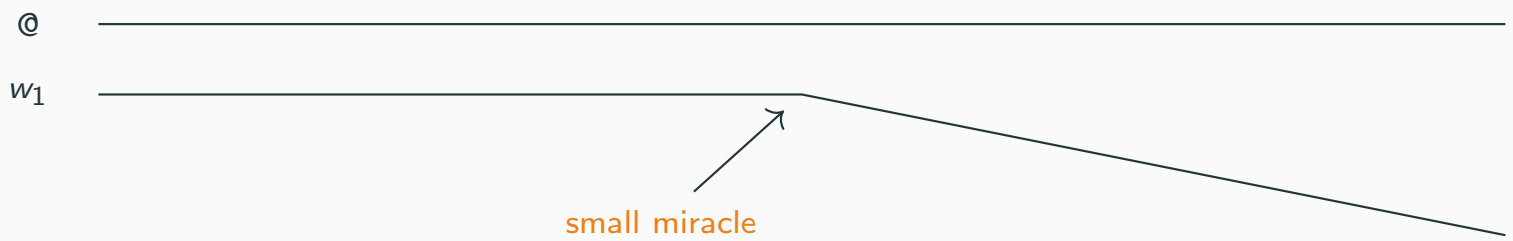
The nature of similarity

Most similar A-worlds to w: worlds with an identical course of history up to shortly before t_A, and where a 'small miracle' brings about A.



The nature of similarity

Most similar A-worlds to w: worlds with an identical course of history up to shortly before t_A, and where a 'small miracle' brings about A.

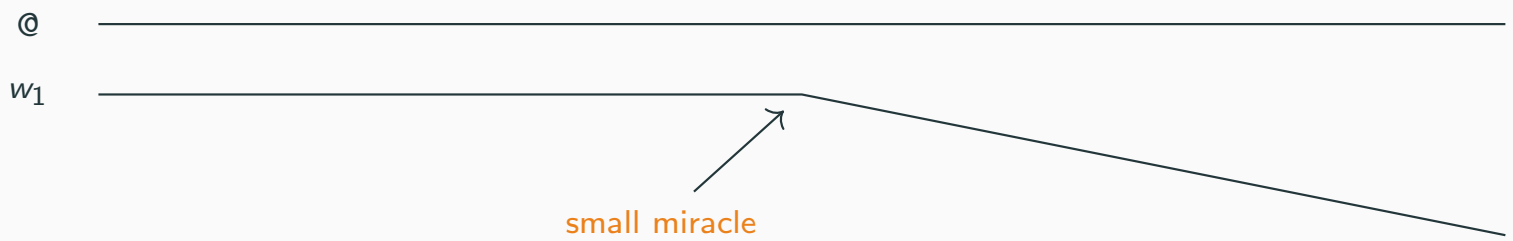


Worlds that are not among the closest:

- Worlds where laws perfectly match actual laws, but have a different history.

The nature of similarity

Most similar A-worlds to w: worlds with an identical course of history up to shortly before t_A, and where a 'small miracle' brings about A.

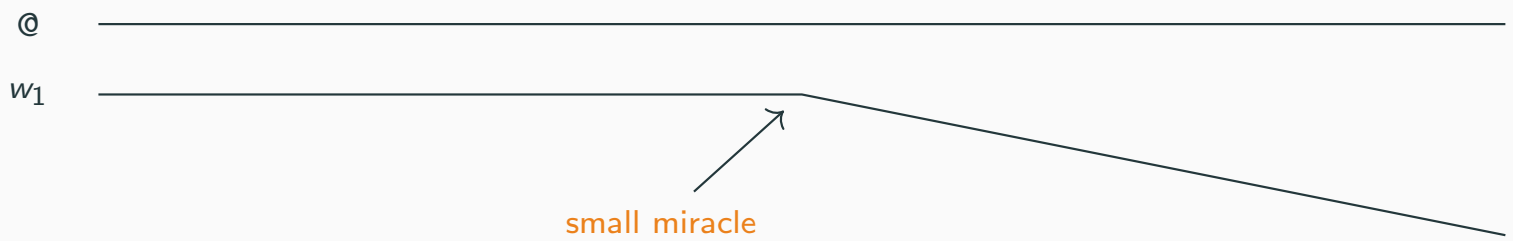


Worlds that are not among the closest:

- Worlds where laws perfectly match actual laws, but have a different history.
- Worlds with abrupt discrepancies in history (bumps).

The nature of similarity

Most similar A-worlds to w: worlds with an identical course of history up to shortly before t_A, and where a 'small miracle' brings about A.



Worlds that are not among the closest:

- Worlds where laws perfectly match actual laws, but have a different history.
- Worlds with abrupt discrepancies in history (bumps).
- Worlds that involve major divergences from actual laws (big miracles).

The nature of similarity

In the Nixon case: closest worlds are worlds where:

- History matches actual history until shortly before antecedent time;
- A ‘small miracle’ leads Nixon to push the button.
- History takes a very different course after that.

The nature of similarity

In the Nixon case: closest worlds are worlds where:

- History matches actual history until shortly before antecedent time;
- A 'small miracle' leads Nixon to push the button.
- History takes a very different course after that.

Hence: it's predicted to be true that, if Nixon had pressed the button, there would have been a nuclear holocaust.

Option #2: no-miracles. This option is defended by a number of philosophers, including Dorr (2016) a.o.

The nature of similarity

Option #2: no-miracles. This option is defended by a number of philosophers, including Dorr (2016) a.o.

On this option, we require:

- Laws of nature hold throughout closest worlds.
- History is so-very-slightly different, starting from the beginning of time.

The nature of similarity

Option #2: no-miracles. This option is defended by a number of philosophers, including Dorr (2016) a.o.

On this option, we require:

- Laws of nature hold throughout closest worlds.
- History is so-very-slightly different, starting from the beginning of time.

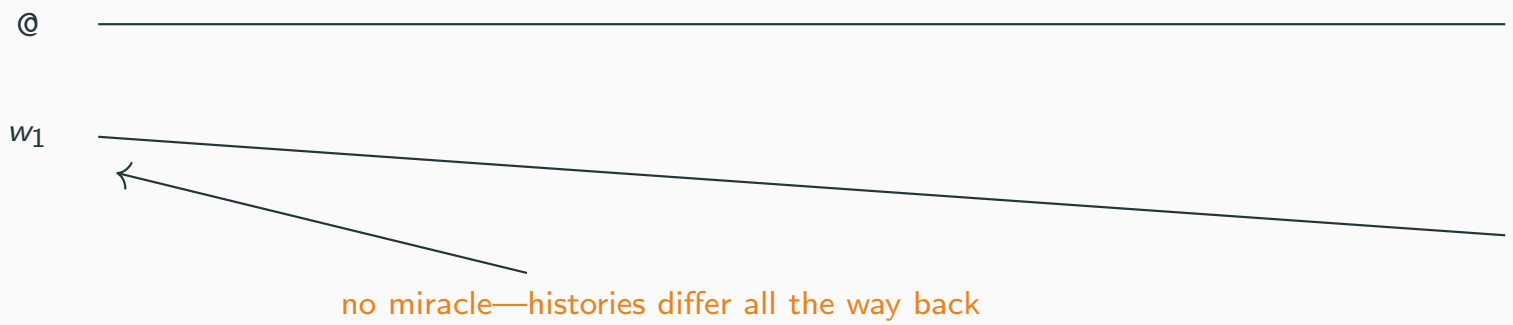
So we choose:

(i) perfect match in laws;

(ii) (possibly) imperfect match in history.

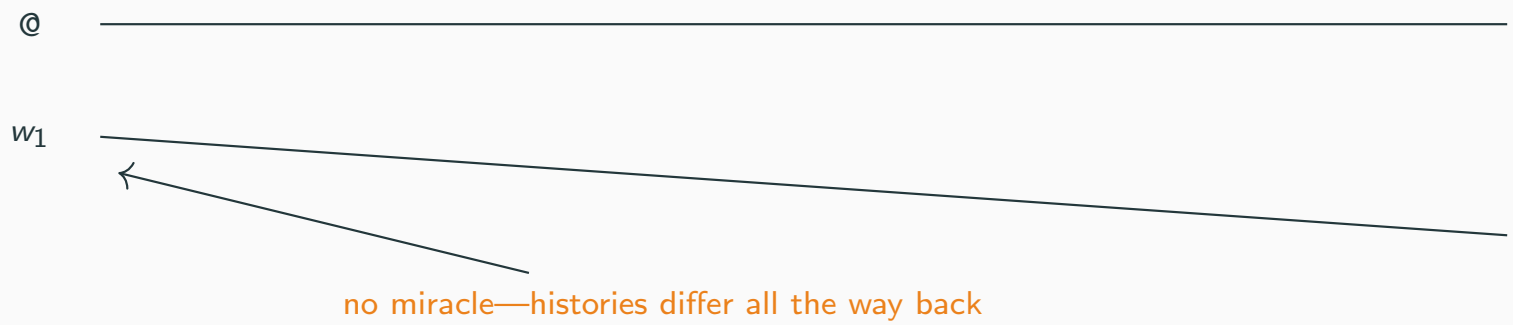
The nature of similarity

Most similar A-worlds to w: worlds with identical laws to w, and with a history that is different from the very beginning (though, possibly, in imperceptible ways).



The nature of similarity

Most similar A-worlds to w: worlds with identical laws to w, and with a history that is different from the very beginning (though, possibly, in imperceptible ways).

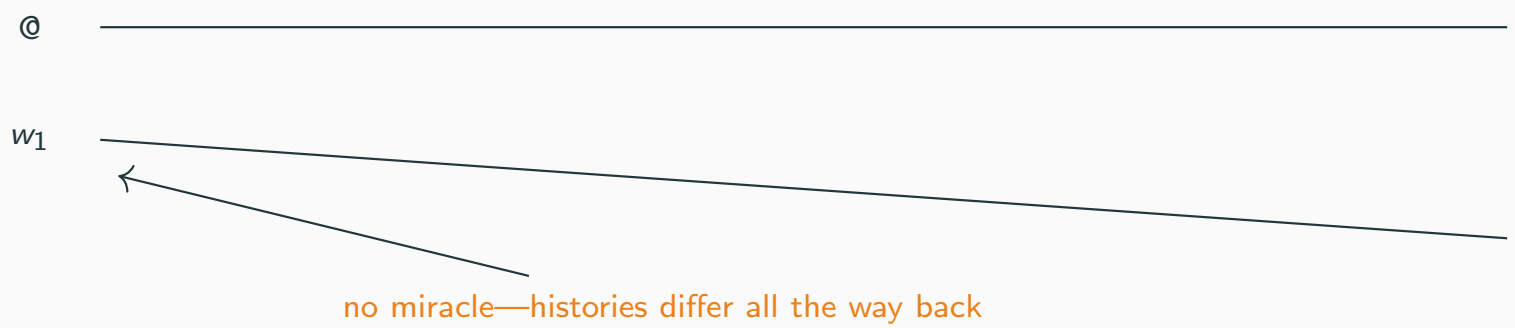


Worlds that are not among the closest:

- Worlds with miracles, big and small.

The nature of similarity

Most similar A-worlds to w: worlds with identical laws to w, and with a history that is different from the very beginning (though, possibly, in imperceptible ways).



Worlds that are not among the closest:

- Worlds with miracles, big and small.
- Worlds with abrupt discrepancies in history (bumps).

The nature of similarity

In the Nixon case: closest worlds are worlds where:

- History was always different, from the beginning of time.
- This divergence does not generate any divergence in macroscopic facts, up to the point at which Nixon to push the button.
- History takes a very different course after that.

The nature of similarity

In the Nixon case: closest worlds are worlds where:

- History was always different, from the beginning of time.
- This divergence does not generate any divergence in macroscopic facts, up to the point at which Nixon to push the button.
- History takes a very different course after that.

Again: it's predicted to be true that, if Nixon had pressed the button, there would have been a nuclear holocaust.

What conditionals for deliberation?

DeRose on conditionals of deliberation

So now that we have truth conditions, we can ask: what conditionals are employed in deliberation—indicative, subjunctive, or both?

DeRose on conditionals of deliberation

So now that we have truth conditions, we can ask: what conditionals are employed in deliberation—indicative, subjunctive, or both?

- Some arguments obviously come from decision-theoretic considerations (such as: Newcomb-type cases).
- But some other arguments can be made directly from the language of deliberation.

DeRose on conditionals of deliberation

So now that we have truth conditions, we can ask: what conditionals are employed in deliberation—indicative, subjunctive, or both?

- Some arguments obviously come from decision-theoretic considerations (such as: Newcomb-type cases).
- But some other arguments can be made directly from the language of deliberation.

One argument we consider today: DeRose (2010).

DeRose on conditionals of deliberation

DeRose (2010) considers (variants of) the following:

(15) a. If I drive to work without filling up the tank, I will run out of petrol.

b. If I leave for my meeting 5 minutes before it starts, I will be late.

c. If the house is not painted, it will soon look quite shabby.

> *These all seem to be conditionals that would be useful in deliberation: to the extent that I have reason to believe one of them, then in so far as I desire its consequent to be true, I have reason to make (or to try to make, in cases where the truth of the antecedent is not completely up to me) its antecedent true.* *(DeRose 2010)*

DeRose on conditionals of deliberation

(15) a. If I drive to work without filling up the tank, I will run out of petrol.
b. If I leave for my meeting 5 minutes before it starts, I will be late.
c. If the house is not painted, it will soon look quite shabby.

(P1) These conditionals are paradigm examples of conditionals of deliberation.

(P2) These conditionals are not subjunctive, but rather indicative.

(C) So conditionals of deliberation are indicative and not subjunctive—contrary to what causal decision theorists suggest.

DeRose on conditionals of deliberation

(15) a. If I drive to work without filling up the tank, I will run out of petrol.
b. If I leave for my meeting 5 minutes before it starts, I will be late.
c. If the house is not painted, it will soon look quite shabby.

(P1) These conditionals are paradigm examples of conditionals of deliberation.

(P2) These conditionals are not subjunctive, but rather indicative.

(C) So conditionals of deliberation are indicative and not subjunctive—contrary to what causal decision theorists suggest.

How good is this argument?

The logical form of will-conditionals

To answer this: we need to say more about two issues:

The logical form of will-conditionals

To answer this: we need to say more about two issues:

- the meaning of *will*;
- the compositional semantics of conditionals.

The logical form of will-conditionals

The meaning of *will*.

- Traditionally, especially in tense logic, *will* is treated as a tense.
- But there are excellent arguments that *will* is actually a modal auxialiary.

We will consider two arguments for the claim that *will* is a modal.

Will: motivating a modal analysis

Argument #1: modal subordination.

Will: motivating a modal analysis

Argument #1: modal subordination.

It is well known that modal expressions can enter anaphoric relations.

Here is a classical example, due to Craige Roberts:

(16) A wolf might come in. It would eat you. (Roberts 1989)

Will: motivating a modal analysis

Argument #1: modal subordination.

It is well known that modal expressions can enter anaphoric relations.

Here is a classical example, due to Craige Roberts:

(16) A wolf might come in. It would eat you. (Roberts 1989)

- The *would*-claim is understood as a conditional: it says that the wolf would eat the interlocutor *in the scenario in which the wolf comes in*.
- Hence the second sentence in the discourse is subordinated, in the relevant sense, to the first sentence.

Will: motivating a modal analysis

Crucially, as Klecha (2014) points out, *will* displays an analogous behavior:

(17) a. If Katie travels to Berkeley, she will shop at Amoeba Records. She will buy a boxed set and a dozen used LPs. (Cariani 2021)

b. The supplies might arrive tomorrow. It will be late in the day. (Cariani and Santorio 2018)

c. Don't go near that bomb! It'll explode. (Binnick 1971)

Will: motivating a modal analysis

Crucially, as Klecha (2014) points out, *will* displays an analogous behavior:

(17) a. If Katie travels to Berkeley, she will shop at Amoeba Records. She will buy a boxed set and a dozen used LPs. (Cariani 2021)

b. The supplies might arrive tomorrow. It will be late in the day. (Cariani and Santorio 2018)

c. Don't go near that bomb! It'll explode. (Binnick 1971)

Parallel sentences involving past tense cannot be forced to undergo subordination. Consider:

(18) a. If Katie traveled to Berkeley, she shopped at Amoeba Records. She bought a boxed set and a dozen used LPs.

b. The supplies might have arrived yesterday. It was late in the day.

c. I hope he didn't go near that bomb! It exploded.

Will: motivating a modal analysis

Argument #2: the acquaintance inference.

Will: motivating a modal analysis

Argument #2: the acquaintance inference.

Background: Predicates of personal taste, like *tasty*, *delicious*, and *good* and *great* in some uses, trigger an inference to the effect that the speaker has had a certain relevant kind of experience.

To see this, notice that the continuations in (19) are all pragmatically odd.

(19) a. That sandwich is delicious, # but I haven't tried it.

b. Joan's book is great, # but I haven't read it.

c. The soup is disgusting, # but I haven't had it.

Will: motivating a modal analysis

Interestingly, this effect is suppressed under modals, including strong epistemic modals like *must* and *have to*.

(20) a. That sandwich must/has to/should/... be delicious, but I haven't tried it.

b. Joan's book must/has to/should/... to be great, but I haven't read it.

c. The soup must/has to/should/... be disgusting, but I haven't had it.

Will: motivating a modal analysis

Interestingly, this effect is suppressed under modals, including strong epistemic modals like *must* and *have to*.

(20) a. That sandwich must/has to/should/... be delicious, but I haven't tried it.

b. Joan's book must/has to/should/... to be great, but I haven't read it.

c. The soup must/has to/should/... be disgusting, but I haven't had it.

Will patterns with modals.

(21) a. That sandwich will be delicious, but I haven't tried it/won't try it.

b. Joan's book will be great, but I haven't read it/won't read it.

c. The soup will be disgusting, but I haven't had it/will not have it.

A modal analysis for will

If *will* is a modal, what analysis should we adopt?

A modal analysis for will

If *will* is a modal, what analysis should we adopt?

The obvious suggestion: an analysis on which it quantifies or selects from metaphysical possibilities, similarly to subjunctive *would*.
(Copley 2009, Cariani and Santorio 2018).

A modal analysis for will

If *will* is a modal, what analysis should we adopt?

The obvious suggestion: an analysis on which it quantifies or selects from metaphysical possibilities, similarly to subjunctive *would*.
(Copley 2009, Cariani and Santorio 2018).

My favorite option: a selection semantics, which makes it very close to the semantics for *would*.

(22) $[\![\textit{will } A]\!]^{w,f,s} = \text{true iff } [\![A]\!]^{s(w,s(f(w)),f,s} = \text{true}$

Conditionals and the restrictor theory

The compositional semantics of conditionals.

- So far, we have presupposed that, at the level of logical form, conditionals simply involve a binary connective.

$$\textit{match struck} \; \Box\!\!\rightarrow \; \textit{match lights}$$

Conditionals and the restrictor theory

The compositional semantics of conditionals.

- So far, we have presupposed that, at the level of logical form, conditionals simply involve a binary connective.

$$\textit{match struck} \; \Box\!\!\rightarrow \; \textit{match lights}$$

Conditionals and the restrictor theory

The compositional semantics of conditionals.

- So far, we have presupposed that, at the level of logical form, conditionals simply involve a binary connective.

$$\textit{match struck} \;\Box\!\!\rightarrow\; \textit{match lights}$$

- But this is not the standard view of conditionals in linguistic semantics.

Conditionals and the restrictor theory

The compositional semantics of conditionals.

- So far, we have presupposed that, at the level of logical form, conditionals simply involve a binary connective.

$$\textit{match struck} \; \Box\!\!\rightarrow \; \textit{match lights}$$

- But this is not the standard view of conditionals in linguistic semantics.
- On the standard view, the restrictor theory (Kratzer 1981, 1986, 2012), conditionals (i) involve a modal auxiliary and (ii) *if*-clauses are used to restrict the modal.

(23) a. If Alice is in Prague, Bob must be in Prague too.

b. must [if Alice is in Prague] [Bob is in Prague]

Conditionals and the restrictor theory

The compositional semantics of conditionals.

- So far, we have presupposed that, at the level of logical form, conditionals simply involve a binary connective.

$$\textit{match struck} \; \Box\!\!\rightarrow \; \textit{match lights}$$

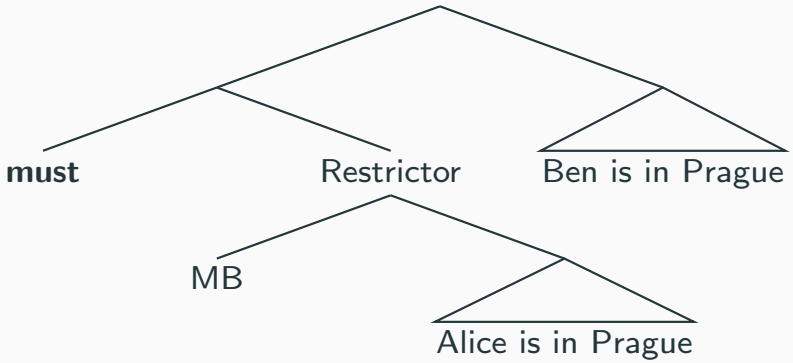
- But this is not the standard view of conditionals in linguistic semantics.
- On the standard view, the restrictor theory (Kratzer 1981, 1986, 2012), conditionals (i) involve a modal auxiliary and (ii) *if*-clauses are used to restrict the modal.

(23) a. If Alice is in Prague, Bob must be in Prague too.

b. must [if Alice is in Prague] [Bob is in Prague]

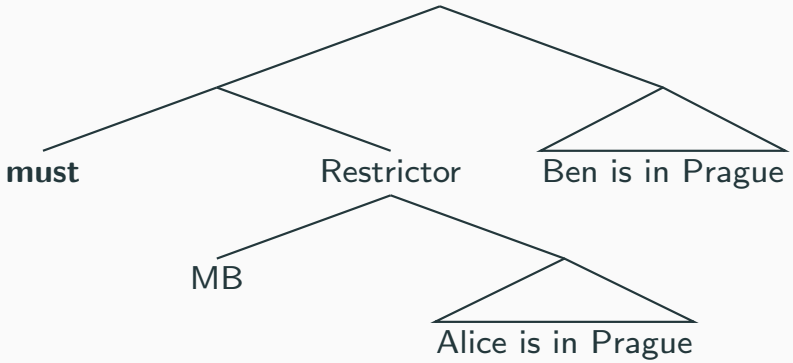
(24) a. If Alice is in Prague, Ben must be in Prague.

b.



(24) a. If Alice is in Prague, Ben must be in Prague.

b.



- This doesn't make a difference at the level of truth conditions.
- Where it matters: predicting how conditionals and overt modals interact.

Conditionals and the restrictor theory

What about conditionals that do not include an overt modal (bare conditionals)?

(25) If Maria danced, Frida danced.

Conditionals and the restrictor theory

What about conditionals that do not include an overt modal (bare conditionals)?

(25) If Maria danced, Frida danced.

- The standard view: they include a covert epistemic modal.
- So e.g. the logical form of (25) is:

(26) MUST [if Maria danced] [Frida danced]

The LF of will-conditionals

Given this treatment of *will* and this LF for conditionals, we can make some observations about *will*-conditionals (following Ciardelli and Ommundsen 2025).

(27) If you strike this match, it will light.

The LF of will-conditionals

Given this treatment of *will* and this LF for conditionals, we can make some observations about *will*-conditionals (following Ciardelli and Ommundsen 2025).

(27) If you strike this match, it will light.

We predict that (28) has two possible logical forms:

- single modal:

  will [you strike the match][match lights]
- double modal:

  MUST [you strike the match][*will* [match lights]]

The LF of will-conditionals

Given this treatment of *will* and this LF for conditionals, we can make some observations about *will*-conditionals (following Ciardelli and Ommundsen 2025).

(27) If you strike this match, it will light.

We predict that (28) has two possible logical forms:

- single modal:

  will [you strike the match][match lights]

- double modal:

  MUST [you strike the match][*will* [match lights]]

Interestingly, these two conditionals will differ in truth conditions!

An ambiguity in will-conditionals

(28) If you strike this match, it will light.

An ambiguity in will-conditionals

(28) If you strike this match, it will light.

Single modal:

will [you strike the match][match lights]

- *will* is the main modal of the conditional;
- the conditional has causal/metaphysical flavor.

An ambiguity in will-conditionals

(28) If you strike this match, it will light.

Single modal:

will [you strike the match][match lights]

- *will* is the main modal of the conditional;
- the conditional has causal/metaphysical flavor.

Double modal:

MUST [you strike the match][*will* [match lights]]

- covert MUST is the main modal of the conditional;
- the conditional has epistemic flavor.

An ambiguity in will-conditionals

(28) If you strike this match, it will light.

Single modal:

will [you strike the match][match lights]

- *will* is the main modal of the conditional;
- the conditional has causal/metaphysical flavor.

Double modal:

MUST [you strike the match][*will* [match lights]]

- covert MUST is the main modal of the conditional;
- the conditional has epistemic flavor.

So the prediction is: systematically, *will*-conditionals have two readings!

An ambiguity in will-conditionals

So: on one reading, *will*-conditionals are almost synonymous with subjunctives!

Indeed, this is what we have observed:

(29) a. If I strike this match, it will light.

b. If I were to strike this match, it would light.

An ambiguity in will-conditionals

So: on one reading, *will*-conditionals are almost synonymous with subjunctives!

Indeed, this is what we have observed:

(29) a. If I strike this match, it will light.

b. If I were to strike this match, it would light.

- The only difference is that (29b) carries a 'remoteness' inference of sorts (Iatridou 2000).
- What this is is a matter for another day.
- The key thing: both (29a) and (29b) have causal/metaphysical flavor.

An ambiguity in will-conditionals

Back to DeRose's observation:

(15) a. If I drive to work without filling up the tank, I will run out of petrol.
b. If I leave for my meeting 5 minutes before it starts, I will be late.
c. If the house is not painted, it will soon look quite shabby.

- DeRose is correct that conditionals in (15) are used in deliberation.
- But we can explain why without concluding that indicatives in general are conditionals of deliberation.
- We know that the conditionals in (15) are systematically ambiguous.
- They are used in deliberation when they are interpreted in their causal/metaphysical flavor (i.e. the same flavor we find in subjunctives).

An ambiguity in will-conditionals

A general takeaway: the classical philosophical distinction between indicatives and subjunctives is way too crude.

An ambiguity in will-conditionals

A general takeaway: the classical philosophical distinction between indicatives and subjunctives is way too crude.

A better taxonomy (to be still refined) exploits two dimensions:

- modal flavor (epistemic vs metaphysical);
- remoteness (which might connect to... mood; see Mendes 2025).

An ambiguity in will-conditionals

A general takeaway: the classical philosophical distinction between indicatives and subjunctives is way too crude.

A better taxonomy (to be still refined) exploits two dimensions:

- modal flavor (epistemic vs metaphysical);
- remoteness (which might connect to... mood; see Mendes 2025).

	No remoteness effect	Remoteness effect
Epistemic	Bare cond, epistemic *will*-cond *If Frida danced, Maria danced* MUST *[if I strike] [will [it lights]]*	[empty?]
Metaphysical	metaphysical will-conditionals *will [if I strike] [it lights]*	*would*-conditionals *if I struck, it would light*