Philosophy of Logic:
Modality, Conditionals, Vagueness

Lecture Notes
2014

Dirk Kindermann · Department of Philosophy · University of Graz
dirk.kindermann@uni-graz.at · www.dirkkindermann.com

Contents

Preface

The following are course notes written for my upper-level undergraduate class *Philosophical Logic: Modality, Conditionals, Vagueness*, taught at the University of Graz in spring 2014. The notes are based mainly on Priest (2008) and Sider (2010). They are best used in tandem with the readings from these and other sources as indicated at the end of every chapter.

I make no claim to originality with these notes. My knowledge of the areas covered here stems in large part from courses I took with Colin Caret, Patrick Greenough, John MacFarlane, Stephen Read, and Jason Stanley. Their slides and handouts have helped shape the way I present the material here – on occasion quite directly. Their influence is gratefully acknowledged. (Any mistakes and shortcomings are of course my responsibility alone.)

I hope you will find these notes useful.

Dirk Kindermann
Graz, July 2014

Part I

Preliminaries

1 Review: Propositional Logic

1.1 Language

Definition 1.1.1. A well-formed formula, or *wff*, of basic propositional logic is defined as follows:

- lowercase letters $p, q, r, s, \ldots$ are atomic formulas
- if A is a wff, so is $\neg A$
- if A and B are wffs, so are $(A \wedge B)$, $(A \vee B)$, $(A \supset B)$, $(A \equiv B)$[1]
- nothing else is a wff.

We sometimes omit writing the outermost brackets around a wff. **N.B.** the only individual letters which count as 'real formulae of our symbolic language are lowercase letters; the uppercase letters are **metavariables** which can represent, schematically, any 'real' formula.

Wffs are expressions in the **object language**; to talk *about* the object language and its various properties, we use a **metalanguage**. For instance, Definition 1.1.1 is formulated in the metalanguage.

We will have lax standards for use and mention. If you would like to re-acquaint yourself with the use-mention distinction, read and do the exercises on the **Use-Mention Handout** on Moodle.

1.2 Truth Tables

The classical theory of meaning of the connectives is captured by the following matrices.

[1]You may be used to writing the material conditional with the arrow ('$\rightarrow$') rather than the horseshoe ('$\supset$'), and the biconditional with a double arrow ('$\leftrightarrow$') rather than three lines ('$\equiv$'). These notations are equivalent, but we will reserve the horseshoe for the truth-conditional, material conditional.

$\neg$	
1	0
0	1

$\wedge$	1	0
1	1	0
0	0	0

$\vee$	1	0
1	1	1
0	1	0

$\supset$	1	0
1	1	0
0	1	1

$\equiv$	1	0
1	1	0
0	0	1

But this information can also be captured in a more 'formulaic way that will give us the flexibility to easily see how it ties into other logics down the line.

1.3 Model Theory

Definition 1.3.1. An **interpretation** ν is a function assigning a truth-value 0 or 1 (false/true) to each atomic formula. We extend this interpretation to all wffs by the following definition:

- $\nu(\neg A) = 1$ iff $\nu(A) = 0$
- $\nu(A \wedge B) = 1$ iff $\nu(A) = 1$ and $\nu(B) = 1$
- $\nu(A \vee B) = 1$ iff $\nu(A) = 1$ or $\nu(B) = 1$
- $\nu(A \supset B) = 1$ iff $\nu(A) = 0$ or $\nu(B) = 1$
- $\nu(A \equiv B) = 1$ iff $\nu(A) = \nu(B)$

Definition 1.3.2. We say that an interpretation ν of the language is a **model** of formula A just in case the given formula is true on that interpretation, i.e. $\nu(A) = 1$. Then ...

- An argument is **valid** iff every model of the premises is a model of the conclusion, which we gloss by saying that valid arguments are *truth-preserving* or have *no counter-models*. We write $\Sigma \vDash_{\mathsf{C}} A$ to mean that the inference from (the set of wffs) Σ to the conclusion A is valid according to classical logic C. The technical definition of this notion is:

  $$\Sigma \vDash_{\mathsf{C}} A \text{ iff for all interpretations } \nu, \text{ if } \nu(B) = 1 \text{ for all } B \in \Sigma, \text{ then } \nu(A) = 1.$$

- $\vDash_{\mathsf{C}} A$, that is, A is a **tautology** iff $\nu(A) = 1$ on every interpretation.
- ν is a counter-model to the inference from Σ to A if $\nu(B) = 1$ for all $B \in \Sigma$ and $\nu(A) = 0$. An argument with a counter-model is *invalid*, which we sometimes write $\Sigma \nvDash_{\mathsf{C}} A$.

1.4 Tableaux

A method for testing whether an argument is valid in classical logic is by constructing a tree derivation, which uses the following *resolution rules* at the nodes of the tree.

$$\wedge\text{-rule} \quad \begin{array}{c} A \wedge B \quad \surd \\ \downarrow \\ A \\ B \end{array} \qquad\qquad \neg\wedge\text{-rule} \quad \begin{array}{ccc} & \neg(A \wedge B) & \quad \surd \\ \swarrow & & \searrow \\ \neg A & & \neg B \end{array}$$

$$\vee\text{-rule} \quad \begin{array}{ccc} & A \vee B & \quad \surd \\ \swarrow & & \searrow \\ A & & B \end{array} \qquad\qquad \neg\vee\text{-rule} \quad \begin{array}{c} \neg(A \vee B) \quad \surd \\ \downarrow \\ \neg A \\ \neg B \end{array}$$

$$\supset\text{-rule} \quad \begin{array}{ccc} & A \supset B & \quad \surd \\ \swarrow & & \searrow \\ \neg A & & B \end{array} \qquad\qquad \neg\supset\text{-rule} \quad \begin{array}{c} \neg(A \supset B) \quad \surd \\ \downarrow \\ A \\ \neg B \end{array}$$

$$\equiv\text{-rule} \quad \begin{array}{ccc} & A \equiv B & \quad \surd \\ \swarrow & & \searrow \\ A & & \neg A \\ B & & \neg B \end{array} \qquad \neg\equiv\text{-rule} \quad \begin{array}{ccc} & \neg(A \equiv B) & \quad \surd \\ \swarrow & & \searrow \\ A & & \neg A \\ \neg B & & B \end{array} \qquad \text{DN-rule} \quad \begin{array}{c} \neg\neg A \quad \surd \\ \downarrow \\ A \end{array}$$

Definition 1.4.1. A branch of a tree for classical logic **closes** if it contains both a wff and its negation (i.e., both A and $\neg A$ for some formula A) The tree **closes** if every branch closes.

Definition 1.4.2. We say that A is **derivable** from (the set of wffs) Σ, written $\Sigma \vdash_{\mathsf{C}} A$, just if there is a closed tree with a starting list that includes the members of Σ as well as $\neg A$.

Definition 1.4.3. A logic L is **sound** if whenever $\Sigma \vdash_{\mathsf{L}} A, \Sigma \models_{\mathsf{L}} A$. A logic L is **complete** if whenever $\Sigma \models_{\mathsf{L}} A, \Sigma \vdash_{\mathsf{L}} A$.

1.4.1 Examples

$$
\begin{array}{ccc}
 & p \supset q & \surd \\
 & \neg(\neg q \supset \neg p) & \surd \\
 & \neg q & \\
 & \neg\neg p & \\
\swarrow & & \searrow \\
\neg p & & q \\
\times & & \times
\end{array}
$$

So we have shown that $p \supset q \vdash_{\mathsf{C}} \neg q \supset \neg p$.

$$
\begin{array}{ccccc}
 & & p \supset s & & \surd \\
 & & q \supset \neg s & & \surd \\
 & & \neg\neg(p \wedge q) & & \surd \\
 & & p \wedge q & & \surd \\
 & & p & & \\
 & & q & & \\
 & \swarrow & & \searrow & \\
\neg p & & & & s \\
 & & & \swarrow & & \searrow \\
\times & & \neg q & & & & \neg s \\
 & & \times & & & & \times
\end{array}
$$

So we have $p \supset s, q \supset \neg s \vdash_{\mathsf{C}} \neg(p \wedge q)$.

$$
\begin{array}{ccccc}
 & & p \supset s & & \surd \\
 & & \neg((p \vee q) \supset s) & & \surd \\
 & & p \vee q & & \surd \\
 & & \neg s & & \\
 & \swarrow & & \searrow & \\
 & \neg p & & & s \\
\swarrow & & \searrow & & \\
p & & q & & \times \\
\times & & & &
\end{array}
$$

There is an open branch, so the argument is invalid.
Counterexample: let $\nu(p) = 0, \nu(q) = 1, \nu(s) = 0$.
So $p \supset s \nvdash_{\mathsf{C}} (p \vee q) \supset s$.

1.5 The Greek Alphabet for Logic

In logic you will often find letters from the Greek alphabet used. For future reference, here is a list of the most commonly used letters. Where the upper case version is too similar to a Roman letter, it is not used:

lower case		upper case	lower case		upper case
α	alpha	A	ν	nu	N
β	beta	B	ξ	xi	Ξ
γ	gamma	Γ	o	omicron	O
δ	delta	Δ	π	pi	Π
ϵ	epsilon	E	ρ	rho	P
ζ	zeta	Z	σ	sigma	Σ
η	eta	H	τ	tau	T
θ	theta	Θ	υ	upsilon	Υ
ι	iota	I	ϕ	phi	Φ
κ	kappa	K	χ	chi	X
λ	lambda	Λ	ψ	psi	Ψ
μ	mu	M	ω	omega	Ω

1.6 Optional Exercises

1. Use Definition 1.1.1 to decide whether the following are well-formed formulae.

   a) $\neg(p \wedge \neg q) \supset (p \supset \neg q)$
   b) $p \wedge \neg p \supset p$
   c) $A \supset (B \supset A)$

   Explain your answers.

2. Fill in the quotes where necessary to make the following sentences true:

   a) New York City refers to New York City.
   b) Graz is in Styria, but Graz isn't in Styria.
   c) Moore's wife called Moore Moore.
   d) There are three words in the previous sentence.

3. Check the truth of each of the following, using tableaux. If the inference is invalid, read off a counter-model from the tree, and check directly that it makes the premises true and the conclusion false:

   (b) $p \supset (q \wedge r), \neg r \vdash_{\mathsf{C}} \neg p$
   (g) $p \wedge (\neg r \vee s), \neg(q \supset s) \vdash_{\mathsf{C}} r$

1.7 Readings

Priest (2008, §§0.1–0.3, 1.1–1.5, 1.12–1.13)

2 Review: Predicate Logic

2.1 Language

Definition 2.1.1. The **basic vocabulary** of the language of first-order logic (or predicate logic) $\mathcal{L}$ includes:

- individual variables x, y, z, with or without numerical subscripts
- individual constants a, b, c, with or without numerical subscripts
- for every natural number n ¿ 0, n-place predicates P^n, Q^n, S^n, with or without numerical subscripts[1]
- unary connective $\neg$ and binary connectives $\wedge, \vee, \supset, \equiv$
- quantifiers $\forall$ and $\exists$
- parentheses (,)

We will call any individual variable or constant a **term**.

Definition 2.1.2. A **well-formed formula**, or wff, of first-order logic is defined as follows:

- if Π is any n-place predicate and $t_1, \ldots, t_n$ are any terms, then $\Pi t_1 \ldots t_n$ is an atomic wff
- if A and B are wffs, so are $\neg A$, $(A \wedge B)$, $(A \vee B)$, $(A \supset B)$, $(A \equiv B)$
- if A is any wff and α is any variable, then $\forall\alpha A$, $\exists\alpha A$ are wffs
- nothing else is a wff.

Definition 2.1.3. An occurrence of a variable α in wff A is **bound** in A if that occurrence is within an occurrence of some wff of the form $\forall\alpha B$ or $\exists\alpha B$ within A. Otherwise the occurrence is **free** in A.
A formulae with no free occurrences of variables is said to be **closed**; otherwise it is **open**.
$A(\alpha/\beta)$ is the formula obtained by substituting β for each free occurrence of α in A.

[1]We may occasionally leave out the subscripts when the adicity of the predicate is obvious from the context.

2.2 Model Theory

Definition 2.2.1. A **model** $\mathcal{M}$ for PL is an ordered pair $\langle \mathcal{D}, \mathcal{J} \rangle$ such that:

- $\mathcal{D}$ is a non-empty set (the *domain*, or *universe*)
- $\mathcal{J}$ is a function (the *interpretation function*) obeying the following constraints:
  - if t is an individual constant then $\mathcal{J}(t) \in \mathcal{D}$
  - if Π is an n-place predicate, then $\mathcal{J}(\Pi)$ is an n-place relation over $\mathcal{D}$

Definition 2.2.2. g is a **variable assignment** for model $\langle \mathcal{D}, \mathcal{J} \rangle$ iff g is a function that assigns to each variable some object in $\mathcal{D}$.
$g^{\alpha/d}$ is the variable assignment that is just like g, except that it assigns d to α, where d is some object in $\mathcal{D}$. Note that $g^{\alpha/d}(\alpha) = d$.

Definition 2.2.3. Let $\mathcal{M}$ $(= \langle \mathcal{D}, \mathcal{J} \rangle)$ be a model, g be a variable assignment, and t be a term. $[t]_{\mathcal{M},g}$, i.e. the **denotation** of t (relative to $\mathcal{M}$ and g), is defined as follows:

$$[t]_{\mathcal{M},g} = \begin{cases} \mathcal{J}(t) & \text{if } t \text{ is a constant} \\ g(t) & \text{if } t \text{ is a variable} \end{cases}$$

Definition 2.2.4. The **valuation** function, $\nu_{\mathcal{M},g}$, for model $\mathcal{M}$ $(= \langle \mathcal{D}, \mathcal{J} \rangle)$ and variable assignment g, is defined as the function that assigns to each wff either 0 or 1 subject to the following constraints:

(i) for any n-place predicate Π and any terms $t_1 \dots t_n$, $\nu_{\mathcal{M},g}(\Pi t_1 \dots t_n) = 1$ iff $\langle [t_1]_{\mathcal{M},g} \dots [t_n]_{\mathcal{M},g} \rangle \in \mathcal{J}(\Pi)$.

(ii) For any wffs A, B and any variable α:

$$\begin{array}{rcl}
\nu_{\mathcal{M},g}(\neg A) = 1 & \text{iff} & \nu_{\mathcal{M},g}(A) = 0 \\
\nu_{\mathcal{M},g}(A \wedge B) = 1 & \text{iff} & \nu_{\mathcal{M},g}(A) = 1 \text{ and } \nu_{\mathcal{M},g}(B) = 1 \\
\nu_{\mathcal{M},g}(A \vee B) = 1 & \text{iff} & \nu_{\mathcal{M},g}(A) = 1 \text{ or } \nu_{\mathcal{M},g}(B) = 1 \\
\nu_{\mathcal{M},g}(A \supset B) = 1 & \text{iff} & \nu_{\mathcal{M},g}(A) = 0 \text{ or } \nu_{\mathcal{M},g}(B) = 1 \\
\nu_{\mathcal{M},g}(A \equiv B) = 1 & \text{iff} & \nu_{\mathcal{M},g}(A) = \nu_{\mathcal{M},g}(B) \\
\nu_{\mathcal{M},g}(\forall \alpha A) = 1 & \text{iff} & \text{for every } d \in \mathcal{D},\ \nu_{\mathcal{M},g^{\alpha/d}}(A) = 1 \\
\nu_{\mathcal{M},g}(\exists \alpha A) = 1 & \text{iff} & \text{for at least one } d \in \mathcal{D},\ \nu_{\mathcal{M},g^{\alpha/d}}(A) = 1
\end{array}$$

Read '$\nu_{\mathcal{M},g}(A) = 1$' as A *is true in (model)* $\mathcal{M}$ *relative to (variable assignment)* g.

Definition 2.2.5. A is **true in model** $\mathcal{M}$ iff $\nu_{\mathcal{M},g}(A) = 1$, for each variable assignment g for $\mathcal{M}$.

Definition 2.2.6. The inference from (the set of wffs) Σ to the conclusion A is **valid according to predicate logic PL** ($\Sigma \vDash_{\mathsf{PL}} A$) iff for every model $\mathcal{M}$ and every variable assignment g for $\mathcal{M}$, if $\nu_{\mathcal{M},g}(B) = 1$ for each $B \in \Sigma$, then $\nu_{\mathcal{M},g}(A) = 1$.
When $\Sigma \vDash_{\mathsf{PL}} A$, we also say that A is a **semantic consequence** in PL of the set of wffs Σ.
A wff A is **valid in PL** ($\vDash_{\mathsf{PL}} A$) iff A is true in all models for PL.

2.3 Tableaux

A method for testing whether an argument is valid in classical first-order logic is by constructing a tree derivation, which uses the following *resolution rules* at the nodes of the tree.

The resolution rules for classical first-order logic include all resolution rules for classical propositional logic (see 'Handout I Propositional Logic', §4), and four rules for the quantifiers:

$\forall$-rule $\quad \forall\alpha A \quad /\beta$
$\quad\quad\downarrow$
$\quad A(\alpha/\beta)$
for any constant β
already on the branch
N.B.: Never check off the node.

$\neg\forall$-rule $\quad \neg\forall\alpha A \quad \surd$
$\quad\quad\downarrow$
$\quad \exists\alpha\neg A$

$\exists$-rule $\quad \exists\alpha A \quad \surd$
$\quad\quad\downarrow$
$\quad A(\alpha/\beta)$
where β is a *new* constant
not yet on the branch

$\neg\exists$-rule $\quad \neg\exists\alpha A \quad \surd$
$\quad\quad\downarrow$
$\quad \forall\alpha\neg A$

Definition 2.3.1. A branch of a tree for classical logic **closes** if it contains both a wff and its negation (i.e., both A and $\neg A$ for some formula A) The tree **closes** if every branch closes.

Definition 2.3.2. We say that A is **derivable** from (the set of wffs) Σ, written $\Sigma \vdash_{\mathsf{PL}} A$, just if there is a closed tree with a starting list that includes the members of Σ as well as $\neg A$.

2.3.1 An Example

$$
\begin{array}{cc}
\forall x(Px \supset Qx) & /a/b \\
\exists x \neg Px & \surd a \\
\neg \forall x \neg Qx & \surd \\
\exists x \neg\neg Qx & \surd b \\
\neg Pa & \\
\neg\neg Qb & \surd \\
Qb & \\
Pa \supset Qa & \surd
\end{array}
$$

$$
\begin{array}{ccc}
& \swarrow \quad \searrow & \\
\neg Pa & & Qa \\
Pb \supset Qb & & Pb \supset Qb \\
\swarrow \quad \searrow & & \swarrow \quad \searrow \\
\neg Pb \qquad Qb & & \neg Pb \qquad Qb
\end{array}
$$

So $\forall x(Px \supset Qx)$, $\exists x \neg Px \nvdash_{\mathsf{PL}} \forall x \neg Qx$. Counter-model: $\langle \mathcal{D}, \mathcal{J} \rangle$ such that

$\mathcal{D} = \{d_a, d_b\}$
$\mathcal{J}(P) = \varnothing$ (importantly, $d_a \notin \mathcal{J}(P)$)
$\mathcal{J}(Q) = \{d_a, d_b\}$

2.4 Optional Exercises

1. Use Definition 2.1.2 to decide whether the following are well-formed formulae. Explain your answers.
   (a) Px
   (b) $\forall x(Px \supset Qa)$
   (c) $\exists y Px \wedge Qx$
   (d) $\forall \alpha(\Pi\alpha \supset \Psi\alpha)$
2. Show that $\nu_{\mathcal{M},g}(\forall \alpha A) = \nu_{\mathcal{M},g}(\neg \exists \alpha \neg A)$. (Hint: Using Definitions 2.1 – 2.4 on *Handout II Predicate Logic*, especially the clauses for the quantifiers and negation in Definition 2.4, reason in a series of biconditionals, starting with '$\nu_{\mathcal{M},g}(\forall \alpha A) = 1$ iff...'. Cf. Sider (2010, 95-6) for more help.)
3. Check the truth of each of the following, using tableaux. If the inference is invalid, use an open branch to specify a counter-model for the inference.
   (a) $\forall x(Px \supset Qx), \exists x(Qx \wedge Sx) \vdash_{\mathsf{PL}} \exists x(Px \wedge Sx)$
   (b) $\forall x(Px \supset Qx), \exists x(Px \wedge Sx) \vdash_{\mathsf{PL}} \exists x(Qx \wedge Sx)$

2.5 Readings

Obligatory reading: Priest (2008, ch. 12)

Optional readings: Sider (2010, §§4.1–4.3), Bell et al. (2001, §§2.1–2.3, 2.5–2.6)

3 Set Theory Tutorial

3.1 Sets

The basic intuition of set theory is that one can group objects together into a collection or **set**, in such a way that, presented with an object, u, and such a set, $\mathcal{A}$, one can sensibly ask whether the object belongs to, or is a **member** of, the set. The basic relation is symbolised by

$$u \in \mathcal{A}$$

If u is not a member of $\mathcal{A}$, we write

$$u \notin \mathcal{A}$$

A set is determined by its members:

Definition 3.1.1. The intuitive principle of extension. Two sets are equal iff they have the same members. We write '$\mathcal{A} = \mathcal{B}$' iff $\mathcal{A}$ and $\mathcal{B}$ are equal, and '$\mathcal{A} \neq \mathcal{B}$' iff $\mathcal{A}$ and $\mathcal{B}$ are inequal.

We can write sets by simply listing its members between curly brackets. Thus, $\{2, 4, 6\} = \{2, 6, 4\}$. Sets may be infinite, which we can write, e.g. in the following way: $\{1, 2, 3, \ldots\}$. Another way to give a set is by specifying the (necessary and sufficient) condition(s) for membership in the set:

Definition 3.1.2. The intuitive principle of abstraction. A formula $P(x)$ defines a set $\mathcal{A}$ by the convention that the members of $\mathcal{A}$ are exactly those objects a such that $P(a)$ is true. We write: $\mathcal{A} = \{x|P(x)\}$.

Example: $\{y|y \text{ is divisible by } 2\}$

$\{x\}$, a **unit set** or **singleton set**, is the set whose sole member is x. The set with no members is the **empty set**, $\varnothing$; that is, for every object u, u is not a member of $\varnothing$ ($u \notin \varnothing$).

3.1.1 Inclusion

If $\mathcal{A}$ and $\mathcal{B}$ are sets, then $\mathcal{A}$ is **included in** $\mathcal{B}$, symbolized by

$$\mathcal{A} \subseteq \mathcal{B},$$

iff each member of $\mathcal{A}$ is a member of $\mathcal{B}$. In this event one also says that $\mathcal{A}$ is a **subset** of $\mathcal{B}$. Further, we agree that $\mathcal{B}$ **includes** $\mathcal{A}$, symbolized by

$$\mathcal{B} \supseteq \mathcal{A},$$

iff $\mathcal{A}$ is included in $\mathcal{B}$. Thus, $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{B} \supseteq \mathcal{A}$ each means that, for all x, if $x \in \mathcal{A}$, then $x \in \mathcal{B}$. The set $\mathcal{A}$ is **properly included** in $\mathcal{B}$, symbolized by

$$\mathcal{A} \subset \mathcal{B},$$

(or, alternatively, $\mathcal{A}$ is a **proper subset** of $\mathcal{B}$, and $\mathcal{B}$ **properly includes** $\mathcal{A}$) iff $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{A} \neq \mathcal{B}$.

3.1.2 Operations for Sets

The **union** (or **sum**) of the sets $\mathcal{A}$ and $\mathcal{B}$, symbolized by $\mathcal{A} \cup \mathcal{B}$, is the set of all objects which are members either of $\mathcal{A}$ or of $\mathcal{B}$; that is,

$$\mathcal{A} \cup \mathcal{B} = \{x | x \in \mathcal{A} \text{ or } x \in \mathcal{B}\}$$

Example: $\{1, 2, 3\} \cup \{1, 3, 4\} = \{1, 2, 3, 4\}$

The **intersection** (or **product**) of the sets $\mathcal{A}$ and $\mathcal{B}$, symbolized by $\mathcal{A} \cap \mathcal{B}$ is the set of all objects which are members of both $\mathcal{A}$ and $\mathcal{B}$; that is,

$$\mathcal{A} \cap \mathcal{B} = \{x | x \in \mathcal{A} \text{ and } x \in \mathcal{B}\}$$

Example: $\{1, 2, 3\} \cap \{1, 3, 4\} = \{1, 3\}$

The **absolute complement** of a set $\mathcal{A}$, symbolized by

$$\overline{\mathcal{A}}$$

is $\{x | x \notin \mathcal{A}\}$. The **relative complement** of $\mathcal{A}$ with respect to a set $\mathcal{B}$ is $\mathcal{B} \cap \overline{\mathcal{A}}$; this is often shortened $\mathcal{B} - \mathcal{A}$. Thus

$$\mathcal{B} - \mathcal{A} = \{x \in \mathcal{B} | x \notin \mathcal{A}\},$$

that is, the set of those members of $\mathcal{B}$ which are not members of $\mathcal{A}$.

3.2 Relations

Monadic predicate letters of first-order logic have as their meaning 1-place relations; dyadic predicate letters have as their meaning 2-place relations; ... tetradic (3-place), and generally n-adic predicate letters have as their meaning n-place) relations.

Definition 3.2.1. An n-place **relation** is a set of n-tuples.

So a dyadic (2-place) relation is a set of ordered pairs. E.g., The *being less-than* relation for positive integers is the set of ordered pairs $\langle m, n\rangle$ such that m is a positive integer less than n, another positive integer. That is, it is the following set:

$$\{\langle 1,\, 2\rangle,\, \langle 1,\, 3\rangle,\, \langle 1,\, 4\rangle,\, \ldots,\, \langle 2,\, 3\rangle,\, \langle 2,\, 4\rangle,\, \ldots\}$$

Definition 3.2.2. The **domain** of a dyadic relation R is the set of all objects which are related by R to something, i.e.,

$$\mathrm{dom}(R) = \{x|\exists yRxy\}.$$

E.g., the domain of the relation of being a daughter of, is the set of all women and girls.

Definition 3.2.3. The **range** (or **co-domain** or **converse domain**) of R is the set of all objects to which anything is related by R, i.e.,

$$\mathrm{range}(R) = \{x|\exists yRyx\}.$$

The range of being a daughter of is the class of all men and women who have a daughter.

Definition 3.2.4. The **field** of R consists of its domain and range.

Often, a relation has the same domain and range, e.g., being married to, or $\leqslant$.

Definition 3.2.5. Given sets $\mathcal{A}_1, \ldots, \mathcal{A}_n$, their **Cartesian product**, $\mathcal{A}_1 \times \ldots \times \mathcal{A}_n$ is the set of all n-tuples, the first member of which is in $\mathcal{A}_1$, the second of which is in $\mathcal{A}_2$, etc. That is,

$$\mathcal{A}_1 \times \ldots \times \mathcal{A}_n = \{\langle \mathcal{A}_1, \ldots, \mathcal{A}_n\rangle \mid a_1 \in \mathcal{A}_1,\text{and } \ldots,\text{ and } a_n \in \mathcal{A}_n\}$$

Definition 3.2.6. Let R be any binary relation over some set $\mathcal{A}$.

- R is **serial** (in $\mathcal{A}$) iff for every $u \in \mathcal{A}$, there is some $v \in \mathcal{A}$ such that Ruv.
- R is **reflexive** (in $\mathcal{A}$) iff for every $u \in \mathcal{A}$, Ruu (e.g., being identical, being the same age as).
  R is **irreflexive** (in $\mathcal{A}$) iff for every $u \in \mathcal{A}$, $\neg Ruu$ (e.g., being next to, being less than).
  R is **non-reflexive** (in $\mathcal{A}$) iff for some $u \in \mathcal{A}$, Ruu, and for some $u \in \mathcal{A}$, $\neg Ruu$ (e.g., being two natural numbers whose product is even; loving).

- R is **symmetric** iff for all u, v, if Ruv then Rvu (e.g., being identical, being adjacent to).
  R is **asymmetric** iff for all u, v, if Ruv then $\neg Rvu$ (e.g., being less than).
  R is **anti-symmetric** iff for all u, v, if Ruv and $u \neq v$ then $\neg Rvu$ (e.g., being less than or equal to).
  R is **non-symmetric** iff for some u, v, Ruv and $\neg Rvu$, and for some u, v Ruv and Rvu (e.g., liking).
- R is **transitive** iff for any u, v, w, if Ruv and Rvw then Ruw (e.g., identity, being less than, being less than or equal to).
  R is **intransitive** iff for any u, v, w, if Ruv and Rvw then $\neg Ruw$ (e.g., being the square of (on the positive integers $\geqslant 2$)).
  R is **non-transitive** iff for some u, v, w, if Ruv and Rvw then $\neg Ruw$, and for some u, v, w, if Ruv and Rvw then $\neg Ruw$ (e.g., being similar, liking).
- R is an **equivalence relation** (in $\mathcal{A}$) iff R is symmetric, transitive, and reflexive (in $\mathcal{A}$) (e.g., being equal to (in $\mathcal{A}=\mathsf{N}$), having the same birthday as (in $\mathcal{A}$ = the set of all people)).

3.3 Functions

Definition 3.3.1. A **function** is a set of ordered pairs, f, obeying the condition that if $\langle u, v\rangle$ and $\langle u, w\rangle$ are both members of f, then $v = w$.
When $\langle u, v\rangle \in f$, we say that u is an **argument** of f, v is a **value** of f, and that f **maps** u to v; we write '$f(u) = v$.' The domain of a function is the set of its arguments, its range is the set of its values. A function is n-place when every member of its domain is an n-tuple.

A function is a binary relation that never relates a single argument to two distinct values. A function is called **one-to-one** if it maps distinct elements to distinct elements; i.e., a function f is one-to-one iff $u \neq v$ implies $f(u) \neq f(v)$. For instance, $f(x) = 2x + 1$ (in N) is one-to-one.

3.4 Readings

Obligatory reading: Priest (2008, §§01.–0.3)

Optional reading: Sider (2010, pp. 12–16)

Part II

Modality

4 Propositional Modal Logic

4.1 Motivating Modal Semantics

- Modal logic is narrowly defined as the logic of necessity and possibility: *it is necessary that. . .* & it it is possible that. . .
- It concerns two **modes** in which propositions (more generally, any truth bearer) can be true or false.
- The notion of **modality** (in contemporary linguistics) is much wider: "modality is the linguistic phenomenon whereby grammar allows one to say things about, or on the basis of, situations which need not be real." (Portner, 2009, 1)
- It's an open research question which features of language are associated with modality. Take for example tense: Are the past and future real? Hence, do past tense and future tense expressions (*-ed*, *will*+verb) have modal meanings?
- Kinds of (English) expressions that have modal meanings (cf. von Fintel (2006))
  1. Modal auxiliaries:
     Sandy must/should/might/may/could be home.
  2. Semimodal Verbs:
     Sandy has to/ought to/needs to be home.
  3. Modal adverbs:
     Perhaps, Sandy is home.
  4. Nouns:
     There is a slight possibility that Sandy is home.
  5. Adjectives:
     It is far from necessary that Sandy is home.
  6. Conditionals:
     If the light is on, Sandy is home.
- Kinds of Modal Meaning:

- **Alethic/logical/metaphysical modality** is hard to find in natural language but matters to philosophy: it concern what is in the widest sense/logically/metaphysically possible or necessary.
- **Epistemic modality** (Greek *episteme*, meaning 'knowledge) concerns what is possible or necessary given what is known and what the available evidence is.

  (4.1) A: Where is Paul?
  B: I don't know. He may be at home.
- **Deontic modality** (Greek: *deon*, meaning 'duty) concerns what is possible, necessary, permissible, or obligatory, given a body of law or a set of moral principles or the like.

  (4.2) He may bring his partner to the dinner.
- **Bouletic modality** concerns what is possible or necessary, given a persons desires.

  (4.3) You should try this cake, given how much you love chocolate.
- **Circumstantial modality** (sometimes called dynamic modality) concerns what is possible or necessary, given a particular set of circumstances.

  (4.4) Tulips can grow here.
- **Teleological modality** (Greek *telos*, meaning 'goal) concerns what means are possible or necessary for achieving a particular goal.

  (4.5) To get to the Isle of Mull, you must take a ferry.
- ⋮

Why a logic of (different kinds of) necessity and possibility?

(4.6) Durs Grünbein isn't *necessarily* going to win a Nobel prize next year.

(4.7) It's *possible* that Durs Grünbein will not win a Nobel prize next year.

- It seems that if (4.6) is true, (4.7) has to be true as well.
- This inference relies crucially on the modal adverb *necessarily* and the sentential modal operator *possible*.

Can we just add modal operators to propositional logic?

- Let $\Box A$ mean *It is necessary that A.*
- Let $\Diamond A$ mean *It is possible that A.*
- Logical operators in propositional logic are truth-functional. What could the truth-functional meanings be of $\Box A$ and $\Diamond A$?

Win	$\Diamond$ Win
1	1
0	?

- Both choices (0 & 1) get things wrong:
  - Its consistent to say that *Grünbein didn't win a Nobel prize, but it was possible for him to win one.*
    So 0 is wrong
  - Also consistent: *Grünbein didn't win a Nobel prize, and it wasnt even possible for him to win it.*
    So 1 is wrong too.
- The truth value of A doesn't determine the truth value of $\Diamond A$.
- Parallel difficulties apply to $\Box A$.
- We need more in our semantics than *truth-functional* propositional operators.

Two important ideas that modal semantics implement:

1. **Possible worlds**: a possible world can be understood as a way the world might have been – a way that the totality of things/events/states of affairs might have been. The way things in fact *are* – the so-called **actual** world – is also a possible world.
2. **Relative possibility**: what is possible (necessary) given how things are may be different from what is possible given how things could be. For example, given how things actually are, it is (physically) necessary that the Earth's standard acceleration due to gravity is $g = 9.80665$ m/s^2. But in a world in which different laws of physics hold sway, and/r Earth has a different mass, Earth's standard acceleration could be necessarily different. Thus, what is (physically) possible/necessary relative to one world need not be what is (physically) possible/necessary relative to another world.

4.2 Language

Definition 4.2.1. A **well-formed formula**, or wff, of propositional modal logic is defined as follows:

- lowercase letters $p, q, r, s, \ldots$ are atomic formulas
- if A is a wff, so are $\neg A$, $\Box A$, $\Diamond A$
- if A and B are wffs , so are $(A \wedge B)$, $(A \vee B)$, $(A \supset B)$, $(A \equiv B)$
- nothing else is a wff.

4.3 Model Theory

- To define a **model** we need a bit of extra machinery to help us implement the Leibnizian equivalence: 'Possibly, φ' $=_{\text{df.}}$ 'φ is true at some possible world'.

- In order to fix the range of quantification when we say '... at some world' we introduce the idea of a relation R (*accessibility*) and read wRx as 'x is possible relative to w'.

Definition 4.3.1. A **model** $\mathcal{M}$ for modal propositional logic is a structure $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ where:

- $\mathcal{W}$ is a non-empty set of objects, intuitively understood as possible worlds
- $\mathcal{R}$ is an accessibility relation between worlds; i.e. $\mathcal{R}$ is a binary relation on $\mathcal{W}$ (so that $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$). We write '$w_1 \mathcal{R} w_2$' for '$w_2$ is accessible from w_1', or 'w_1 sees w_2', which means intuitively that w_2 is possible given/relative to w_1.
- $\mathcal{J}$ is a function assigning a truth-value to each atomic formula relative to each world. That is, for any propositional letter α, and any $w \in \mathcal{W}$, $\mathcal{J}(\alpha, w)$ is either 1 or 0. We will sometimes equivalently write $\mathcal{J}_w(\alpha)$.

Definition 4.3.2. A **frame** $\mathcal{F}$ is an ordered pair $\langle \mathcal{W}, \mathcal{R} \rangle$, where $\mathcal{W}$ is a non-empty set of objects (possible worlds) and $\mathcal{R}$ is an accessibility relation between worlds. A model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ is said to be **based** on the frame $\langle \mathcal{W}, \mathcal{R} \rangle$.

Definition 4.3.3. Where $\mathcal{M}(= \langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle)$ is any model for modal propositional logic, the **valuation** for $\mathcal{M}$, $\nu_{\mathcal{M}}$, is defined as the two-place function that assigns either 0 or 1 to each wff relative to each member of $\mathcal{W}$, subject to the following constraints, where α is any propositional letter, A and B are any wffs, and w is any member of $\mathcal{W}$:

$$\begin{array}{rcl}
\nu_{\mathcal{M},w}(\alpha) & = & \mathcal{J}_{\mathcal{M},w}(\alpha) \\
\nu_{\mathcal{M},w}(\neg A) = 1 & \text{iff} & \nu_{\mathcal{M},w}(A) = 0 \\
\nu_{\mathcal{M},w}(A \wedge B) = 1 & \text{iff} & \nu_{\mathcal{M},w}(A) = 1 \text{ and } \nu_{\mathcal{M},w}(B) = 1 \\
\nu_{\mathcal{M},w}(A \vee B) = 1 & \text{iff} & \nu_{\mathcal{M},w}(A) = 1 \text{ or } \nu_{\mathcal{M},w}(B) = 1 \\
\nu_{\mathcal{M},w}(A \supset B) = 1 & \text{iff} & \nu_{\mathcal{M},w}(A) = 0 \text{ or } \nu_{\mathcal{M},w}(B) = 1 \\
\nu_{\mathcal{M},w}(A \equiv B) = 1 & \text{iff} & \nu_{\mathcal{M},w}(A) = \nu_{\mathcal{M},w}(B) \\
\nu_{\mathcal{M},w}(\Box A) = 1 & \text{iff} & \nu_{\mathcal{M},x}(A) = 1 \text{ at all worlds } x \text{ such that } w\mathcal{R}x \\
\nu_{\mathcal{M},w}(\Diamond A) = 1 & \text{iff} & \nu_{\mathcal{M},x}(A) = 1 \text{ at some world } x \text{ such that } w\mathcal{R}x
\end{array}$$

Where the context makes the model clear, we will sometimes write $\nu_w(A)$ instead of $\nu_{\mathcal{M},w}(A)$.

Definition 4.3.4. We say that a world w of model $\mathcal{M}(= \langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle)$ **models** formula A just in case the given formula is true at that world on that model, i.e. $\nu_{\mathcal{M},w}(A) = 1$.

Let $\mathcal{M}$ be a model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$. We say that a formulae A is **true in** $\mathcal{M}$ iff for every world $w \in \mathcal{W}$, $\nu_{\mathcal{M},w}(A) = 1$.

Using K (for Kripke) to refer to our basic modal logic, we say that an inference is **valid in system K** iff every world of every model that models the premises also models the conclusion; i.e.

$$\begin{array}{c}
\Sigma \vDash_{\mathsf{K}} A \text{ iff for all worlds } w \in W \text{ of all models } \langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle\text{:} \\
\text{if } \nu_{\mathcal{M},w}(B) = 1 \text{ for all the premises } B \in \Sigma, \text{ then } \nu_{\mathcal{M},w}(A) = 1
\end{array}$$

When $\Sigma \vDash_{\mathsf{K}} A$, we also say that A is a **semantic consequence in K** of the set of wffs Σ.

$\vDash_{\mathsf{K}} A$, that is, A is **valid** iff $\nu_{\mathcal{M},w}(A) = 1$ for every world w of every model $\mathcal{M}$.

Model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ with world w gives a **counter-model** to the inference from Σ to A if $\nu_{\mathcal{M},w}(B) = 1$ for all $B \in \Sigma$ but $\nu_{\mathcal{M},w}(A) = 0$. This makes the inference *invalid*, $\Sigma \nvDash_{\mathsf{K}} A$.

4.4 Tableaux

A method for testing validities in basic modal logic K is by constructing a tree derivation using the following *resolution rules*. The letters i, j, k, etc. stand for numeric *world-indices*.

$$\wedge\text{-rule}\quad \begin{array}{c} A \wedge B, i \quad \checkmark \\ \downarrow \\ A, i \\ B, i \end{array} \qquad\qquad \neg\wedge\text{-rule}\quad \begin{array}{cc} \multicolumn{2}{c}{\neg(A \wedge B), i \quad \checkmark} \\ \swarrow & \searrow \\ \neg A, i & \neg B, i \end{array}$$

$$\vee\text{-rule}\quad \begin{array}{cc} \multicolumn{2}{c}{A \vee B, i \quad \checkmark} \\ \swarrow & \searrow \\ A, i & B, i \end{array} \qquad\qquad \neg\vee\text{-rule}\quad \begin{array}{c} \neg(A \vee B), i \quad \checkmark \\ \downarrow \\ \neg A, i \\ \neg B, i \end{array}$$

$$\supset\text{-rule}\quad \begin{array}{cc} \multicolumn{2}{c}{A \supset B, i \quad \checkmark} \\ \swarrow & \searrow \\ \neg A, i & B, i \end{array} \qquad\qquad \neg\supset\text{-rule}\quad \begin{array}{c} \neg(A \supset B), i \quad \checkmark \\ \downarrow \\ A, i \\ \neg B, i \end{array}$$

$$\equiv\text{-rule}\quad \begin{array}{cc} \multicolumn{2}{c}{A \equiv B, i \quad \checkmark} \\ \swarrow & \searrow \\ A, i & \neg A, i \\ B, i & \neg B, i \end{array} \qquad \neg\equiv\text{-rule}\quad \begin{array}{cc} \multicolumn{2}{c}{\neg(A \equiv B), i \quad \checkmark} \\ \swarrow & \searrow \\ A, i & \neg A, i \\ \neg B, i & B, i \end{array} \qquad \text{DN-rule}\quad \begin{array}{c} \neg\neg A, i \quad \checkmark \\ \downarrow \\ A, i \end{array}$$

$\Diamond$-rule

$$\begin{array}{c} \Diamond A, i \quad \checkmark \\ \downarrow \\ irj \\ A, j \end{array}$$

using some *new* index j

$\neg\Diamond$-rule

$$\begin{array}{c} \neg\Diamond A, i \quad \checkmark \\ \downarrow \\ \Box\neg A, i \end{array}$$

$\Box$-rule

$$\begin{array}{c} \Box A, i \\ irj \\ \downarrow \\ A, j \end{array}$$

for *every* j such that irj is already on the branch. N.B. we *never* check off a $\Box$ line

$\neg\Box$-rule

$$\begin{array}{c} \neg\Box A, i \quad \checkmark \\ \downarrow \\ \Diamond\neg A, i \end{array}$$

An optional, but useful trick is to draw a slash next to a $\Box$ line. Each time you apply the $\Box$-rule write down the index you are applying it to so you know you don't have to do that one again.

Definition 4.4.1. A branch of a tree for modal logic **closes** if it contains a wff and its negation with the same *world-index* (i.e., both A, k and $\neg A, k$) The tree **closes** if every branch does.

Definition 4.4.2. We say that there is a modal tableaux **proof** from Σ to A, written $\Sigma \vdash_{\mathsf{TK}} A$, just if the tree whose starting list includes $B, 0$ for each $B \in \Sigma$ as well as $\neg A, 0$ closes.

Definition 4.4.3. We say that there is A is a **theorem of K**, written $\vdash_{\mathsf{TK}} A$, just if the tree whose starting list includes $\neg A, 0$ closes.

4.4.1 Examples

(i)

$$\begin{array}{ccc}
\multicolumn{2}{c}{\Box(p \supset q), 0} & \\
\multicolumn{2}{c}{\neg(\Diamond p \supset \Diamond q), 0} & \checkmark \\
\multicolumn{2}{c}{\Diamond p, 0} & \checkmark \\
\multicolumn{2}{c}{\neg\Diamond q, 0} & \checkmark \\
\multicolumn{2}{c}{\Box\neg q, 0} & \\
\multicolumn{2}{c}{0r1} & \\
\multicolumn{2}{c}{p, 1} & \\
\multicolumn{2}{c}{p \supset q, 1} & \checkmark \\
\swarrow & \searrow & \\
\neg p, 1 & q, 1 & \\
\times & \neg q, 1 & \\
 & \times &
\end{array}$$

So $\Box(p \supset q) \vdash_{\mathsf{TK}} \Diamond p \supset \Diamond q$

(ii)

$$\begin{array}{cccc}
\multicolumn{3}{c}{\neg(\Box(p \wedge q) \supset (\Box p \wedge \Box q)), 0} & \checkmark \\
\multicolumn{3}{c}{\Box(p \wedge q), 0} & \\
\multicolumn{3}{c}{\neg(\Box p \wedge \Box q), 0} & \checkmark \\
\swarrow & & \searrow & \\
\neg\Box p, 0 \ \checkmark & & \neg\Box q, 0 & \checkmark \\
\Diamond\neg p, 0 \ \checkmark & & \Diamond\neg q, 0 & \checkmark \\
0r1 & & 0r1 & \\
\neg p, 1 & & \neg q, 1 & \\
p \wedge q, 1 \ \checkmark & & p \wedge q, 1 & \checkmark \\
p, 1 & & q, 1 & \\
\times & & \times &
\end{array}$$

So $\vdash_{\mathsf{TK}} \Box(p \wedge q) \supset (\Box p \wedge \Box q)$

(iii)

$$\begin{array}{cccc}
\multicolumn{3}{c}{\neg(\Box(p \vee q) \supset (\Box p \vee \Box q)), 0} & \checkmark \\
\multicolumn{3}{c}{\Box(p \vee q), 0} & \\
\multicolumn{3}{c}{\neg(\Box p \vee \Box q), 0} & \checkmark \\
\multicolumn{3}{c}{\neg\Box p, 0} & \checkmark \\
\multicolumn{3}{c}{\neg\Box q, 0} & \checkmark \\
\multicolumn{3}{c}{\Diamond\neg p, 0} & \checkmark \\
\multicolumn{3}{c}{\Diamond\neg q, 0} & \checkmark \\
\multicolumn{3}{c}{0r1} & \\
\multicolumn{3}{c}{\neg p, 1} & \\
\multicolumn{3}{c}{p \vee q, 1} & \checkmark \\
\swarrow & \multicolumn{2}{c}{\searrow} & \\
p, 1 & \multicolumn{2}{c}{q, 1} & \\
 & \multicolumn{2}{c}{0r2} & \\
 & \multicolumn{2}{c}{\neg q, 2} & \\
 & \multicolumn{2}{c}{p \vee q, 2} & \\
 & \swarrow & \searrow & \\
 & p, 2 & q, 2 & \\
 & & \times &
\end{array}$$

An open branch, so invalid.
Counterexample:

$$\begin{array}{cccc}
 & & w_1 & q \\
 & \nearrow & & \neg p \\
w_0 & & & \\
 & \searrow & & \\
 & & w_2 & p \\
 & & & \neg q
\end{array}$$

That is, $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ s.t.

$$\begin{aligned}
&\mathcal{W} = \{w_0, w_1, w_2\} \\
&\mathcal{R} = \{\langle w_0, w_1 \rangle, \langle w_0, w_2 \rangle\} \\
&\mathcal{J}(q, w_1) = 1 \\
&\mathcal{J}(p, w_1) = 0 \\
&\mathcal{J}(p, w_2) = 1 \\
&\mathcal{J}(q, w_2) = 0
\end{aligned}$$

So $\nvdash_{\mathsf{TK}} \Box(p \vee q) \supset (\Box p \vee \Box q)$

(iv) $\neg(\Box p \supset p), 0$ ✓
$\Box p, 0$
$\neg p, 0$

Invalid. Counterexample:

w_0 $\neg p$

That is, $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ s.t.

$$\mathcal{W} = \{w_0\}$$
$$\mathcal{R} = \varnothing$$
$$\mathcal{J}(p, w_0) = 0$$

So $\not\models_{\mathsf{TK}} \Box p \supset p$

Note that in this frame there is no world accessible from w_0, not even w_0 itself. That's why the counterexample works: $\mathcal{J}(\Box p, w_0) = 1$, since there is no world accessible from w_0 where p is false—the only world, and indeed the only world where p is false, is not accessible from w_0.
The fact that $\Box p \supset p$ is invalid in K shows that K is not really a theory of necessity and possibility. Later we will extend K to stronger systems ($T = K\rho$, $S4 = K\rho\tau$) where wffs like $\Box p \supset p$ are valid and which are more plausibly theories of necessity and possibility.

4.5 Possible Worlds – Ontological Positions

- Realism:
  - Modal realism: this world is just one of many equally real and concrete worlds, which are causally and spatiotemporally unrelated (D. Lewis, McCall)
  - Moderate realism (a form of actualism according to Priest): this world is actual and concrete, others merely possible, abstract ways this world might have been (Stalnaker, Plantinga)
- Actualism:
  - Linguistic ersatzism (called thusly by D. Lewis): (other) worlds are sets of sentences (Cresswell)
  - Combinatorialism: (other possible) combinations of atoms from this world (Armstrong)

4.6 Optional Exercises

1. Show that the truth value of $\neg\Box A$ at a world (and relative to a model) is the same as that of $\Diamond\neg A$ (see Priest (2008, §2.3.9) for help).
2. Show that the following formulae are valid in K (i.e. $\models_{\mathsf{K}} A$). (Hint: See Sider (2010, Example 6.1, §6.3.2) for a very similar validity proof.)

   2. $\models_{\mathsf{K}} \Box(p \supset p)$
   1. $\models_{\mathsf{K}} \neg\Diamond(p \wedge \neg p)$
   3. $\models_{\mathsf{K}} \Box(p \supset q) \supset (\Box p \supset \Box q)$ (often called 'K')

3. Test the following, using tableaux. Where the tableau does not close, use it to define a counter-model, and draw this, as in Priest (2008, §2.4.8).

   1. $\vdash_{\mathsf{TK}} (\Box p \wedge \Box q) \supset \Box(p \wedge q)$
   2. $\vdash_{\mathsf{TK}} \Diamond(p \wedge q) \supset (\Diamond p \wedge \Diamond q)$
   3. $\Box p, \Box\neg q \vdash_{\mathsf{TK}} \Box(p \supset q)$
   4. $\Diamond p, \Diamond q \vdash_{\mathsf{TK}} \Diamond(p \wedge q)$

4.7 Readings

Obligatory reading: Priest (2008, ch. 2 & §§3.1–3.6)

Optional readings: Sider (2010, §§6.1-6.3); on possible worlds: Read (1994, 96–109)

5 Normal Propositional Modal Logics

5.1 Introduction

- So far, we studied the system K (for Kripke) of propositional modal logic. A characteristic theorem of system K is K:

  K: $\Box(A \supset B) \supset (\Box A \supset \Box B)$
- K is plausible enough for necessity: If it's necessary that B follows from A, then necessarily B follows from necessarily A.
- Now consider the formula D:

  D: $\Box A \supset \Diamond A$
- D also seems plausible for necessity: If A is necessary, then it is possible. But D is not a theorem of K. Is K the right system for necessity?
- There are many systems of modal logics, some of which are more plausibly capturing (a particular kind of) necessity and possibility than others (cf. the kinds of modal meanings on *Handout III-1*).
- We are looking at some of the more famous ones: normal modal logics D, T, B, $\mathsf{S4}$, $\mathsf{S5}$.

5.2 Normal Systems of Modal Logic

Definition 5.2.1. A **system of modal logic**, K_n, is a set of premises-conclusion pairs, $\langle \Sigma, A \rangle$, (where Σ can be $\varnothing$) such that $\Sigma \vdash_{\mathsf{K}_n} A$. ($\mathsf{K}_n$ is the set of inferences derivable in it.) We also call $\mathsf{K}n$ a modal logic.

Definition 5.2.2. A system of modal logic, K_n, is an **extension** of a system K_m just in case if $\Sigma \vdash_{\mathsf{K}_m} A$, then $\Sigma \vdash_{\mathsf{K}_n} A$. That is, every inference derivable in K_m is derivable in K_n, and every theorem of K_m is a theorem of K_n.

Note that by the soundness and completeness of K_m and K_n (cf. *Handout I*, Definition 4.3), it also holds that K_n is an extension of K_m just in case if $\Sigma \vDash_{\mathsf{K}_m} A$, then $\Sigma \vDash_{\mathsf{K}_n} A$.

That is, every inference that is valid in K_m is valid in K_n, and every logical truth of K_m ($\vDash_{\mathsf{K}_m} A$) is a logical truth of K_n; the set of inferences valid in K_m (logical truths of K_m) is a subset of the inferences valid in K_n (logical truths in K_n).

When K_n is an extension of K_m, we also say that K_n is **at least as strong** than K_m. (When it is a proper extension of K_m, we say that it is **stronger** than K_m.)

Definition 5.2.3. A system of modal logic is **normal** iff it is an extension of K (i.e., iff it is at least as strong as K).

5.2.1 System D (Kη)

- A system stronger than K is D (Priest calls it $\mathsf{K}\eta$).
- A characteristic theorem of D is D:

  D: $\Box A \supset \Diamond A$

- On a deontic reading of the modal operators, where '$\Box$ means 'it is obligatory that' and '$\Diamond$ means 'it is permissible that,' '$\Box A \supset \Diamond A$ is essentially the principle 'ought implies can.' So D looks like a reasonable candidate for deontic modality. Note that in a deontic logic, we dont want '$\Box A \supset A$, since often what ought to be the case isnt the case.

Model Theory

- We get a stronger notion of validity by restricting the models $\mathcal{M}$ we quantify over to those with an accessibility relation $\mathcal{R}$ that satisfies some restriction. (The fewer models we consider for truth-preservation, the easier it is to preserve truth from premises to conclusion; the fewer models a logical truth has to be true in, the easier it is for it to be true in all of them.)
- The restriction that $\mathcal{R}$ has to satisfy to be a model we quantify over in the definition of $\vDash_{\mathsf{D}}$ is serialness (Priest calls it 'extendability' and uses the Greek letter 'η'):

  serialness η for all worlds w, there is some world w' such that $w\mathcal{R}w'$
  (every world can see *some* world)

- Let a **D-model**, or **serial model**, be a model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ whose accessibility relation $\mathcal{R}$ is serial.

Definition 5.2.4. An inference is **valid in system D** iff every world of every serial model that models the premises also models the conclusion; i.e.

$$\begin{aligned}&\Sigma \vDash_{\mathsf{D}} A \text{ iff for all worlds } w \in \mathcal{W} \text{ of all D-models } \langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle:\\ &\text{if } \nu_{\mathcal{M},w}(B) = 1 \text{ for all the premises } B \in \Sigma, \text{ then } \nu_{\mathcal{M},w}(A) = 1\end{aligned}$$

$\vDash_{\mathsf{D}} A$, that is, A is **valid in D** iff $\nu_{\mathcal{M},w}(A) = 1$ for every world w of every D-model $\mathcal{M}$.

- D is stronger than K: there are K-models that are not D-models.

- There are no D-models that are counter-models to D ($\Box A \supset \Diamond A$), but there are K-models that are counter-models to D. (A counter-model to a formula is one in which the formula is false.)

Optional Exercise: Find a K-model in which D is false.

Tableaux

- The tableaux rules for D include all of the rules for K, plus the following:

η-rule

$$\begin{array}{c} . \\ \downarrow \\ irj \end{array}$$

for any integer i already on the branch, provided there is not already something of the form irj on the branch, and with j being *new* on the branch.

- Example: $\vdash_{\mathsf{D}} \Box p \supset \Diamond p$

$$\begin{array}{c} \neg(\Box p \supset \Diamond p), 0 \checkmark \\ \downarrow \\ \Box p, 0 \\ \neg\Diamond p, 0 \checkmark \\ \downarrow \\ \Box\neg p, 0 \\ \downarrow \\ 0r1 \\ \downarrow \\ p, 1 \\ \downarrow \\ \neg p, 1 \\ \times \end{array}$$

5.2.2 System T ($\mathsf{K}\rho$)

- A characteristic theorem of T is T:

  T: $\Box A \supset A$

- A **T-model**, or **reflexive model**, is a model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ whose accessibility relation $\mathcal{R}$ is reflexive.

  reflexivity ρ $w\mathcal{R}w$ for all worlds w (every world can see itself)

Definition 5.2.5. An inference is **valid in system T**, $\Sigma \vDash_{\mathsf{T}} A$, iff every world of every reflexive model that models the premises also models the conclusion. $\vDash_{\mathsf{T}} A$, that is, A is **valid in T** iff $\nu_{\mathcal{M},w}(A) = 1$ for every world w of every T-model $\mathcal{M}$.

- T is stronger than K: there are K-models that are not T-models.
- There are no T-models that are counter-models to T ($\Box A \supset A$), but there are K-models that are counter-models to T.

Optional Exercise: Find a D-model in which $\Box A \supset A$ is false.

- The tableaux rules for T include all of the rules for K, plus the following:

ρ-rule

$$\begin{array}{c} . \\ \downarrow \\ iri \end{array}$$

for any i on the tree

- Example: $\vdash_{\mathsf{T}} \Box p \supset p$

$$\begin{array}{c} \neg(\Box p \supset p), 0 \checkmark \\ \downarrow \\ 0r0 \\ \downarrow \\ \Box p, 0 \\ \neg p, 0 \\ \downarrow \\ p, 0 \\ \times \end{array}$$

5.2.3 System B ($\mathsf{K}\rho\sigma$)

- A characteristic theorem of B is B:

  B: $A \supset \Box\Diamond A$

- A **B-model** is a model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ whose accessibility relation $\mathcal{R}$ is reflexive and symmetric.

reflexivity	ρ	$w\mathcal{R}w$ for all worlds w (every world can see itself)
symmetry	σ	if $w\mathcal{R}w'$, then $w'\mathcal{R}w$ for all worlds w, w' (if a world can see another world, that world can see the first too)

Definition 5.2.6. An inference is **valid in system B**, $\Sigma \models_{\mathsf{B}} A$, iff every world of every B-model that models the premises also models the conclusion. $\models_{\mathsf{B}} A$, that is, A is **valid in B** iff $\nu_{\mathcal{M},w}(A) = 1$ for every world w of every B-model $\mathcal{M}$.

- B is stronger than K: there are K-models that are not B-models.
- There are no B-models that are counter-models to B $(A \supset \Box\Diamond A)$, but there are K-models that are counter-models to B.

Optional Exercise: Find a B-model in which $\Diamond A \supset \Box\Diamond A$ is false.

- The tableaux rules for T include all of the rules for K, plus the following:

$$\begin{array}{ccc} \rho\text{-rule} & \qquad & \sigma\text{-rule} \\ & & \\ . & & irj \\ \downarrow & & \downarrow \\ iri & & jri \end{array}$$

for any i on the tree

- Example: $\vdash_{\mathsf{B}} p \supset \Box\Diamond p$

$$\begin{array}{c} \neg(p \supset \Box\Diamond p), 0 \surd \\ \downarrow \\ p, 0 \\ \neg\Box\Diamond p, 0 \surd \\ \downarrow \\ \Diamond\neg\Diamond p, 0 \surd \\ \downarrow \\ 0r1 \\ \neg\Diamond p, 1 \surd \\ \downarrow \\ 1r0 \\ \downarrow \\ \Box\neg p, 1 \\ \downarrow \\ \neg p, 0 \\ \times \end{array}$$

5.2.4 System $\mathsf{S4}$ ($\mathsf{K}\rho\tau$)

- A characteristic theorem of $\mathsf{S4}$ is 4:

  4: $\Box A \supset \Box\Box A$

- A **S4-model** is a model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ whose accessibility relation $\mathcal{R}$ is reflexive and transitive.

reflexivity	ρ	$w\mathcal{R}w$ for all worlds w (every world can see itself)
transitivity	τ	if $w_1\mathcal{R}w_2$ and $w_2\mathcal{R}w_3$, then $w_1\mathcal{R}w_3$ for all worlds w_1, w_2, w_3 (if a world can see another, which can see a third, then the first world can see the third)

Definition 5.2.7. An inference is **valid in system S4**, $\Sigma \models_{\mathsf{S4}} A$, iff every world of every S4-model that models the premises also models the conclusion. $\models_{\mathsf{S4}} A$, that is, A is **valid in S4** iff $\nu_{\mathcal{M},w}(A) = 1$ for every world w of every S4-model $\mathcal{M}$.

- S4 is stronger than K: there are K-models that are not S4-models.
- There are no S4-models that are counter-models to 4 ($\Box A \supset \Box\Box A$), but there are K-models that are counter-models to 4.

Optional Exercise: (i) Find a B-model in which $\Box A \supset \Box\Box A$ is false.
(ii) Find a S4-model in which $\Diamond A \supset \Box\Diamond A$ is false.

- The tableaux rules for T include all of the rules for K, plus the following:

$$\begin{array}{ccc} \rho\text{-rule} & & \tau\text{-rule} \\ & & \\ . & & irj \\ \downarrow & & jrk \\ iri & & \downarrow \\ & & irk \\ \text{for any } i \text{ on the tree} & & \end{array}$$

- Example: $\vdash_{\mathsf{S4}} \Box p \supset \Box\Box p$

$$\begin{array}{c} \neg(\Box p \supset \Box\Box p), 0 \surd \\ \downarrow \\ \Box p, 0 \\ \neg\Box\Box p, 0 \surd \\ \downarrow \\ \Diamond\neg\Box p, 0 \surd \\ \downarrow \\ 0r1 \\ \neg\Box p, 1 \surd \\ \downarrow \\ \Diamond\neg p, 1 \surd \\ \downarrow \\ 1r2 \end{array}$$

$$\begin{array}{c} \neg p, 2 \\ 0r2 \\ p, 2 \\ \times \end{array}$$

5.2.5 System S5 (K$\rho\sigma\tau$)

- A characteristic theorem of S5 is 5:

  5: $\Diamond A \supset \Box\Diamond A$

- A **S5-model** is a model $\langle \mathcal{W}, \mathcal{R}, \mathcal{J} \rangle$ whose accessibility relation $\mathcal{R}$ is reflexive, symmetric and transitive.

reflexivity	ρ	$w\mathcal{R}w$ for all worlds w (every world can see itself)
symmetry	σ	if $w\mathcal{R}w'$, then $w'\mathcal{R}w$ for all worlds w, w' (if a world can see another world, that world can see the first too)
transitivity	τ	if $w_1\mathcal{R}w_2$ and $w_2\mathcal{R}w_3$, then $w_1\mathcal{R}w_3$ for all worlds w_1, w_2, w_3 (if a world can see another, which can see a third, then the first world can see the third)

Definition 5.2.8. An inference is **valid in system S5**, $\Sigma \vDash_{\mathsf{S5}} A$, iff every world of every S5-model that models the premises also models the conclusion. $\vDash_{\mathsf{S5}} A$, that is, A is **valid in S5** iff $\nu_{\mathcal{M},w}(A) = 1$ for every world w of every S5-model $\mathcal{M}$.

- S5 is stronger than K: there are K-models that are not S5-models.
- There are no S5-models that are counter-models to 5 ($\Diamond A \supset \Box\Diamond A$), but there are K-models that are counter-models to 5.

Optional Exercise: Find a S5-model in which $\Diamond A \supset \Box A$ is false.

- The tableaux rules for S5 include all of the rules for K, plus the following:

ρ-rule	σ-rule	τ-rule
iri for any i in the tree	$\begin{array}{c} irj \\ \downarrow \\ jri \end{array}$	$\begin{array}{c} irj \\ jrk \\ \downarrow \\ irk \end{array}$

- Example: $\vdash_{\mathsf{S5}} \Diamond p \supset \Box\Diamond p$

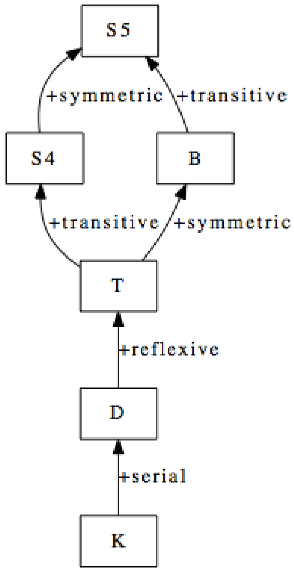
$$\begin{array}{c} \neg(\Diamond p \supset \Box\Diamond p), 0 \checkmark \\ \downarrow \\ \Diamond p, 0 \\ \neg\Box\Diamond p, 0 \checkmark \end{array}$$

$$\begin{array}{c}
\downarrow \\
0r1 \\
p, 1 \\
\downarrow \\
1r0, 0r0, 1r1 \\
\downarrow \\
\Diamond\neg\Diamond p, 0\surd \\
\downarrow \\
0r2 \\
\neg\Diamond p, 2\surd \\
2r0, 2r2, 1r2, 2r1 \\
\downarrow \\
\Box\neg p, 2 \\
\downarrow \\
\neg p, 2 \\
\downarrow \\
\neg p, 1 \\
\downarrow \\
\neg p, 0 \\
\times
\end{array}$$

- Consider a **universal** model, in which $\mathcal{R}$ relates every world $w \in \mathcal{W}$ to every world $w' \in \mathcal{W}$ (unrestricted access across all worlds). A formula is true in all S5-models just in case it is true in all universal models. So with the system S5, we can ignore the accessibility relation in the semantics of '$\Box$' and '$\Diamond$'.

5.3 Summary of the Main Systems of Normal Propositional Modal Logic

Logic		Restrictions on accessibility relation	Characteristic theorem
K		–	$\Box(A \supset B) \supset (\Box A \supset \Box B)$
D	Kη	serialness/extendability	$\Box A \supset \Diamond A$
T	Kρ	reflexivity	$\Box A \supset A$
B	K$\rho\sigma$	reflexivity, symmetry	$A \supset \Box\Diamond A$
S4	K$\rho\tau$	reflexivity, transitivity	$\Box A \supset \Box\Box A$
S5	K$\rho\sigma\tau$	reflexivity, symmetry, transitivity	$\Diamond A \supset \Box\Diamond A$



(Diagram due to John MacFarlane)

5.4 Optional Exercises

1. Show that K is a theorem and a logical truth of K; i.e. show that

   (a) $\vDash_{\mathsf{K}} \Box(A \supset B) \supset (\Box A \supset \Box B)$
   (Hint: Show that K is true at a randomly chosen world of a randomly chosen model.)
   (b) $\vdash_{\mathsf{TK}} \Box(A \supset B) \supset (\Box A \supset \Box B)$
   (Use tableaux)

2. Using tableaux, show that D is not a theorem of K. That is, show that $\nvdash_{\mathsf{TK}} \Box A \supset \Diamond A$.
3. For each of the following arguments

   (a) $\vdash \Box p \supset p$ and
   (b) $\vdash \Box p \supset \Box\Box p$

   does it hold in $\mathsf{K}\rho$, $\mathsf{K}\sigma$, or $\mathsf{K}\tau$? Check with appropriate tableaux and if the tableau does not close, define and draw a counter-model in the usual way.
4. Show the following in $\mathsf{K}\rho$ using tableaux:

   $\vdash (\Diamond\neg A \vee \Diamond\neg B) \vee \Diamond(A \vee B)$
5. Does the following hold in $\mathsf{K}\rho\tau$ using tableaux?

   $\vdash \Diamond\Box p \supset \Box\Diamond p$

5.5 Readings

Obligatory reading: Priest (2008, §§3.1–3.6)

Optional readings: Sider (2010, §§6.1-6.3)

6 Quantified Modal Logic

6.1 Introduction

Previously, we introduced modal operators by adding *possible worlds* to our semantics, letting the truth-values of sentences vary *relative* to each world, and letting the modal operators 'look' across worlds. Interpretations/valuations of propositional logic are based on the truth-values of atomics, so the basic idea behind possible worlds semantics is to allow any atomic sentence to receive different truth-values at different worlds. Predicate logic valuations, on the other hand, work differently: they are based on the meaning assigned to terms and predicates. Once these assignments of meaning are settled, the truth-values of all sentences are settled under that valuation.

We get quantified modal logic, or predicate modal logic, by combining modal logic and predicate logic: we add bells and whistles to our possible worlds semantics that allow the meaning assignments of terms and predicates to vary relative to each world and, as a result, allow the truth-values of sentences to vary relative to each world.

With quantified modal logic, we can — *very, very roughly* — represent natural language sentences such as:

> 'Necessarily, all musicians play an instrument': $\Box\forall x(Px \supset Qx)$
>
> 'Some violias could be cellos': $\Diamond\exists x(Px \wedge Qx)$
>
> 'Paul could have been a drummer and Ringo could have been a singer': $\Diamond Pa \wedge \Diamond Qb$

With the help of quantified modal logic, we can also shed light on a number of philosophical concerns about necessity, possibility, essence, and other modal notions. (More on this below)

6.2 Language

Definition 6.2.1. The vocabulary of modal predicate logic, or quantified modal logic, includes the following symbols.

- individual variables x, y, z, with or without numerical subscripts
- individual constants a, b, c, with or without numerical subscripts
- for every natural number n ¿ 0, n-place predicates P^n, Q^n, S^n, with or without numerical subscripts
- logical operators: $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$, $\Box$, $\Diamond$, $\forall$, $\exists$
- parentheses (,)

Definition 6.2.2. A well-formed formula, or **wff**, of predicate logic is defined inductively:

- if Π is any n-place predicate and $t_1, \ldots, t_n$ are any terms, then $\Pi t_1 \ldots t_n$ is an atomic wff
- if A is a wff, so are $\neg A$, $\Box A$, and $\Diamond A$
- if A and B are wffs, so are $(A \wedge B)$, $(A \vee B)$, $(A \supset B)$, and $(A \equiv B)$
- if A is any wff and α is any variable, then $\forall \alpha A$, $\exists \alpha A$ are wffs
- nothing else is a wff.

6.3 Model Theory

The model theory for quantified modal logic combines the model theory of predicate logic with that of modal logic. (If you're not sure what the model theory of predicate logic looks like, have a look at *Handout II-1*, §2.) We will first look at one way of doing this: **constant domain** semantics, in which there is just one domain of objects that first-order quantifiers quantify over. On the next handout, we will encounter variable domain semantics, in which the domain of first-order quantification can vary from possible world to possible world.

Definition 6.3.1. A **model** for constant domain quantified modal logic is a structure $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ where:

- $\mathcal{W}$ is a non-empty set of objects (intuitively understood as *possible worlds*)
- $\mathcal{D}$ is a non-empty set of objects (the *domain*)
- $\mathcal{R}$ is an accessibility relation between worlds; i.e. $\mathcal{R}$ is a binary relation on $\mathcal{W}$ (so that $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$). We write '$w_1 \mathcal{R} w_2$' for '$w_2$ is accessible from w_1', or 'w_1 sees w_2', which means intuitively that w_2 is possible given/relative to w_1.
- $\mathcal{J}$ is a function (the *interpretation function* such that:
  - if t is an individual constant then $\mathcal{J}(t) \in \mathcal{D}$
  - if Π is an n-place predicate, then $\mathcal{J}(\Pi, w)$ is an n-place relation over $\mathcal{D}$, with $w \in W$.

As before in predicate logic, each constant t is assigned a referent $\mathcal{J}(t)$ in the (constant) domain $\mathcal{D}$. So the reference of constants does not vary with possible worlds; a constant has the same referent in every possible world. Unlike in standard predicate logic, however, we now want the properties of objects to vary across worlds, so for each predicate Π *and* each world w, the predicate is given an *extension* $\mathcal{J}(\Pi, w)$ relative to that world (an extension is still a collection of n-place sequences of objects).

The definitions of variable assignment and denotation are almost identical to those for standard predicate logic (cf. *Handout II-1*, §2). We do not relativize them to worlds, in the same way that we do not relativize the referents of constants to worlds.

Definition 6.3.2. g is a **variable assignment** for model $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ iff g is a function that assigns to each variable some object in $\mathcal{D}$.
$g^{\alpha/d}$ is the variable assignment that is just like g, except that it assigns d to α, where d is some object in $\mathcal{D}$. Note that $g^{\alpha/d}(\alpha) = d$.

Definition 6.3.3. Let $\mathcal{M}$ $(= \langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle)$ be a model, g be a variable assignment, and t be a term. $[t]_{\mathcal{M},g}$, i.e. the **denotation** of t (relative to $\mathcal{M}$ and g), is defined as follows:

$$[t]_{\mathcal{M},g} = \begin{cases} \mathcal{J}(t) & \text{if } t \text{ is a constant} \\ g(t) & \text{if } t \text{ is a variable} \end{cases}$$

Definition 6.3.4. The **valuation** function, $\nu_{\mathcal{M},g,w}$, for model $\mathcal{M}$ $(= \langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle)$, variable assignment g, and world w is defined as the function that assigns either 0 or 1 to each wff relative to each world $w \in W$, subject to the following constraints:

(i) for any n-place predicate Π and any terms $t_1 \dots t_n$, $\nu_{\mathcal{M},g,w}(\Pi t_1 \dots t_n) = 1$ iff $\langle [t_1]_{\mathcal{M},g} \dots [t_n]_{\mathcal{M},g} \rangle \in \mathcal{J}(\Pi, w)$.

(ii) For any wffs A, B and any variable α:

$$\begin{array}{rcl}
\nu_{\mathcal{M},g,w}(\neg A) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 0 \\
\nu_{\mathcal{M},g,w}(A \wedge B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 1 \text{ and } \nu_{\mathcal{M},g,w}(B) = 1 \\
\nu_{\mathcal{M},g,w}(A \vee B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 1 \text{ or } \nu_{\mathcal{M},g,w}(B) = 1 \\
\nu_{\mathcal{M},g,w}(A \supset B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 0 \text{ or } \nu_{\mathcal{M},g,w}(B) = 1 \\
\nu_{\mathcal{M},g,w}(A \equiv B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = \nu_{\mathcal{M},g,w}(B) \\
\nu_{\mathcal{M},g,w}(\forall \alpha A) = 1 & \text{iff} & \text{for every } d \in \mathcal{D},\ \nu_{\mathcal{M},g^{\alpha/d},w}(A) = 1 \\
\nu_{\mathcal{M},g,w}(\exists \alpha A) = 1 & \text{iff} & \text{for at least one } d \in \mathcal{D},\ \nu_{\mathcal{M},g^{\alpha/d},w}(A) = 1 \\
\nu_{\mathcal{M},g,w}(\Box A) = 1 & \text{iff} & \nu_{\mathcal{M},g,x}(A) = 1 \text{ at all worlds } x \text{ such that } w\mathcal{R}x \\
\nu_{\mathcal{M},g,w}(\Diamond A) = 1 & \text{iff} & \nu_{\mathcal{M},g,x}(A) = 1 \text{ at some world } x \text{ such that } w\mathcal{R}x
\end{array}$$

Read '$\nu_{\mathcal{M},g,w}(A) = 1$' as *A is true in (model)* $\mathcal{M}$ *relative to (variable assignment) g and (possible world) w*. We will also write '$\nu_{\mathcal{M},g}(A, w) = 1$' to mean the same thing.

Definition 6.3.5. Because we are currently dealing with constant domain semantics, we call our basic quantified modal logic CK or 'Constant Domain K'.

We say that a world w of model $\langle W, R, D, \mathcal{J}\rangle$ **models** sentence A iff $\nu_{\mathcal{M},g,w}(A) = 1$.

An argument is **valid in CK** iff every world of every model that models the premises also models the conclusion. We write '$\vDash_{\mathsf{CK}}$' for this semantic consequence relation, and we define this notion precisely as follows.

$$\begin{gathered}\Sigma \vDash_{\mathsf{CK}} A \;\text{ iff }\; \text{for all worlds } w \in \\ W \text{ of all models } \langle W, R, D, \mathcal{J}\rangle \text{ and all variable assignments } g \text{ for } \mathcal{M}\text{:} \\ \text{if } \nu_{\mathcal{M},g,w}(B) = 1 \text{ for all the premises } B \in \Sigma, \text{ then } \nu_{\mathcal{M},g,w}(A) = 1\end{gathered}$$

$\vDash_{\mathsf{CK}} A$, that is, A is **valid in CK** iff $\nu_{\mathcal{M},g,w}(A) = 1$ for every world w of every model $\mathcal{M}$ and every variable assignment g for $\mathcal{M}$.

Model $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J}\rangle$ with world w gives a **counter-model** to the inference from Σ to A if $\nu_{\mathcal{M},g,w}(B) = 1$ for all $B \in \Sigma$ but $\nu_{\mathcal{M},g,w}(A) = 0$. This makes the inference *invalid*, $\nvDash_{\mathsf{CK}} A$.

Remark. In propositional modal logic, truth-functional connectives are interpreted 'locally' in the sense that the truth-value at w of a sentence whose main operator is a truth-functional connective is fully determined by the truth-values at w of its parts. In quantified modal logic the quantifiers are also interpreted 'locally' in the sense that what determines the truth-value of a formula like '$\forall x Px$' at world w is just the properties of objects at w.

6.4 Additions to Modal Tableaux

∀-rule	**¬∀-rule**	**∃-rule**	**¬∃-rule**
$\forall \alpha A, i$ $\downarrow$ $A(\alpha/\beta), i$	$\neg\forall \alpha A, i$ ✓ $\downarrow$ $\exists \alpha \neg A, i$	$\exists \alpha A, i$ ✓ $\downarrow$ $A(\alpha/\beta), i$	$\neg\exists \alpha A, i$ ✓ $\downarrow$ $\forall \alpha \neg A, i$
for *any* constant β already on the branch, or else introduce a new one. Never check off.		where β is a *new* constant not yet on the branch	

Definition 6.4.1. A branch of a tree for quantified modal logic **closes** if it contains a wff and its negation with the same *world-index* (i.e., both A, k and $\neg A, k$) The tree **closes** if every branch does.

Definition 6.4.2. We say that there is a modal tableaux **proof** from Σ to A, written $\Sigma \vdash_{\mathsf{CK}} A$, just in case the tree whose starting list includes $B, 0$ for each $B \in \Sigma$ as well as $\neg A, 0$ closes.

Definition 6.4.3. We say that there is A is a **theorem of CK**, written $\vdash_{\mathsf{CK}} A$, just if the tree whose starting list includes $\neg A, 0$ closes.

6.5 The Barcan Formula

What we have done above is *fuse together* two semantic frameworks—a domain of objects taken over from predicate logic, with a set of worlds taken over from modal logic—and furthermore, we have combined them in the simplest way imaginable. Let's consider one of the most infamous ramifications of fusing things together in this way: the Barcan formula (BF), $\forall x \Box Px \supset \Box \forall x Px$. This formula, named after Ruth Barcan Marcus, is a theorem and logical truth in CK as seen below.

$$
\begin{array}{cl}
\neg(\forall x \Box Px \supset \Box \forall x Px), 0 & \checkmark \\
\downarrow & \\
\forall x \Box Px, 0 & /\mathrm{a} \\
\neg \Box \forall x Px, 0 & \checkmark \\
\downarrow & \\
\Diamond \neg \forall x Px, 0 & \checkmark \\
\downarrow & \\
0r1 & \\
\neg \forall x Px, 1 & \checkmark \\
\downarrow & \\
\exists x \neg Px, 1 & \checkmark \\
\downarrow & \\
\neg Pa, 1 & \\
\downarrow & \\
\Box Pa, 0 & /1 \\
\downarrow & \\
Pa, 1 & \\
\times &
\end{array}
$$

It is highly contentious whether this is a correct principle of quantified modal logic. If you think about it, our search for a counterexample only fails because we assume that the object a which falsifies the consequent also falls within the range of the quantifier in the antecedent. This means we are implicitly assuming that *all possible individuals are actual*. Indeed, that is roughly what BF says: if everything must be P, then necessarily everything is P. The kind of semantics we are using is called constant domain semantics precisely because of the fact that the domain of objects does not change between worlds (just the properties of objects can vary). It seems intuitive that the consequent of BF could be false even when the antecedent is true, so long as there are worlds where some things exist that do not actually exist. Our semantic framework doesn't allow this, but if we agree that *which objects exist* should be able to vary between possible worlds, then we should agree that BF is false.

Consider the formula $\Diamond\exists xPx \supset \exists x\Diamond Px$. You can test it to see that it is also a theorem and logical truth, and for the same reasons; it is, in fact, equivalent to BF. The 'diamond' version makes the criticism especially clear. An instance of this formula, for example, is the claim that if it is possible that someone jumps nine meters there is someone who can jump nine meters. Or for another instance, if it is possible that there be a Leprechaun, then something exists which could have been a Leprechaun. To many people these sound very counter-intuitive.

Although this discussion may make it sound obvious how to adjust the semantics to avoid validating these formulae, in practice it is complicated. Next week we will discuss an alternative approach called **'variable domain' semantics**, which is an attempt to make good on our intuition that which objects exist may vary between possible worlds.

Here is another example in CK, this time an invalid inference.

$\forall x\Box(Px \supset Qx) \nvdash_{\mathsf{CK}} \forall x(Px \supset \Box Qx)$.

$$
\begin{array}{ccc}
 & \forall x\Box(Px \supset Qx), 0 & /a \\
 & \neg\forall x(Px \supset \Box Qx), 0 & \checkmark \\
 & \downarrow & \\
 & \exists x\neg(Px \supset \Box Qx), 0 & \checkmark \\
 & \downarrow & \\
 & \neg(Pa \supset \Box Qa), 0 & \checkmark \\
 & \downarrow & \\
 & Pa, 0 & \\
 & \neg\Box Qa, 0 & \checkmark \\
 & \downarrow & \\
 & \Diamond\neg Qa, 0 & \checkmark \\
 & \downarrow & \\
 & 0r1 & \\
 & \neg Qa, 1 & \\
 & \Box(Pa \supset Qa), 0 & /1 \\
 & \downarrow & \\
 & Pa \supset Qa, 1 & \checkmark \\
\swarrow & & \searrow \\
\neg Pa, 1 & & Qa, 1 \\
 & & \times
\end{array}
$$

Counter-model: $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$ such that

$\mathcal{W} = \{w_0, w_1\}$
$\mathcal{D} = \{\delta_a\}$
$\mathcal{R} = \{\langle w_0, w_1\rangle\}$
$\mathcal{I}(a) = \delta_a, \quad \mathcal{I}(P, w_0) = \{\delta_a\}, \quad \mathcal{I}(P, w_1) = \varnothing$ and $\mathcal{I}(Q, w_1) = \varnothing$

6.6 Optional Exercises

1. Test using tableaux. Construct a counter-model if the tree is open.
    1. $\forall x\Box(Qx \vee Hx) \vdash \Box\forall x Hx$
    2. $\exists x\Box(Px \wedge Qx) \vdash \Box\exists x Px$
    3. $\Box\Diamond\exists x Px \vdash \Box\exists x\Diamond(Px \vee Qx)$
2. Consider the inference above: $\forall x\Box(Px \supset Qx) \vdash \forall x(Px \supset \Box Qx)$

    What happens if we add the ρ constraint? Test this using a tree with the ρ rule (cf. *Handout III-2*, §2.2). Does this have any impact on the result? Does the inference come out valid in this system?

6.7 Readings

Obligatory reading: Priest (2008, ch .14)

Optional readings: Sider (2010, §§9.1-9.5); on some philosophical issues: Lowe (2002, pp. 79–84)

7 Quantified Modal Logic: Variable Domains

7.1 Constant Domain Quantified Modal Logic

7.1.1 Review

- **Model** for constant domain Quantified Modal Logic: $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$
- This is a very simple way of fusing predicate and modal logic semantics:

  A model has a **single** domain of objects $\mathcal{D}$ (hence 'constant domain' semantics)

  - Quantifiers quantify over the same domain of objects, no matter at which world we evaluate a quantified wff for truth
  - If '$\exists x(x = a)$' is true (false) at some world, it is true (false) at *every* world.
  - Whatever exists at this world exists at every other world.
  - Whatever exists at any world exists at this (the actual) world.

7.1.2 Challenges to Constant Domain Quantified Modal Logic

The Barcan Formula

- The Barcan Formula (BF) is a logical truth of CK: $\vDash_{\mathsf{CK}} \forall x \Box Px \supset \Box \forall x Px$
- Many find it objectionable; an example for materialists:

  Everything is necessarily material $\not\Rightarrow$ Necessarily, everything is material.

  $\forall x \Box Mx$ $\qquad\qquad$ $\Box \forall x Mx$

  Everything in our world may be such that it cannot fail to be material. Yet it seems possible that *other things* existed which are immaterial.

Necessary Existence

- With a single domain of objects, every object exists at every world. Thus, it exists **necessarily** (it cannot fail to exist).

- But it seems that existence is **contingent**: I could not have come into existence, the laptop I'm writing on may not have been manufactured, the cow whose meat you've had for dinner last night could have never been conceived, etc.
- But even if it's a hard metaphysical question whether existence is necessary or contingent: Should its answer be a matter of logical truth?
- On constant domain semantics, it is:

  $\vDash_{\mathsf{CK}} \forall x \Box \exists y (y = x)$
  (where '=' is a predicate that receives the same interpretation in *every* model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{J} \rangle$: $\mathcal{J}(=, w) = \{\langle d, d \rangle : d \in \mathcal{D}\}$)

  $\Rightarrow$ '*Everything necessarily exists*' is a matter of logical truth.

A Dilemma: What should be in the single domain $\mathcal{D}$?

1. **Parsimonious**: $\mathcal{D}$ contains only *really* existing things – concrete/spatio-temporal//... objects in the (actual) universe (people, tables, chairs, planets, electrons, ...)

   We take $\forall$ and $\exists$ to have **strong ontological import**: whatever is in the domain $\mathcal{D}$ can truly be said to *really* exist (not 'exist' in some derivative way).

   If I want to say that there (really!) are cyclones, I do this as follows: $\exists x Cx$

   Issue 1: $\vDash_{\mathsf{CK}} \forall x \Box \exists y (y = x)$

   All (concretely/physically/...) things exist necessarily. $\lightning$

   Issue 2: If we only allow into the domain the objects that *actually* exist (exist at the actual world), then we cannot *truly say* that there could be things that do not (actually) exist: $\Diamond \exists x (x = a) \wedge \neg \exists x (x = a)$

   $\Rightarrow$ For any a, either it is in the domain (then it also exists at the actual world), or it is not (then it doesn't exist at any world).

   But couldn't there be things that do not actually exist? Couldn't there be a son of L. Wittgenstein?

   That is: Someone who doesn't actually exist but *could have* existed? (Note that this is different from someone having the property of possibly being Wittgenstein's son)

2. **Prodigal**: In addition to concretely/physically/fundamentally existing things, $\mathcal{D}$ contains things that exist in some derivative/extraordinary/abstract way (ghosts, golden mountains, talking donkeys, Wittgenstein's son, ...)

   We do not take $\forall$ and $\exists$ to have strong ontological import: objects in the domain include things we do not call 'existing' in the strong (concrete) way.

   Then we can throw anything in our single domain that we want to say exists at *some* world or other: our domain contains Wittgenstein's son as well as ghosts.

   $\vDash_{\mathsf{CK}} \forall x \Box \exists y (y = x)$ is not so bad: everything exists necessarily **in a concrete or abstract way**.

Issue 1: Hefty metaphysical commitments! Many philosophers consider the postulation of such extraordinary things to be obviously false, conceptually incoherent, or worse (cf. the popular criticism of Meinong's 'non-existent objects').

Issue 2: Not clear how we can say that something *really* exists (at the actual world) (rather than just in this extraordinary way which includes ghosts): '$\exists x(x = a)$' says only that a exists *in a concrete* ***or abstract way***

Ways out of the dilemma? (cf. Sider (2010, §9.5); Williamson (1998))

7.1.3 In Defence of Constant Domain Quantified Modal Logic

1. Embrace the first horn (parsimony): Logic is a more reliable guide to (modal) metaphysics than our intuitions about what exists. (Bite the bullet on counter-intuitive commitments.)
2. Embrace the second horn (prodigality): one big domain of concrete and merely possible things.
   - Express concrete existence by adding a **(concrete) existence predicate,** E, which at every possible world applies to all and only those things that exist concretely at that world.
   - 'a (concretely) exists': $\exists x(Ex \wedge x = a)$
   - 'Everything exists necessarily (concretely)' is not a logical truth: $\nvDash_{\mathsf{CK'}} \forall x \Box \exists y(Ey \wedge y = x)$

   Objections to 2:

   1. Hefty metaphysical commitments remain
   2. Rendering English into logic is less direct: E must be added to all translations of sentences expressing concrete existence.
   3. The proper role of quantifiers is to record robust ontological commitment (Quine).

7.2 Variable Domain Quantified Modal Logic: Model Theory

- Another way out of the dilemma: Give up the commitment to a single constant domain
- Make quantification **world-relative** . . .
- . . . by introducing **variable domains** into our semantics: a domain for every possible world.
- $\forall$ and $\exists$ quantify over different domains, depending on the world at which they are being evaluated for truth/falsity.

Definition 7.2.1. A **model** for variable domain quantified modal logic is a structure $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathscr{Q}, \mathcal{J} \rangle$ where:

- $\mathcal{W}$ is a non-empty set of objects (intuitively understood as *possible worlds*)
- $\mathcal{R}$ is an accessibility relation between worlds; i.e. $\mathcal{R}$ is a binary relation on $\mathcal{W}$ (so that $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$). We write '$w_1 \mathcal{R} w_2$' for '$w_2$ is accessible from w_1', or 'w_1 sees w_2', which means intuitively that w_2 is possible given/relative to w_1.
- $\mathcal{D}$ is a non-empty set of objects (the *super-domain*)
- $\mathscr{Q}$ is a function that assigns to any $w \in \mathcal{W}$ a subset of $\mathcal{W}$. We will refer to $\mathscr{Q}(w)$ as '$\mathcal{D}_w$'. Think of $\mathcal{D}_w$ as w's *sub-domain* — the set of objects that exist at w.
- $\mathcal{J}$ is a function (the *interpretation function*) such that:
  - if t is an individual constant then $\mathcal{J}(t) \in \mathcal{D}$
  - if Π is an n-place predicate, then $\mathcal{J}(\Pi, w)$ is an n-place relation over $\mathcal{D}$, with $w \in W$.

Note that $\mathcal{D}$ (the super-domain) still includes *all* possible individuals. But quantifiers $\forall$ and $\exists$ range over subsets of $\mathcal{D}$ – one for each world w: $\mathcal{D}_w$. When we evaluate a quantified sentence such as '$\forall x P x'$ for truth/falsity, we do so at a world w – **and we interpret** $\forall x$ **as ranging over** w**'s subdomain,** $\mathcal{D}_w$.

Definition 7.2.2. g is a **variable assignment** for model $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathscr{Q}, \mathcal{J} \rangle$ iff g is a function that assigns to each variable some object in $\mathcal{D}$.
$g^{\alpha/d}$ is the variable assignment that is just like g, except that it assigns d to α, where d is some object in $\mathcal{D}$. Note that $g^{\alpha/d}(\alpha) = d$.

Definition 7.2.3. Let $\mathcal{M}$ $(= \langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathscr{Q}, \mathcal{J} \rangle)$ be a model, g be a variable assignment, and t be a term. $[t]_{\mathcal{M},g}$, i.e. the **denotation** of t (relative to $\mathcal{M}$ and g), is defined as follows:

$$[t]_{\mathcal{M},g} = \begin{cases} \mathcal{J}(t) & \text{if } t \text{ is a constant} \\ g(t) & \text{if } t \text{ is a variable} \end{cases}$$

Definition 7.2.4. The **valuation** function, $\nu_{\mathcal{M},g,w}$, for model $\mathcal{M}$ $(= \langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathscr{Q}, \mathcal{J} \rangle)$, variable assignment g, and world w is defined as the function that assigns either 0 or 1 to each wff relative to each world $w \in W$, subject to the following constraints:

(i) for any n-place predicate Π and any terms $t_1 \ldots t_n$, $\nu_{\mathcal{M},g,w}(\Pi t_1 \ldots t_n) = 1$ iff $\langle [t_1]_{\mathcal{M},g} \ldots [t_n]_{\mathcal{M},g} \rangle \in \mathcal{J}(\Pi, w)$.

(ii) For any wffs A, B and any variable α:

$$\begin{array}{rcl}
\nu_{\mathcal{M},g,w}(\neg A) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 0 \\
\nu_{\mathcal{M},g,w}(A \wedge B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 1 \text{ and } \nu_{\mathcal{M},g,w}(B) = 1 \\
\nu_{\mathcal{M},g,w}(A \vee B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 1 \text{ or } \nu_{\mathcal{M},g,w}(B) = 1 \\
\nu_{\mathcal{M},g,w}(A \supset B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = 0 \text{ or } \nu_{\mathcal{M},g,w}(B) = 1 \\
\nu_{\mathcal{M},g,w}(A \equiv B) = 1 & \text{iff} & \nu_{\mathcal{M},g,w}(A) = \nu_{\mathcal{M},g,w}(B) \\
\nu_{\mathcal{M},g,w}(\forall \alpha A) = 1 & \text{iff} & \text{for every } d \in \mathcal{D}_w,\ \nu_{\mathcal{M},g^{\alpha/d},w}(A) = 1 \\
\nu_{\mathcal{M},g,w}(\exists \alpha A) = 1 & \text{iff} & \text{for at least one } d \in \mathcal{D}_w,\ \nu_{\mathcal{M},g^{\alpha/d},w}(A) = 1 \\
\nu_{\mathcal{M},g,w}(\Box A) = 1 & \text{iff} & \nu_{\mathcal{M},g,x}(A) = 1 \text{ at all worlds } x \text{ such that } w\mathcal{R}x \\
\nu_{\mathcal{M},g,w}(\Diamond A) = 1 & \text{iff} & \nu_{\mathcal{M},g,x}(A) = 1 \text{ at some world } x \text{ such that } w\mathcal{R}x
\end{array}$$

Read '$\nu_{\mathcal{M},g,w}(A) = 1$' as A *is true in (model)* $\mathcal{M}$ *relative to (variable assignment)* g *and (possible world)* w. We will also write '$\nu_{\mathcal{M},g}(A, w) = 1$' to mean the same thing.

Definition 7.2.5. We call our basic quantified modal logic **VK** or '**Variable Domain K**'.

We say that a world w of model $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{Q}, \mathcal{J} \rangle$ **models** sentence A iff $\nu_{\mathcal{M},g,w}(A) = 1$.

An argument is **valid in VK** iff every world of every model that models the premises also models the conclusion. We write '$\vDash_{\mathsf{VK}}$' for this semantic consequence relation, and we define this notion precisely as follows.

$$\begin{gathered}
\Sigma \vDash_{\mathsf{VK}} A \ \text{ iff } \ \text{for all worlds } w \in \\
\mathcal{W} \text{ of all models } \langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{Q}, \mathcal{J} \rangle \text{ and all variable assignments } g \text{ for } \mathcal{M}\text{:} \\
\text{if } \nu_{\mathcal{M},g,w}(B) = 1 \text{ for all the premises } B \in \Sigma, \text{ then } \nu_{\mathcal{M},g,w}(A) = 1
\end{gathered}$$

$\vDash_{\mathsf{VK}} A$, that is, A is **valid in VK** iff $\nu_{\mathcal{M},g,w}(A) = 1$ for every world w of every model $\mathcal{M}$ and every variable assignment g for $\mathcal{M}$.

Model $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathcal{Q}, \mathcal{J} \rangle$ with world w gives a **counter-model** to the inference from Σ to A if $\nu_{\mathcal{M},g,w}(B) = 1$ for all $B \in \Sigma$ but $\nu_{\mathcal{M},g,w}(A) = 0$. This makes the inference *invalid*, $\nvDash_{\mathsf{VK}} A$.

We get stronger systems of variable domain quantified modal logic by adding constraints on the accessibility relation $\mathcal{R}$ — in the same way you do with normal propositional modal logics (e.g. D, T, B, S4, S5) – and defining validity by means of models whose $\mathcal{R}$ satisfy the constraint (D-models, T-models, . . .). Cf. *Handout III-2.*

7.3 Solving CK's Issues

- What exists at the actual world @ need not exist at any other world: $\mathcal{D}_@$ need not be a subset of any other w's domain $\mathcal{D}_w$.
- What exists at some other world w need not exist at the actual world @: $\mathcal{D}_@$ need not be a superset of any other w's domain $\mathcal{D}_w$.

- BF is not valid in VK: $\not\models_{\mathsf{VK}} \forall x \Box Px \supset \Box \forall x Px$

  Take a random world w of model $\mathcal{M}$:

  - let the antecedent, '$\forall x \Box Px$', be true at w (given $\mathcal{M}$ and g): all $d \in \mathcal{D}_w$ satisfy P at every world that w accesses.
  - the consequent, '$\Box \forall x Px$', can still be false: there may be a world w' (accessible from w) **whose domain $\mathcal{D}_{w'}$ contains at least one object d' not in $\mathcal{D}_w$, which does not satisfy** P, so '$\forall x Px$' is false at w' and thus '$\Box \forall x Px$' is false.

- No necessary existence: $\not\models_{\mathsf{VK}} \forall x \Box \exists y (y = x)$

  The formula is false at a world w if w accesses some world w' **whose domain $\mathcal{D}_{w'}$ fails to contain at least one object d' that $\mathcal{D}_w$ contains**.

7.4 Tableaux

7.4.1 A Complication

- Consider the tableaux rule for **Universal Instantiation**:

$\forall$-rule

$$\forall \alpha A, i$$
$$\downarrow$$
$$A(\alpha/\beta), i$$

for **any** constant β already on the branch, or else introduce a new one. Never check off.

- But what if β does not exist at world w_i (i.e. $\beta \notin \mathcal{D}_{w_i}$)?

  For instance, it seems wrong to derive from the truth of '$\forall x Px$' at w the truth at w of 'Pa' for some a that doesn't exist at w.
- Note: We cannot just write into the $\forall$-rule that β must be in $\mathcal{D}_{w_i}$: Our proof theory (tableaux rules) must not depend on *semantic* notions such as '$\mathcal{D}$'.
- One way to solve the complication: **free logic**
- Free logic was developed (inter alia) to allow our language to contain *empty names* such as 'Sherlock Homes', 'Pegasus', etc.: names that do not have a referent. We can think of our complication along the same lines: Some individual constants have as referents objects (from $\mathcal{D}$) that do not exist at a given world w (are not in $\mathcal{D}_w$ — they do not have a referent at w.
- We introduce an **existence predicate** $\mathscr{E}$ into our language. It is defined as follows:

  $\mathcal{J}(\mathscr{E}, w) = \mathcal{D}_w$
- Next, we introduce free logic tableaux rules for $\forall$ and $\exists$ (modalized versions):

7.4.2 Tableaux Rules for VK

Replace CK's rules for $\forall$ and $\exists$ by the following two:

$\forall$-rule

$$\begin{array}{ccc} & \forall\alpha A, i & \\ \swarrow & & \searrow \\ \neg\mathscr{E}\beta, i & & A(\alpha/\beta), i \end{array}$$

for *any* constant β already on the branch, or else introduce a new one. Never check off.

$\exists$-rule

$$\begin{array}{c} \exists\alpha A, i \ \checkmark \\ \downarrow \\ \mathscr{E}\beta, i \\ A(\alpha/\beta), i \end{array}$$

where β is a *new* constant not yet on the branch

Definition 7.4.1. A branch of a tree for quantified modal logic **closes** if it contains a wff and its negation with the same *world-index* (i.e., both A, k and $\neg A, k$) The tree **closes** if every branch does.

Definition 7.4.2. We say that there is a modal tableaux **proof** from Σ to A, written $\Sigma \vdash_{\mathsf{VK}} A$, just in case the tree whose starting list includes $B, 0$ for each $B \in \Sigma$ as well as $\neg A, 0$ closes.

Definition 7.4.3. We say that there is A is a **theorem of CK**, written $\vdash_{\mathsf{VK}} A$, just if the tree whose starting list includes $\neg A, 0$ closes.

7.4.3 An Example: BF

$\nvdash_{\mathsf{VK}} \forall x \Box Px \supset \Box \forall x Px$

$$
\begin{array}{ccc}
\neg(\forall x\Box Px \supset \Box\forall x Px), 0 & & \checkmark \\
\downarrow & & \\
\forall x\Box Px, 0 & & /a \\
\neg\Box\forall x Px, 0 & & \checkmark \\
\downarrow & & \\
\Diamond\neg\forall x Px, 0 & & \checkmark \\
\downarrow & & \\
0r1 & & \\
\neg\forall x Px, 1 & & \checkmark \\
\downarrow & & \\
\exists x\neg Px, 1 & & \checkmark \\
\downarrow & & \\
\mathscr{E}a, 1 & & \\
\neg Pa, 1 & & \\
\swarrow \quad \searrow & & \\
\neg\mathscr{E}a, 0 \qquad \Box Pa, 0 & & /1 \\
\qquad\qquad \downarrow & & \\
\qquad\qquad Pa, 1 & & \\
\qquad\qquad \times & &
\end{array}
$$

Counter-model read off from left branch: $\langle \mathcal{W}, \mathcal{R}, \mathcal{D}, \mathscr{Q}, \mathscr{I} \rangle$ such that

$\mathcal{W} = \{w_0, w_1\}$
$\mathcal{R} = \{\langle w_0, w_1 \rangle\}$
$\mathcal{D} = \{\delta_a\}$
$\mathscr{Q}(w_0) = \mathcal{D}_{w_0} = \mathscr{I}(\mathscr{E}, w_0) = \varnothing, \quad \mathscr{Q}(w_1) = \mathcal{D}_{w_1} = \mathscr{I}(\mathscr{E}, w_1) = \{\delta_a\}$
$\mathscr{I}(a) = \delta_a, \quad \mathscr{I}(P, w_1) = \varnothing, \quad \mathscr{I}(P, w_0)$ doesn't matter

7.5 Optional Exercises

1. Test the following inference using tableaux. If the tree does not close, use an open branch to define a counter-model.

   (a) $\vdash_{\mathsf{VK}} (\Box\forall x Px \wedge \Box\forall x Qx) \supset \Box\forall x(Px \wedge Qx)$
   (b) $\vdash_{\mathsf{VK}} \Diamond\exists Px \supset \Diamond\exists x(Px \vee Qx)$
   (c) $\vdash_{\mathsf{VK}} \Box\exists x Px \supset \exists x\Box Px$

2. Can we define the logic VK (understood as the set of premises-conclusion pairs $\langle \Sigma, A \rangle$ such that $\Sigma \vdash_{\mathsf{VK}} A$) by using constant domain QML plus the existence predicate? How would we do this? Are there any philosophical costs to this maneuver? (Cf. Priest (2008, §15.8))

7.6 Readings

Obligatory reading: Priest (2008, ch .15)

Optional readings:

- Priest (2008, §§13.1-13.6)
- Sider (2010, §9.6)
- Garson, James (2013.'Modal Logic'. *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = ¡http://plato.stanford.edu/archives/spr2013/entries/logic-modal/¿

Part III

Conditionals

8 Material and Strict Implication

8.1 What Are Conditionals?

Priest (2008, 11) — a characterization from philosophical logic:

> 'Conditionals relate some proposition (the *consequent*) to some other proposition (the *antecedent*) on which, in some sense, it depends.'

von Fintel (2011, 2) — a characterization from linguistic semantics:

> 'Conditionals are sentences that talk about a possible scenario that may or may not be actual and describe what (else) is the case in that scenario; or, considered from "the other end", conditionals state in what kind of possible scenarios a given proposition is true. The canonical form of a conditional is a two-part sentence consisting of an "antecedent" (also: "premise", "protasis") [in English] marked with *if* and a "consequent" ("apodosis") sometimes marked with *then*...'

8.1.1 Conditionals in Natural Language (English)

Conditionals can be expressed in different ways in English. The 'canonical' form of a conditional is the one in (8.1), but see also (8.2)-(8.7)

(8.1) If Grijpstra played his drum, de Gier played his flute.

(8.2) Had he admitted his guilt, he would have gotten off easier.

(8.3) Take another step and I'll knock you down.

(8.4) You won't eat those and live.

(8.5) He was pushed or he wouldn't have fallen down the cliff.

(8.6) Without you, I would be lost.

(8.7) *I* would have beaten Kasparov.

Nota bene: The English word *if* can occur in idiomatic expressions that despite their appearance are not conditionals; e.g. in 'If I may say so, you have a nice ear-ring.'

The grammar of conditionals in English imposes certain requirements on the tense (past, present, future) and mood (indicative, subjunctive) of the sentences expressing them. A textbook of English grammar will tell you this:[1]

Type	**If clause**	**Main clause**
I	Simple present	*will* + infinitive (future), or modal, or sometimes simple present
II	Simple past	*would* + infinitive (*could/may/might/should/must*)
III	Past perfect	*would* + *have* + past participle (*could/may/might/should/must*)

8.1.2 Types Of Conditionals

We can also distinguish between different types of conditionals.

1. It is common to distinguish between **indicative conditionals** like (10.1a) and **counterfactual** or **subjunctive conditionals** like (10.1b).

   (8.8) a. If Oswald didn't kill Kennedy, someone else did.
   b. If Oswald hadn't killed Kennedy, someone else would have. (Adams, 1970)

   There are both differences and similarities between indicative and subjunctive conditionals.

   – Indicative conditionals convey that the truth of the antecedent is an open issue whereas subjunctive conditionals (typically, but not invariably) convey that the antecedent is false (hence the term 'counterfactuals').
   – The meaning of indicative conditionals like (10.1a) (unlike subjunctive conditionals) seems to correspond to the so-called **"Ramsey Test"** (Ramsey, 1931):

     > If two people are arguing "If p will q?" and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q.

   – However, indicative and subjunctive conditionals exhibit similar inference patterns; e.g. failure of **strengthening the antecedent**:

[1] If you would like to brush up your knowledge of the grammar of English conditionals, try one of these sources:
http://learnenglish.britishcouncil.org/en/grammar-reference/conditionals-1
http://www.englisch-hilfen.de/en/grammar/if.htm

$$A \to B \not\Rightarrow (A \wedge C) \to B)$$

(Let's agree on a useful notational convention and say that '$\to$' stands for the conditional, independently (i) of the natural language expressions used to express it (e.g. *if*) and/or (ii) of the logical connective used to express the conditional.)

– Here are examples of the failure of antecedent strengthening:

(8.9) a. If I light this match, it will burn.
b. $\not\Rightarrow$ If I light this match and it is wet, it will burn.

(8.10) a. If Piedro had come to the parade, he would have seen the dancers.
b. $\not\Rightarrow$ If Piedro had come to the parade and got stuck behind a tall person, he would have seen the dancers.

2. There are more types of conditionals than indicative/subjunctive conditionals. For instance, there are **biscuit conditionals** (or '**speech act conditionals**') like (8.11) and (8.12.

(8.11) There are biscuits on the sideboard if you want them. (Austin)
(8.12) I paid you back yesterday, if you remember. (P.T. Geach p.c. to Austin)

Biscuit conditionals do not state conditions under which the consequent is true. Instead, they seem to operate on a speech act level.

3. There are also **premise conditionals** like (8.13) that often echo someone else's introduction of the antecedent.

(8.13) If you're so clever, why don't you do this problem on your own.

- Furthermore, besides conditional statements, there are conditional commands, promises, offers, questions, etc.:

(8.14) If it gets cold, close the window and turn on the heating.
(8.15) I promise to buy you a car if you pass the driving test this time.
(8.16) What should I do if she doesn't say 'Yes'?

- Finally, note that there is still controversy over how to best classify conditionals. For instance, some theorists doubt that the indicative/subjunctive distinction is tracking a real semantic distinction (see e.g. Gibbard (1981, 222-6)).

8.1.3 Contrapositive, Converse, Inverse

- The **contrapositive** of $A \to B$ is $\neg B \to \neg A$
- N.B.: $\models_C A \supset B \equiv \neg B \supset \neg A$.
- The **converse** of $A \to B$ is $B \to A$.

- The **inverse** of $A \to B$ is $\neg A \to \neg B$.

Sometimes, we also use these terms when $\vdash$ or $\models$ are at issue (rather than the conditional $\to$). Thus, the contrapositive of $A \vdash B$ is $\neg B \vdash \neg A$.

8.2 Material Implication

- Think semantically in truth-functional terms. Which function from two truth-values (for antecedent and consequent could express the conditional?
- The best candidate seems to be material implication. Here is the familiar truth table that gives the meaning of '$\supset$' (the 'horseshoe', or 'hook', symbol stands for material implication):

$\supset$	1	0
1	1	0
0	1	1

8.2.1 Arguments In Favour Of Material Implication

1. Given that 'if... then' is a truth-functional binary connective, '$\supset$' is the most plausible candidate.

   > 'A professor who declares that *If I am healthy, I will come to class* can only be said to have broken her promise if she is healthy but doesnt come to class. Clearly, if she is healthy and comes to class, shell have spoken the truth. And if she is sick, it is immaterial whether she comes to class (going beyond the call of duty and beyond what she promised) or doesnt neither case constitutes a breaking of the promise.' (von Fintel, 2011, §3.1; example attributed to Suber)

2. A quick argument:

   If A then B = Either not-A or B

   $$= \neg A \vee B$$

   $$= A \supset B$$

   (Cf. Priest (2008, §§1.10.2-1.10.5) for a longer version of this argument)

3. The Direct Argument (cf. Stalnaker (1975)):

   "Either the butler or the gardener did it. Therefore, if the butler didn't do it, the gardener did."

(i) '$A \vee B$' entails 'If $\neg A$, then B'. Substitute '$\neg A$' for 'A'. '$\neg A \vee B$' entails 'if A then B'; i.e. '$A \supset B$' entails 'if A then B'.

(ii) It is widely accepted that the indicative conditional entails the material conditional; that is, 'If A then B' entails '$A \supset B$.

So we get the equivalence of 'If A then B' and '$A \supset B$'.

8.2.2 Arguments Against Material Implication

1. The analysis of conditionals as expressing material implication gives rise to the so-called **"paradoxes of material implication"**.[2] Take another look at the truth table: the falsity of the antecedent is sufficient for the truth of the conditional, and so is the truth of the consequent.

Positive Paradox of MI:	$B \models_{\mathsf{C}} A \supset B$
Negative Paradox of MI:	$\neg A \models_{\mathsf{C}} A \supset B$

   Ad 1: Suppose (correctly) that the SPÖ is in the Austrian government. Does it follow from this that if people go to the elections, then the SPÖ is in the Austrian government?

   Ad 2: Suppose (correctly) that there aren't 25 people in the room. Does it follow from this that if there are 25 people in the room, then there are 50 people in the room?

2. The analysis also has problems with **embedded conditionals**. For instance, consider cases where indicative conditionals are embedded under nominal quantifiers:

   (8.17) a. Every student will succeed if he works hard.
   b. No student will succeed if he goofs off.

   The material implication analysis may make the correct predictions for (9.6a), but it predicts that (9.6b) means that every students goofs off and does not succeed (Higginbotham, 1986): The analysis of (9.6b) is $\neg\exists x(Sx \wedge Gx \supset Ux)$ — equivalently, $\forall x \neg(Sx \wedge Gx \supset Ux)$ and $\forall x(Sx \wedge Gx \wedge \neg Ux)$.

3. There are **indicative/subjunctive** pairs such as (10.1a) and (10.1b), which have the same antecedent and consequent, yet one may be true and the other false. So they cannot both be expressed by material implication.

[2] In logic, a proposition or statement counts as paradoxical if it is self-contradictory and its obvious alternatives are either self-contradictory or very costly (e.g. the Liar paradox). A thumbnail definition of 'paradox' in philosophy: a paradox is a set of propositions, or sentences, all of which seem true/acceptable but which are mutually inconsistent. The 'paradoxes of material implications' are not strictly paradoxes in either of these senses, but rather shortcomings in the match between the formal analysis and our intuitive judgments regarding (instances of) conditionals (in natural language).

4. Material implication validates a number of inference patterns that intuitively aren't valid for conditionals.

   (8.18) $A \supset B \vDash_{\mathsf{C}} (A \wedge C) \supset B$ (strengthening of the antecedent)
   (8.19) $(A \wedge B) \supset C \vDash_{\mathsf{C}} (A \supset C) \vee (B \supset C)$
   (8.20) $(A \supset B) \wedge (C \supset D) \vDash_{\mathsf{C}} (A \supset D) \vee (C \supset B)$
   (8.21) $\neg(A \supset B) \vDash_{\mathsf{C}} A$
   (8.22) $\neg(A \supset B) \vDash_{\mathsf{C}} \neg B$

5. Hunter's counterexample is an instance of (9.4):

   Tax man: If you filled in your tax form incorrectly and owe the Government money, you intended to cheat the Government.
   Tax payer: Thats not true!
   Tax man: So you admit you filled in your tax form incorrectly and owe the Government money!

8.2.3 A Sophisticated Defence of Material Implication

Next time well will look at a famous attempt at defending the claim that *if. . . (then)* means $\supset$. It draws a principled distinction between semantics and pragmatics and argues that *if. . . (then) semantically* means $\supset$ (even if pragmatically, it conveys more). This attempt is due to Grice (1975) and Jackson (1979).

8.3 Strict Implication

- Material implication makes implication a contingent affair: If A and B happen to be true, then both $A \supset B$ and $B \supset A$ are also true.

- Lewis (1917, 355): "'Proof' requires that a connection of content or meaning or logical connection be established. And this is not done for the postulates and theorems in material implication . . . For a relation which does not indicate relevance of content is merely a connection of 'truth-values', not what we mean by a 'logical' relation or 'inference'."

- According to C. I. Lewis, material implication does not capture any 'logical' connection between antecedent and consequent. But it seems that a conditional says that the consequent **must** follow from the premises – there is a connection of **necessity**. What captures the idea that A implies B is $A \supset B$'s being **necessarily** true. Strict implication, $p \prec q$ means: $A \supset B$ is necessarily true.

$$A \prec B \;=_{\text{df.}}\; \Box(A \supset B)$$

- Strict implication is immune to a number of the objections raised against the material conditional above. For example:

  (8.23) $B \not\models_{\mathsf{K}} A \strictif B$ (cf. paradox of MI)
  (8.24) $\neg A \not\models_{\mathsf{K}} A \strictif B$ (cf. paradox of MI)
  (8.25) $(A \wedge B) \strictif C \not\models_{\mathsf{K}} (A \strictif C) \vee (B \strictif C)$
  (8.26) $(A \strictif B) \wedge (C \strictif D) \not\models_{\mathsf{K}} (A \strictif D) \vee (C \strictif B)$
  (8.27) $\neg(A \strictif B) \not\models_{\mathsf{K}} A$
  (8.28) $\neg(A \strictif B) \not\models_{\mathsf{K}} \neg B$

  These are all valid with $\supset$ in place of $\strictif$. However, there is an analogous set of concerns that can be raised about the strict conditional.

8.3.1 The Paradoxes of Strict Implication and Other Problems

1. The Paradoxes of Strict Implication:

   Positive Strict Paradox: $\Box B \models_{\mathsf{K}} (A \strictif B)$

   Negative Strict Paradox: $\neg\Diamond A \models_{\mathsf{K}} (A \strictif B)$

   Counter-examples: If April is rainy, then 2 is the only even prime number.
   If there are no even prime numbers, the Greens will win the next elections.

2. Reasoning with necessities and impossibilities.

   By the Positive Strict Paradox, as soon as we reason with consequents that are necessary, we get true conditionals – no matter the antecedent. And by the Negative Strict Paradox, as soon as we reason with antecedents that are impossible, we get true conditionals – no matter the consequent. But reasoning with necessary and/or impossible propositions is essential to many areas of inquiry: mathematics, logic, (parts of) philosophy, etc.

3. Explosion:

   $A \wedge \neg A$ is a contradiction; it isn't true at any world. So it is impossible. Hence $\models_{\mathsf{K}} (A \wedge \neg A) \strictif B$. By modus ponens, we have $A \wedge \neg A \models_{\mathsf{K}} \strictif B$. This means that contradictions entail everything. But sometimes, we *may* need to reason with contradictions. A logic in which anything follows from contradictions is of no use in this case (cf. Priest (2008, §4.8) for examples of reasoning with contradictions).

8.3.2 In Defense of Strict Implication

C.I. Lewis accepted the Positive and Negative Strict Paradoxes as "sound logical principles", even though they at first appear "not in accord with the 'proper' sense of *implies*" (C.I. Lewis, *A Survey of Symbolic Logic*, 1918, p. 339). His argument for Negative Strict Paradox runs (Lewis and Langford, *Symbolic Logic* pp. 250-1):

Suppose	$A \wedge \neg A$	
then	$\neg A$	by conjunction elimination
so	$\neg A \vee B$	by disjunction introduction
But	A	by conjunction elimination again
So	B	by disjunctive syllogism $(A, \neg A \vee B \vdash B)$

which assumes that any impossible proposition can be put in the form of an explicit contradiction, $A \wedge \neg A$. He also produced a proof for Positive Strict Paradox:

Suppose	A	
then	$(A \wedge B) \vee (A \wedge \neg B)$	
so	$A \wedge (B \vee \neg B)$	by distribution
So	$B \vee \neg B$	by conjunction elimination

this time assuming that any necessary proposition can be put in the form $B \vee \neg B$.

8.4 Optional Exercises

1. (a) Show that the inference patterns in §2.2, (9.1) – (9.5), are valid in classical propositional logic.
   (b) Find English (or German) natural language instances of the inference patterns in §2.2, (9.1) – (9.5), which show that the patterns are intuitively invalid.
   (c) Consider the inference patterns in §2.2, (9.1) – (9.5): Are the premises also semantic consequences of the conclusions? (That is, e.g., does the following hold: $(A \wedge C) \supset B \vDash_{\mathsf{C}} A \supset B$?) In each case, is this result in accordance with your judgments about conditionality?
2. Check, by using tableaux, whether the inference patterns in (8.23) – (8.28) are invalid in normal modal logics stronger than K.
   (Hint: (i) Replace any formula '$A \multimap B$' with '$\Box(A \supset B)$' on the tree. (ii) Try $\mathsf{S5}$ first and go from stronger to weaker logics: If an inference pattern is invalid in a stronger system, it is invalid in a weaker system.)
3. Which of the objections against material implication are also strong objections against strict implication?
4. Can you think of further inference patterns that are valid with (a) material implication or (b) strict implication and which are intuitively invalid? Give examples to illustrate their intuitive invalidity.

8.5 Readings

Sources of this handout:

- von Fintel (2011) — optional reading
- Priest (2008, §§1.6–1.10, 4.5–4.9) — obligatory reading
- Edgington, Dorothy: Conditionals. *The Stanford Encyclopedia of Philosophy* (May 2014), Edward Zalta (ed.), URL = http://plato.stanford.edu/entries/conditionals/, sections 1-2

9 Grice's Defense of Material Implication

9.1 Material Implication

- Material implication: '$A \supset B$' is true if either A is false or B is true; it is false if A is true and B is false.
- Among the truth-functional, two-place connectives, material implication is the most plausible candidate for the conditional.

> 'A professor who declares that *If I am healthy, I will come to class* can only be said to have broken her promise if she is healthy but doesnt come to class. Clearly, if she is healthy and comes to class, shell have spoken the truth. And if she is sick, it is immaterial whether she comes to class (going beyond the call of duty and beyond what she promised) or doesnt neither case constitutes a breaking of the promise.' (von Fintel, 2011, §3.1; example attributed to Suber)

- For arguments in favour of the material implication account of conditionals, see *Handout V-1*, §2.1.

9.2 The Equivalence Thesis

9.2.1 The Unsupplemented Equivalence Thesis and its Problems (Review)

Unsupplemented Equivalence Thesis: The meaning of natural-language indicative conditionals is material implication.

Some of the main problems with the Unsupplemented Equivalence Thesis (see *Handout V-1* for details):

1. **Paradoxes of material implication**.

Positive Paradox:	$B \vDash_{\mathsf{C}} A \supset B$
Negative Paradox:	$\neg A \vDash_{\mathsf{C}} A \supset B$

2. Material implication validates a number of **intuitively invalid inference patterns**.

(9.1) $A \supset B \vDash_{\mathsf{C}} (A \wedge C) \supset B$ (strengthening of the antecedent)
(9.2) $(A \wedge B) \supset C \vDash_{\mathsf{C}} (A \supset C) \vee (B \supset C)$
(9.3) $(A \supset B) \wedge (C \supset D) \vDash_{\mathsf{C}} (A \supset D) \vee (C \supset B)$
(9.4) $\neg(A \supset B) \vDash_{\mathsf{C}} A$
(9.5) $\neg(A \supset B) \vDash_{\mathsf{C}} \neg B$

3. **Embedded conditionals**. E.g., indicative conditionals embedded under nominal quantifiers:

(9.6) a. Every student will succeed if he works hard.
b. No student will succeed if he goofs off.

9.2.2 The Supplemented Equivalence Thesis

Grice (1975) and Jackson (1979) defend the following thesis:

The Supplemented Equivalence Thesis: The *semantic, conventional* meaning of natural-language indicative conditionals is material implication.

The basic idea: 'If A then B' *literally* (semantically, conventionally) means $A \supset B$. However, it *pragmatically* implicates that there is a connection between A and B. So a literally true indicative conditional may still fail to be *assertable* because it implicates a false proposition.

9.3 Grice on Communication

- **What is said**: The notion of what someone has said is closely related to the conventional (=semantic) meaning of the words (the sentence) she has uttered.

  What the speaker said involves *some* pragmatic mechanisms: For a full identification of what the speaker of "He is in the grip of a vice" has said, one needs to know what 'he' refers to, the time of utterance, the meaning, on the particular occasion of utterance, of the phrase *in the grip of a vice* (resolution of ambiguity).

- **Implicature**: The implicatures generated by an utterance of a sentence are the propositions that are intentionally communicated by the speaker with her utterance, but are not part of what is said by the sentence in the context of use.

- **Conversational implicature**:

- Conversational implicatures are propositions communicated by an utterance in virtue of what is said together with general facts about the context and conversational norms. Conversational implicatures are *cancelable* and *non-detachable*.
- Grice (1975, 31):

  > A man who, by (in, when) saying (or making as if to say) that p has implicated that q, may be said to have conversationally implicated that q, provided that (1) he is to be presumed to be observing the conversational maxims, or at least the Cooperative Principle; (2) the supposition that he is aware that, or thinks that, q is required in order to make his saying or making as if to say that p (or doing so in *those* terms) consistent with this presumption; and (3) the speaker thinks (and would expect the hearer to think that the speaker thinks) that it is within the competence of the hearer to work out, or grasp intuitively, that the supposition mentioned in (2) is required.

- Examples:

  (9.7) In a letter of reference for a philosophy PhD student, "Smith has beautiful handwriting and was never late for class."
  $\Rightarrow$ Implicature: Smith is not a good philosopher.

  (9.8) A: I am out of petrol.
  B: There is a garage around the corner.
  $\Rightarrow$ Implicature: The garage is open and selling petrol.

- The general principle: **The Cooperative Principle**

  > Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

- **4 Categories of Maxims**

  1. **Maxim of Quantity**:
     a) Make your contribution as informative as is required (for the current purposes of the exchange).
     b) Do not make your contribution more informative than is required.
  2. **Maxim of Quality**:
     Supermaxim: "Try to make your contribution one that is true"
     a) Do not say what you believe to be false.
     b) Do not say that for which you lack adequate evidence.
  3. **Maxim of Relation**:
     Be relevant.
  4. **Maxim of Manner**:
     Supermaxim: "Be perspicuous"

a) Avoid obscurity of expression.
b) Avoid ambiguity.
c) Be brief (avoid unnecessary prolixity).
d) Be orderly.

- Four ways in which one can fail to fulfill a maxim:

  1. **Violate** a maxim: quietly and unostentatiously, in some cases liable to mislead
  2. **Opt out** from the operation of the maxim and of the Cooperative Principle by saying/indicating that one is unwilling to cooperate in the way the maxim requires.
  3. Be faced with a **clash**: be unable to fulfill one maxim without violating another.
  4. **Flout** a maxim: blatantly fail to fulfill the maxim.
     "On the assumption that the speaker is able to fulfill the maxim and to do so without violating another maxim (because of a clash), is not opting out, and is not, in view of the blatancy of his performance, trying to mislead, the hearer is faced with a minor problem: How can the speaker's saying what he did say be reconciled with the supposition that he is observing the overall Cooperative Principle? This situation is one that characteristically gives rise to a conversational implicature; and when a conversational implicature is generated in this way, I shall say that a maxim is being **exploited**." (Grice, 1975, 30)

- **Two primary features of conversational implicatures**:

  1. Cancelability
     A conversational implicature can be canceled, explicitly or 'contextually' (Grice, 1975, 39): It can be (made) clear that the speaker is opting out.

     (9.9) a. A: X is meeting a woman this evening.
     $\Rightarrow$ Usual generalized conversational implicature: the person X is meeting is not X's wife, mother, sister, or perhaps even close platonic friend.
     b. B: X is meeting a woman this evening, who is in fact his sister.

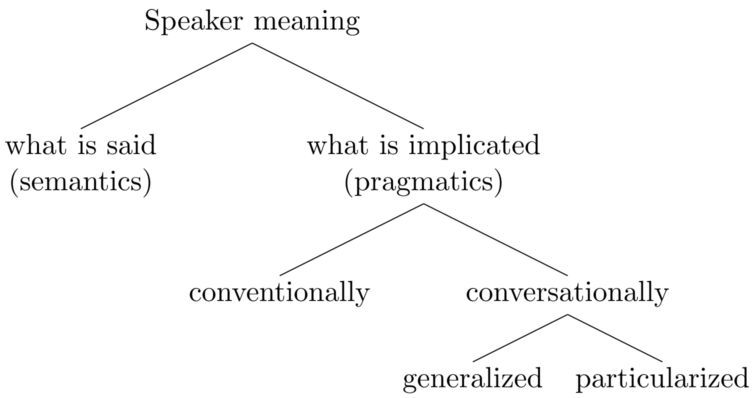
  2. Non-detachability
     "it will not be possible to find another way of saying the same thing, which simply lacks the implicature in question" (Grice, 1975, 39)

     (9.10) a. A: X is meeting a woman this evening.
     b. B: X is meeting a member of the female gender this evening.
     $\Rightarrow$ Generalized conversational implicature in *both* cases: the person X is meeting is not X's wife, mother, sister, or perhaps even close platonic friend.

- **Conventional implicature**: the conventional meaning of the words used determine (besides helping to determine what the speaker says) what is implicated.

Example: "He is an Englishman; he is, therefore, brave"
Conventional implicature = *his being brave is a consequence of his being an Englishman.* The latter isn't part of what the speaker said since, were the consequence not to hold, the utterance of the above sentence would not be, strictly speaking, false.

- **The Gricean Picture**:



9.4 Grice's Pragmatic Defense of the Supplemented Equivalence Thesis (Grice, 1989a)

- The problems with the Unsupplemented Equivalence Thesis show that it is easy for a material implication to be true (or inferable), so there are many cases where material implication is true (inferable) yet intuitively, the corresponding conditional is false (not inferable).

- Grice's strategy is to explain why we have the intuition of falsity (non-inferabibility) by pointing to a false (non-inferable) conversational implicature that many uses of conditionals in communication have.

- The **Indirectness Condition**:

  "that p would, in the circumstances, be a good reason for q"
  "that q is inferable from p"
  "that there non-truth-functional grounds for accepting $p \supset q$" (Grice, 1989a, 58)

- Grice's thesis:

  "in standard cases to say 'if p then q' is to be conventionally committed to (to assert or imply in virtue of the meaning of 'if') both the proposition that $p \supset q$ and the Indirectness Condition.' (Grice, 1989a, 58)

$\Rightarrow$ Grice ultimately argues that in 'standard' uses of 'if p then q' a speaker literally says that $p \supset q$ and conversationally implicates the Indirectness Condition.

- Grice (1989a, 61-2):

> To say that $p \supset q$ is to say something logically weaker than to deny that p or to assert that q, and is thus less informative; to make a less informative rather than a more informative statement is to offend against the first maxim of Quantity, provided that the more informative statement, if made, would be of interest. There is a general presumption that in the case of "$p \supset q$", a more informative statement would be of interest. No one would be interested in knowing that a particular relation (truth-functional or otherwise) holds between two propositions concerned, unless his interest were of an academic or theoretical kind [. . .] An infringement of the first maxim of Quantity, given the assumption that the principle of conversational helpfulness is being observed, is most naturally explained by the supposition of a clash with the second maxim of Quality ("Have adequate evidence for what you say"), so it is natural to assume that the speaker regards himself as having evidence for the less informative statement (that $p \supset q$)—that is, non-truth-functional evidence. So standardly he implicates that there is non-truth-functional evidence when he says that $p \supset q$;

- Let's spell this out:

1. '$\neg p$' and 'q' are more informative than '$p \supset q$.'
2. In general, we are not interested in the connection between the propositions that p and that q without being interested in their truth values.
3. So in asserting '$p \supset q$', a speaker standardly fails to observe the maxim of Quantity ("make your contribution as informative as is required. . . "): it will standardly give the hearer less information than she would like.
4. If it can be supposed that the speaker is trying to observe the Cooperative Principle, the failure to comply with Quantity is best explained by the fact that she would infringe on the maxim of Quality ("do not say that for which you lack adequate evidence"), were she to make a stronger claim than '$p \supset q$.'
5. So the speaker must (take herself to) have only non-truth-functional evidence for a connection between the propositions that p and that q.
6. So the speaker is conversationally implicating that there is non-truth-functional evidence for a connection between the propositions that p and that q (Indirectness Condition).

Note the similarities with what Grice (1975, 32-3, example 3) labels 'Group B' examples of conversational implicatures, in which the maxim of Quantity is infringed upon and where this infringement is to be explained by the supposition of a clash with the maxim of Quality.

- Grice contends that the Indirectness Condition fulfills the two primary features of conversational implicatures:

  1. Cancelability:

     a) Explicit cancelation: "To say 'If Smith is in the library, he is working' would normally carry the implication of the Indirectness Condition; but I might say (opting out) 'I know just where Smith is and what he is doing, but all I will tell you is that if he is in the library he is working."' (Grice, 1989a, 59)

     b) Contextual cancelation: there are cases in which the Indirectness Condition is simply absent:

        (9.11) Perhaps if he comes, he will be in a good mood.
        (9.12) See that, if he comes, he gets his money. (Grice, 1989a, 60)

        Note: Biscuit conditionals are also uses of conditionals that do not convey the Indirectness Condition.

  2. Non-detachability:

     (9.13) Either Smith is not in London, or he is attending the meeting.

     (9.14) It is not the case that Smith is both in London and not attending the meeting.

     According to Grice, (9.13) and (9.14) – both of which say the same as 'If Smith is in London, he's attending the meeting' (they're truth-functionally equivalent with 'Smith is in London $\supset$ Smith is attending the meeting') – also implicate the Indirectness Condition.

Grice's Defense and the Problems of the Unsupplemented Equivalence Thesis

1. **Paradoxes of material implication**

   Explanandum for Grice: why instances of the paradoxes strike us as unassertable/false despite their literal truth.

   (9.15) If there are 25 people in the room, then there are 50 people in the room.

   If the speaker has the information that there aren't 25 people in the room (negative paradox), an assertion of (9.15) infringes on the maxim of Quantity. Similarly, if she has the information that there are 50 people in the room (positive paradox). In either case, a *false* conversational implicature is generated: that there is evidence for a non-truth-functional connection between there being 25 people in the room and there being 50 people in the room (Indirectness Condition). It is this implicature that explains our intuitive judgment that (9.15) is false/unassertible.

9.5 Problems with Grice's Defense

1. **Conditionals in thought** (Edgington, 1995, 245):

   > But the difficulties with the truth-functional conditional cannot be explained away in terms of what is an inappropriate conversational remark. They arise at the level of belief. Believing that John is in the bar does not make it logically impermissible to disbelieve "if he's not in the bar he's in the library". Believing you won't eat them, I may without irrationality disbelieve "if you eat them you will die". Believing that the Queen is not at home, I may without irrationality reject the claim that if she's home, she will be worried about my whereabouts. As facts about the norms to which people defer, these claims can be tested. But, to reiterate, the main point is not the empirical one. We need to be able to discriminate believable from unbelievable conditionals whose antecedent we think false. The truth-functional account does not allow us to do this.

   Note that Edgington's criticism applies generally to pragmatic explanations of our judgments of truth (falsity) in terms of judgements about (un)assertability that target conversational implicatures. The point seems to be independent of a pragmatic defense of *conditionals*.

2. **Contraposition**

   (9.16) a. Even if the Bible was divinely inspired, it is still not literally true.
   b. If the Bible is literally true, then it is not divinely inspired. (Bennett)

   (9.17) a. If it rains, it will not rain heavily.
   b. If it rains heavily, it will not rain. (Jackson)

   Material implication validates **contraposition**: $A \supset B \vDash_{\mathsf{C}} \neg B \supset \neg A$.

   (9.16a) and (9.16b) make it intuitively plausible that 'If A then B' and 'If not-B then not-A' do not have the same truth conditions: (9.16a) can be true and (9.16b) false (the same holds of (9.17a/b)). However, Grice's strategy cannot explain *this* difference in truth/assertability judgments ('If A then B' and 'If not-B then not-A' are equally 'logically strong' and so equally informative).

3. **Disanalogies in assertibility between *if...then* and *or*** (Grice, 1975, 63)

   '$A \supset B$' is logically equivalent with '$\neg A \vee B$' and according to Grice, the literal meaning of 'or' is truth-functional disjunction ('$\vee$'). Thus, a speaker should standardly implicate the Indirectness Condition with uses of 'or' just as much as she implicates it with standard uses of 'if...(then)'. But there are important differences, as Grice (1989a, 63) himself remarks: "whereas ...a disjunctive statement which has been *advanced* on non-truth-functional grounds can be *confirmed* truth-functionally, by establishing one of its disjuncts, the parallel idea with regard to conditionals is not acceptable." For illustration:

(9.18) a. Either Wilson won the elections or Thorpe did.
b. That's right. In fact, Wilson won.

(9.19) a. If Wilson didn't win, then Thorpe did.
b. ? That's right. In fact, Wilson won.

4. **Negations of indicative conditionals**

The literal meaning of a negated conditional 'It is not the case that if A, then B' is $\neg(A \supset B)$. The latter is logically equivalent to $A \wedge \neg B$. However, it seems that we often assert the negation of a conditional and say, or implicate, neither A nor not-B.

(9.20) Speaker A: If God exists, we are free to do whatever we like.
Speaker B: That's not the case.

We can take speaker B's assertion to be short for 'It is not the case that if God exists, we are free to do whatever we like.' But it seems unreasonable to understand B as saying (implicating) that God exists and we are not free to do whatever we like. (Grice (1989a, 80-1) noticed this problem and attempted to solve it by arguing that with negations of conditionals, we either in fact assert '$A \supset \neg B$' (the negation takes narrow scope) or deny the conversational implicature that there is non-truth-functional evidence for the material implication.)

9.6 Readings

- Grice (1975)
- Optional: Bennett (2003, pp. 20–27) [N.B.: In his exposition of Grice on indicative conditionals on pp. 24-5, Bennett is mistakenly appealing to the maxims of quantity (1a) and manner (4c) instead of quantity (1a) and quality (2b). See Grice (1989a, 61-2) for this point.]
- More on Grice's defense of the Supplemented Equivalence Thesis: Grice (1989a)
- Jackson's pragmatic defense of the equivalence thesis: Jackson (1979)

10 Stalnaker's Theory of Conditionals

10.1 The Direct Argument

- Stalnaker's example: 'Either the butler or the gardener did it. Therefore, if the butler didn't do it, the gardener did.' (Stalnaker, 1975, 63)
- (i) or-to-if: '$A \vee B$' entails 'If $\neg A$, then B'. Substitute '$\neg A$' for 'A'. '$\neg A \vee B$' entails 'if A then B'; i.e. '$A \supset B$' entails 'if A then B'.

  (ii) It is widely accepted that the indicative conditional entails the material conditional; that is, 'If A then B' entails '$A \supset B$.
  So we get the equivalence of 'If A then B' and '$A \supset B$' (cf. *Handout V-1*, §2.1).
- But there are well-known problems with the claim that 'If A then B' and '$A \supset B$' are equivalent – for example, the paradoxes of material implication are valid.
- Stalnaker: there are two options that make use of the distinction between semantics and pragmatics.

  1. Grice's strategy: Defend the *semantic* equivalence of 'If A then B' and '$A \supset B$', and explain *pragmatically* why those inferences (e.g. the paradoxes of MI) seem invalid (see *Handout V-2*).
  2. Stalnaker's strategy: Reject the material conditional analysis in favour of a stronger semantic account of *If... (then)*, and explain *pragmatically* why the direct argument seems valid.

     Stalnaker defines a semantic notion of *entailment* and a pragmatic notion of *reasonable inference* and shows that, while the premise of the direct argument ('A or B') doesn't *entail* the conclusion ('If not-A, B'), the conclusion can *reasonably be inferred* from the premise.

10.2 Pragmatics

- Inquiring, deliberating, ..., exchanging information (communication): one is essentially distinguishing between alternative ways the world could be – between different possible worlds.
- **Propositions** $=_{\text{df.}}$ functions from possible worlds into truth values (*true* and *false*), or equivalently sets of possible worlds (those at which the proposition is true)

  Example: The proposition *that snow is white* is a function that takes every world at which snow is white into *true*, and every other world into *false*. Equivalently: $\{w$: snow is white at $w\}$
- **Context set** $=_{\text{df.}}$ the set of possible worlds not ruled out by the presupposed background information

  In conversation, participants make assumptions about what is common ground among all the participants: they presuppose certain propositions. Suppose participants A and B both presuppose (in the above sense) propositions p and q. Then their context set is the *intersection of* p and q – it's the set of possible worlds at which both p and q are true. If they haven't established whether r, then their context set contains some r-worlds and some *not*-r-worlds
- **Assertion**: It is an essential effect of assertion that they add the asserted proposition to the presupposed background information (if they're accepted by all participants). That is, successful assertions of r have the effect of intersecting the context set with r. (To add r to the common ground p, q is to intersect the context set $C_{p,q}$ with r: the context set after the successful assertion of r is $C_{p,q} \cap r$.)
- A proposition is **compatible** with a context iff it is true in some of the worlds in the context set.
  Intuitively: the proposition is among the presuppositions corresponding to the context set (or is a consequence of those presuppositions).
- A proposition is **entailed** by a context iff it is true in all of the worlds in the context set.
  Intuitively: the presuppositions corresponding to the context set determine that the proposition is true.
- In linguistic communication, speakers try to establish information by distinguishing between possible worlds in the context set. An increase in information corresponds to a reduction of possible worlds in the context set.
- Pragmatics has to do with assertability & acceptability of utterances. Semantics has to do with truth and falsity of sentences (in context).

10.3 Semantics for Conditionals

- Stalnaker gives the same semantic analysis for indicative and subjunctive conditionals. Their difference is a pragmatic one.

- Stalnaker's basic idea:

> 'The idea of the analysis is this: a conditional statement, *if A, then B*, is an assertion that the consequent is true, not necessarily in the world as it is but in the world as it would be if the antecedent were true.'

- Let's make this formally more precise. We're going to be a little more sloppy than usual with the formal details (since we've got *a lot* to cover today). A rigorous presentation of Stalnaker's logic of conditionals, called C_2, is given in his 'A Theory of Conditionals' (1968), which we follow in this presentation. (See also Priest (2008, ch. 5) and Sider (2010, ch. 8).)
- C_2 is a propositional modal logic (see *Handout III-1* and *III-2*).

Definition 10.3.1. A **model structure** M_S is a structure $\langle \mathcal{W}, \mathcal{R}, \lambda \rangle$, where

(i) $\mathcal{W}$ is a non-empty set of objects, intuitively understood as possible worlds
(ii) $\mathcal{R}$ is an accessibility relation between worlds
(iii) λ is the absurd world, at which contradictions and all their consequences are true. The following conditions apply: (a) $\lambda \in \mathcal{W}$. (b) It is not accessible from any worlds nor does it access any worlds: for any w, $\langle \lambda, w \rangle \notin \mathcal{R}$ and $\langle w, \lambda \rangle \notin \mathcal{R}$.
(λ is needed for the interpretation of conditionals with impossible antecedents.)

- **Selection function** f: In addition to a model structure, the semantic apparatus includes a selection function, f, which takes a proposition p and a world w as arguments and maps them to a unique world w': $f(p, w) = w'$.
Intuitively, $f(p, w)$ is the "closest", or "most similar" world (to w) at which p is true.
- The semantic clause for the conditional, written '>', can be stated as follows:

$$\begin{array}{lll} A > B \text{ is true at } w & \text{if} & B \text{ is true at } f(A, w); \\ A > B \text{ is false at } w & \text{if} & B \text{ is false at } f(A, w); \end{array}$$

- The semantic clauses for the atomic case, for propositional connectives ($\neg, \wedge, \vee, \supset, \equiv$) and modal operators ($\Box, \Diamond$) are the standard ones (see Definition 3.3 on *Handout III-1*).

Definition 10.3.2. An inference is **valid in system C_2** iff every world of every model $\mathcal{M}$ of model structure M_S at which the premises are true is one at which the conclusion is also true; i.e.

$$\Sigma \vDash_{\mathsf{C}_2} A \text{ iff for all worlds } w \in W \text{ of all models } \mathcal{M} \text{ of model structure } M_S\text{:}$$
$$\text{if } B \text{ is true at } w \text{ in } \mathcal{M} \text{ for all the premises } B \in \Sigma, \text{ then } A \text{ is true at } w \text{ in } \mathcal{M}.$$

- Stalnaker lays down a number of conditions on the selection function f that are meant to ensure that it yields the most similar world:
$f(A, w) = w'$: call A the *antecedent*, w the *base world*, and w' the *selected world*.

(1) For all antecedents A and base worlds w, A must be true at $f(A, w)$.

(2) For all antecedents A and base worlds w, $f(A, w) = \lambda$ only if there is no world accessible from w at which A is true.

N.B.: The absurd world is selected only when the antecedent is impossible.

(3) For all antecedents A and base worlds w, if A is true at w, then $f(A, w) = w$.

N.B.: No world is more similar to w than w itself. So if A is true at w, w itself is selected for the consideration of whether or not B is the case.

(4) For all antecedents B and B' and base worlds w, if B is true at $f(B', w)$ and B' is true at $f(B, w)$, then $f(B, w) = f(B', w)$.

N.B.: 'The fourth condition ensures that the [similarity] ordering among possible worlds is consistent in the following sense: if any selection established [world] β as prior to β' in the ordering (with respect to a particular base world α), then no other selection (relative to that α) may establish β' as prior to β.' (Stalnaker, 1968, 105)

10.3.1 The Context-Dependence of Conditionals

- What does "closest", or "most similar" mean? It will depend on the context: 'Relevant respects of similarity are determined by the context.' (Stalnaker, 1968, 69)

- Stalnaker (1975, 69) adds a fifth, *contextual* condition:

  (5) If w is in the context set, then $f(A, w)$ must, if possible, be within the context set.

  N.B.: 'all worlds within the context set are closer to each other than any worlds outside it.' 'The idea is that when a speaker says "If A," then everything he is presupposing to hold in the actual situation is presupposed to hold in the hypothetical situation in which A is true.'

- Stalnaker (1975, 69):

  > 'The motivation for the principle is this: normally a speaker is concerned only with possible worlds within the context set, since this set is defined as the set of possible worlds among which the speaker wishes to distinguish. So it is at least a normal expectation that the selection function should turn first to these worlds before considering counterfactual worlds–those presupposed to be non-actual.'

- This makes the truth-conditions – here: the semantic content – of (indicative) conditionals **context-dependent**: they depend on the context in a way similar to how the content, or reference, of a definite description or pronoun depends on the context (on who/what is being picked out by the expression in the context of utterance).[1]

[1] Note that Stalnaker uses the word 'pragmatic' in (Stalnaker, 1975, 69) in the sense introduced by R. Montague: pragmatics fills the gap between the context-invariant features of sentences and what a speaker

10.3.2 Indicative vs Subjunctive Conditionals

- Stalnaker' gives a single semantics for both indicative and subjunctive conditionals. But as we have seen (cf. example (8) on *Handout V-1*), two conditionals can differ in truth value when they otherwise differ only in that one is in indicative mood, and the other in subjunctive:

  (10.1) a. If Oswald didn't kill Kennedy, someone else did.
  b. If Oswald hadn't killed Kennedy, someone else would have. (Adams, 1970)

- So how does Stalnaker distinguish between indicative and subjunctive conditionals?
- The difference for Stalnaker is in their context-dependence. The subjunctive mood, he claims, suspends the pragmatic condition (5) on the selection function, which is 'only a defeasible presumption and not a universal generalization':

  > 'I take it that the subjunctive mood in English and some other languages is a conventional device for indicating that presuppositions are being suspended, which means in the case of subjunctive conditional statements, that the selection function is one that may reach outside of the context set. Given this conventional device, I would expect that the pragmatic principle stated above should hold without exception for indicative conditionals.' (Stalnaker, 1975, 70)

- So indicative conditionals in context are associated with different selection functions than subjunctive conditionals: Condition (5) holds for indicative conditionals, but not for subjunctive conditionals.
- Consider again (10.1): for most of us with knowledge of recent US history, the context set is one in which all worlds are worlds in which Kennedy got killed. As concerns the indicative (10.1a), f(*Oswald didn't kill Kennedy*,@) is in the context set. So it is a world in which Kennedy got killed — by someone other than Oswald.

  In contrast, (10.1b) is assertable/acceptable only if f(*Oswald didn't kill Kennedy*,@) reaches outside of this context set, plausibly to a world in which Kennedy didn't get killed. So in that world, it's not the case that someone other than Oswald killed Kennedy.

---

says by uttering a sentence. Contrast this Montagovian sense with the Gricean, where pragmatics fills the gap between what is said and what is communicated (between what a speaker says and what (s)he means). So on the Gricean understanding, Montagovian pragmatics does its job in contributing to what is said, and thus falls on the Gricean side of semantics. This is why we treat it here in the section 'Semantics.'

10.4 Semantic Entailment vs Reasonable Inference

Let's come back to Stalnaker's strategy of dealing with the direct argument. He gives the following 'rough informal' definitions of the pragmatic notion of a reasonable inference and the semantic notion of entailment:

- **Reasonable inference**:

  > 'an inference from a sequence of assertions or suppositions (the premises) to an assertion or hypothetical assertion (the conclusion) is *reasonable* just in case, in every context in which the premises could appropriately be asserted or supposed, it is impossible for anyone to accept the premises without committing himself to the conclusion'.

- **Semantic entailment**:

  > 'a set of propositions (the premises) *entails* a proposition (the conclusion) just in case it is impossible for the premises to be true without the conclusion being true as well.' (Stalnaker, 1975, 65)

- We can now see that the or-to-if inference – the first part of the direct argument: 'A or B; therefore, if $\neg A$ then B' – is not valid in C_2: Take an arbitrary world w and suppose that A is true at w and B is false at w. Then $A \vee B$ is true at w. Suppose further that $f(\neg A, w)$ is a $\neg B$-world. Then $\neg A > B$ is false at w.

- Since Stalnaker accepts the second part (ii above) – that the indicative conditional entails the material conditional – his strategy is to show why the or-to-if inference *seems to be* valid. According to Stalnaker, the or-to-if inference is a *reasonable inference.*

- To show that, he adds the following Gricean principle for the appropriateness of disjunctive assertions:

  (D) 'A disjunctive statement is appropriately made only in a context which allows either disjunct to be true without the other. That is, one may say *A or B* only in a situation in which both *A and not-B* and *B and not-A* are open possibilities.' (Stalnaker, 1975, 71)

  (D) reflects the Gricean maxims of Quantity and Quality: Suppose first that in a conversation, it is already established that A is the case (or that B is the case or that *A and B* is the case). Then saying *A or B* adds no information, infringing on Quantity. Suppose second that in a conversation, it is already established that neither A nor B is the case. Then saying *A or B* violates Quality: one says something one takes to be false.

- Stalnaker's argument that the or-to-if inference is reasonable:

  We need to show that in every context in which $A \vee B$ is assertable and accepted, $\neg A > B$ must be accepted. That is, we need to show that for any context in which $A \vee B$ is assertable and subsequently asserted, the resulting context entails $\neg A > B$.

  1. Suppose $A \vee B$ is assertable.
  2. Given (D), $\neg A \wedge B$ is compatible with the context (i.e. there is at least one world in the context set at which $\neg A$ and B are true).
  3. Suppose $A \vee B$ is accepted. Then the resulting context set contains some $(\neg A \wedge B)$-worlds, but not $(\neg A \wedge \neg B)$-worlds.
  4. Take any world w in the context set.
  5. By the pragmatic condition on f – (5) above –, $f(\neg A, w)$ is in the context set.
  6. Since $A \vee B$ is accepted, it is true at all worlds in the context set, and so true at $f(\neg A, w)$.
  7. By condition (1) on f, $\neg A$ is true at $f(\neg A, w)$.
  8. Hence, B is true at $f(\neg A, w)$ (since there are no $(\neg A \wedge \neg B)$-worlds in the context set).
  9. So $\neg A > B$ is true at w.
  10. Since w is any arbitrary world in the context set, it follows that $\neg A > B$ is true at all worlds in the context set, i.e. it is accepted in the context.

- Stalnaker (1975, 72)

  > '...the indicative conditional and the material conditional are equivalent in the follow- ing sense: in any context whether either might appropriately be asserted, the one is accepted, or entailed by the context, if and only if the other is accepted, or entailed by the context. This equivalence explains the plausibility of the truth-functional analysis of indicative conditionals, but it does not justify that analysis since the two propositions coincide only in their assertion and acceptance conditions, and not in their truth-conditions.'

10.5 Prominent Validities and Invalidities in C_2

The following are validities in Stalnaker's logic C_2:

(10.2) $\Box(A \supset B) \supset (A > B)$

(10.3) $(A > B) \supset (A \supset B)$

(10.4) $(A > (B \vee C)) \supset ((A > B) \vee (A > C))$

(10.5) Conditional excluded middle: $(A > C) \vee (A > \neg C)$

Comments:

- (10.2) and (10.3) together entail that the conditional is the 'intermediate between strict implication and the material conditional.' (Stalnaker, 1968, 106)
- Ad (10.4): It is instructive to see why (10.4) is valid: Suppose $A > (B \vee C)$ is true at a random world w. Then $f(A, w)$ is a world at which $B \vee C$ is true. So either B or C is true at $f(A, w)$. Suppose it is B. Then $A > B$ is true at w. Suppose it is C. Then $A > C$ is true at w. Hence, $(A > B) \vee (A > C)$ is true at w.
- Ad (10.5): You can check the validity of conditional excluded middle easily: At every world w, either C is true or C is false, i.e. $\neg C$ is true. So for any w' and any A, $f(A, w')$ is either a $C - world$ or a $\neg C$-world. If $f(A, w')$ is a C-world, then $A > C$ is true at w'; if $f(A, w')$ is a $\neg C$-world, then $A > \neg C$ is true at w'. So either way, $(A > B) \supset (A \supset B)$ is true at w'.

Stalnaker's logic C_2 invalidates the following inference patterns:

(10.6) Antecedent Strengthening: $A > B \not\models_{\mathsf{C}_2} (A \wedge C) > B$

(10.7) Hypothetical Syllogism/Transitivity: $A > B, B > C \not\models_{\mathsf{C}_2} A > C$

(10.8) Contraposition: $A > B \not\models_{\mathsf{C}_2} \neg B > \neg A$

Comments:

- A counterexample to (10.6):

  (10.9) a. If this match were struck, it would light.
  b. Therefore, if this match were struck and had been soaked in water, it would light.

- Ad (10.7): Let $f(A, w)$ be a B and $\neg C$-world, and $f(B, w)$ be a C-world. Then $A > B$ and $B > C$ are both true at w, but $A > C$ is false at w.

  Stalnaker (1968, 106) gives the following example to support the invalidity in C_2 of hypothetical syllogism:

  (10.10) a. If J. Edgar Hoover had been born a Russian, he would have been a communist.
  b. If he had been a communist, he would have been a traitor.
  c. Therefore, if he had been born a Russian, he would have been a traitor. (version in Lewis (1973, 33))

- Ad (10.8): Let $f(A, w)$ be a B-world; and let $f(\neg B, w)$ be an A-world. Then $A > B$ is true at w, but $\neg B > \neg A$ is false at w.

  Contraposition also has counterexamples:

(10.11) Lewis (1973, 35): Suppose Boris meant to go the party, but stayed away solely to avoid Olga. Olga did go to the party and would have liked it even better, had Boris been there. Then (a) seems true and (b) false.

a. If Boris had gone to the party, Olga would still have gone.
b. If Olga had not gone, Boris would still not have gone.

10.6 Readings

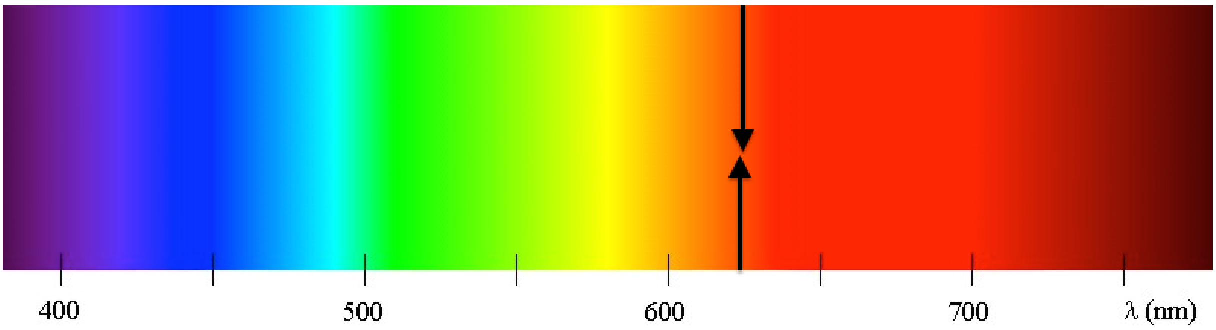
- Stalnaker (1975)
- Optional: Priest (2008, ch. 5) and Sider (2010, ch. 8: especially §§8.2-8.3, 8.6-8.9) are good presentations of the formal details of Stalnaker's and Lewis's logics of counterfactuals.
- Further reading:
  - Stalnaker's most explicit presentation of his formal system is in Stalnaker (1968).
  - Perhaps the best, most comprehensive presentation of Stalnaker's pragmatic framework is in his famous paper 'Assertion' (Stalnaker (1978)).
  - Lewis's logic of counterfactuals, which resembles Stalnaker's C_2 in many respects, and his discussion of Stalnaker's theory can be found in Lewis (1973).

Part IV

Vagueness

11 The Sorites Paradox & Many-valued Logic

11.1 Vagueness



- Where would you draw the line between red and not-red? Between 620nm and 621nm (arrows)? Which is the shortest wavelength to which 'green' applies?

  How many grains of sand does it take to make a heap? What is the maximum number of hairs on a bald head?

- A moment's reflection suffices to realize that vagueness is a pervasive phenomenon in natural language. But what is vagueness?

- **Vagueness** is standardly defined as the possession of borderline cases.

- A **borderline case** is a case in which we do not know whether to apply the word or not, even though we have all the information that would normally be sufficient to settle the matter.

- Wavelength 620nm is a borderline case of 'red': Despite having all the physical information about color, there doesn't seem to be a matter of fact that settles whether it is red or not. (It's arguably also a borderline case of 'orange.')

Contrast this with wavelengths 680nm and 500nm: 680nm is clearly red, 500nm clearly not red.

- Is vagueness a feature of the meaning of natural languages? Or is it a feature of the world? Are there vague objects?

11.1.1 What Vagueness is Not

1. Vagueness ≠ individuals' linguistic incompetence:

   Experts tell us that the visible color spectrum is continuous, with no sharp boundaries between one color and the next. Are we *all* incompetent? This would cry for an explanation.

2. Vagueness ≠ ambiguity:

   Lexical ambiguity:[1] individual words have more than one meaning in the language to which they belong (where 'meaning' refers to what could be captured by a dictionary).[2] Some clear cases: *bank, bill, stage, still...* Consider an ambiguous adjective:

   (11.1) The storm was terrific.

   Even after clarifying the intended meaning of 'terrific' in an utterance of (11.1), vagueness remains: there are borderline cases of 'terrible/frightening' as well as of 'marvelous/wonderful.'

3. Vagueness & context-sensitivity: vagueness does not reduce to dependence on the comparison class picked out in the context of use.

   Consider Danny, who is 1,75m in height. Is Danny tall?

   (11.2) Context: Danny's dad is talking to other dads about how fast her 12year-old son Danny has been growing recently.

   Glen: Danny is tall.
   Paul: That's right! He's been growing so fast since he turned 12 in the fall.

   (11.3) Context: Danny's mom is talking to other moms about NBA basketball.

   Sue: Danny isn't tall. He's very talented but just wouldn't stand a chance against those guys.

   Why do both Glen's use of 'Danny is tall' in his context and Sue's use of 'Danny isn't tall' in her context seem true? What is said with a comparative adjective in context depends on the comparison class implicitly referred to in context (Kennedy,

[1] Contrast lexical ambiguity with structural ambiguity; e.g. 'Visiting relatives can be boring', 'He saw that gasoline can explode' [(a) He saw an explosion of a can of gasoline, (b) He recognized the fact that gasoline is explosive]

[2] Not all variation in word-meaning is ambiguity. A defeasible test is as follows: For a given word W that has more than one meaning, if there is another natural language that has different words for those meanings, this is positive evidence for W's being ambiguous.

2007; Richard, 2004). (11.2) says, roughly, that Danny is tall for a 12year-old (true). (11.3) says, roughly, that Danny is tall for a NBA basketball player (false).

But there are still borderline cases for ‘tall for a 12year-old’ and ‘tall for an NBA basketball player’, no matter how specific the comparison class (‘tall for a 12year-old male Caucasian in 2014’).

11.2 The Sorites Paradox

11.2.1 The Argument

Let’s name the patches along the spectrum. For instance, call the patch at wavelength 700nm ‘patch 700’. Patch 700 is clearly red. But if patch 700 is red, then so is patch 699. After all, a difference of 1nm in wavelength cannot be a difference between red and non-red. If patch 699 is red, then so is patch 698. And if patch 698 is red, so is patch 697. And so on. Therefore, patch 500 is red.

Let ‘P_{700}’ be short for ‘Patch 700 is red.’ Then the argument form of the above reasoning can be displayed as follows:

Sorites paradox[3] (argument form 1)

$$\dfrac{\dfrac{\dfrac{P_{700} \qquad P_{700} \rightarrow P_{699}}{P_{699}} \qquad P_{699} \rightarrow P_{698}}{P_{698}} \qquad P_{698} \rightarrow P_{697}}{\ddots}$$

$$\dfrac{P_{501} \qquad P_{501} \rightarrow P_{500}}{P_{500}}$$

Sorites paradox (argument form 2)

Base step:	A one day old human being is a child.
Induction step:	If an n day old human being is a child, then that human being is also a child when it is $n + 1$ days old.
Conclusion:	Therefore, a 36,500 day old human being is a child.

[3]The Sorites paradox is also called the paradox of the heap (from Greek ‘soros’: ‘heap’).

11.2.2 Responses to the Sorites Paradox

- A **paradox** is an apparently valid argument which has apparently true premises but which has an absurd consequence.

- Argument form 1: P_{700} seems clearly true. ✓
  The $P_{700} \rightarrow P_{699}$ is also extremely plausible and so true. ✓
  P_{699} follows from the above by deductive reasoning. ✓
  $\vdots$
  But the conclusion, P_{500}, is *clearly* false. $\lightning$ Paradox

- Why is the conditional plausible?

  Tolerance Vague expressions are tolerant: small changes do not affect the applicability of the word. For instance, if two color patches are indistinguishable in color, then both or neither are red.

- **Ways of resolving a paradox**:

  1. **Reject one or more of the premises**

     Epistemicism (Sorensen, 1988; Williamson, 1994)
     Supervaluationism (van Fraassen, 1966; Kamp, 1975; Fine, 1975; Keefe, 2000)

  2. **Reject the reasoning as defective (e.g., invalid)**

     Many-valued logic and degrees of truth (Ł ukasiewicz and Tarski, 1930; Zadeh, 1965)
     Contextualism (Raffman, 1996; Soames, 1999; Shapiro, 2006)[4]

  3. **Bite the bullet and accept the conclusion**

     Vague concepts are flawed. (Unger, 1979)

- Accept the conclusion: Unger's view

  Sorites paradoxes show that vague concepts are deeply flawed: they commit us to absurdities. They are a flawed way to cut up the world. Flawed concepts are ones under which nothing can fall. So even if the world is full of stuff, there are no heaps, bald people, red things, mountains, etc.

- Reject a premise: Epistemicism

  – Vagueness is nothing but ignorance. Vague expressions have precise meanings: they cut the world along its sharp boundaries (the heaps are separated from the non-heaps by a sharp line). But we fail to know where those boundaries fall. Why? Our cognitive mechanisms require a margin for error to deliver knowledge. But even if I believe truly that patch 620 is red and patch 619 is not, I do not know this: for all I know, the concept I employ could be *red**, whose boundary is between wavelengths 620nm and 621nm.

[4]A good overview of contextualism can be found in Akerman and Greenough (2010).

  - In the Sorites argument (form 1), exactly one conditional premise is false (but we don't know which): the conditional such that the sharp boundary between red and non-red things falls between the patch in the antecedent and the patch in the consequent.

- We will get to supervaluationism in the next session.

- Reject the reasoning: Contextualism

  - Vague predicates like 'red' are context-sensitive: given a context, they have sharp cut-off points. But given the principle of tolerance, any context picked out is one in which the cut-off point isn't between the indistinguishable instances considered in the context.
  - The reasoning is a series of applications of modus ponens. Each step is valid, *given its context*. But the reasoning commits the fallacy of *equivocation*: the context is being changed along the series: there is no one context in which all the steps (and thus the argument as a whole) are valid.

11.3 Many-valued Logic

- **Principle of Bivalence** Every sentence (in a context of use) is either true or false.
- Two assumptions of classical logic:

  1. **Bivalence**: Every well-formed formula is either true (1) or false (0).
  2. **Validity** (semantic consequence) is **preservation of truth** from premises to conclusion.

- Many-valued logics break with both assumptions:

1. Not all sentences are either true or false. There are(finitely/infinitely) many truth values.

- A generalization from **classical logic**, which is defined by a structure $\langle \mathcal{V}, \mathcal{D}, \{f_c; c \in C\}\rangle$, where

  - $\mathcal{V}$ is the set of truth values $\{1, 0\}$
  - $\mathcal{D}$ is the set of **designated** values $\{1\}$—the set of values preserved in valid inferences
  - for every connective c, f_c is the truth function it denotes

- An **valuation** ν is a map from propositional letters to $\mathcal{V}$. It is extended to a map from all formulas into $\mathcal{V}$ by applying the appropriate truth functions recursively.

- Many-valued logic generalizes this structure, $\langle \mathcal{V}, \mathcal{D}, \{f_c; c \in C\}\rangle$:

  - $\mathcal{V}$ is the set of truth values, with any number of members ($\geqslant 1$)

- $\mathcal{D}$ is a subset of $\mathcal{V}$, the set of **designated** values
- for every connective c, f_c is the truth function it denotes

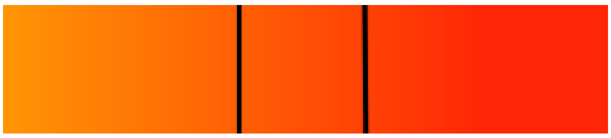
2. **Validity** (semantic consequence) is **preservation of designated truth values** from premises to conclusion.

$$\Gamma \models A \text{ iff there is no valuation } \nu \text{ such that for all } B \in \Gamma, \nu(B) \in \mathcal{D}, \text{ but } \nu(A) \notin \mathcal{D}.$$

$$A \text{ is a logical truth iff } \varnothing \models A, \text{ i.e. for every valuation } \nu(A) \in \mathcal{D}.$$

11.3.1 Three-valued Logic (Kleene, Łukasiewicz)

- There are three truth values: true, false, and *neither true nor false*. $\mathcal{V} = \{1, i, 0\}$.
- Think of 'neither true nor false' either as the lack of truth value (there is no matter of the fact as to whether the predicate applies to the object or not) or as a third truth value: *indefinite*
- Borderline cases (between the two black lines) are neither true nor false.



- What is the meaning of the logical connectives, given three truth values?

$f_\neg$	
1	0
i	i
0	1

$f_\wedge$	1	i	0
1	1	i	0
i	i	i	0
0	0	0	0

$f_\vee$	1	i	0
1	1	1	1
i	1	i	i
0	1	i	0

$f_\supset$	1	i	0
1	1	i	0
i	1	i	i
0	1	1	1

- These connectives of **Strong Kleene 3-valued logic, $\mathsf{K_3}$**, behave like classical connectives for the inputs 1 and 0. You can see this in the corners of all tables. For inputs i: understand *neither true nor false* as insufficient input to compute the (classical) value; e.g. if the input to $f_\supset$ is $\langle 1, i\rangle$, then the value is neither true nor false: given that the antecedent is 1, if the consequent were 0, the material conditional would be 0; if the consequent were 1, the conditional would be 1. But for input $\langle$0,i$\rangle$, there is sufficient information: if the antecedent of a material conditional is 0, then no matter the truth value of the consequent, the conditional is always 1.
- The set of designated values, $\mathcal{D}$, is just $\{1\}$. So validity is still preservation of truth.

- The Law of Excluded Middle (LEM) is not valid in K_3:

  $\not\models_{\mathsf{K}_3} A \vee \neg A$

  Counter-model: $\nu(A) = i$.
  You might think this is good for vagueness: for borderline cases, it's not true that either it is red or not red.
- The Law of Identity is also not valid:

  $A \not\models_{\mathsf{K}_3} A$

  Counter-model: $\nu(A) = i$.
- But we may want the Law of Identity to be a logical truth. We can get it by changing $f_\supset$ for the input $\langle i, i\rangle$:

$f_\supset$	1	i	0
1	1	i	0
i	1	**1**	i
0	1	1	1

- The 3-valued logic resulting from this change was originally given by Łukasiewicz and is often called **Ł3**.
- K_3 and **Ł3** share a solution to the Sorites paradox: Somewhere along the line of reasoning, we are going to get a case where P_{n+1} is a clear case and P_n is a borderline case. So the conditional '$P_{n+1} \supset P_n$' is computed from $\langle 1, i\rangle$ and thus is i (neither true nor false). However, we need not take ourselves to be committed to a premise that isn't true. (In K_3 all the steps in the borderline region – where both antecedent and consequent have the truth value i – receive the value i and thus need not be accepted.) So the Sorites reasoning fails.

11.4 Fuzzy Logics and Degrees of Truth

11.4.1 Semantics

- But why have a sharp boundary between clear cases and borderline cases? May it not be itself a vague matter whether something is a clear case or a borderline case? (Higher-order vagueness)
- Continuum-valued logics, or fuzzy logics, introduce a continuum of truth values: $\mathcal{V}$ is the set of real numbers between 0 (completely false) and 1 (completely true), $\{x : 0 \leqslant x \leqslant 1\}$ or $[0, 1]$.
  Truth comes in degrees. Insofar as patch 700 is more clearly red than patch 620, P_{700} has a higher degree of truth than P_{620}.

- The meaning of the connectives:

  $f_\neg(x) = 1 - x$

  $f_\wedge(x, y) = Min(x, y)$

  $f_\vee(x, y) = Max(x, y)$

  $f_\rightarrow(x, y) = x \ominus y$

  where *Min* means 'the minimum (lesser) of; *Max* means 'the maximum (greater) of'; and $x \ominus y$ is a function defined as follows:

  if $x \leqslant y$, then $x \ominus y = 1$

  if $x > y$, then $x \ominus y = 1 - (x - y) \quad (= 1 - x + y)$

- So as the truth value of a formula does down, the truth value of its negation goes up. A conjunction is as true as its least true conjunct. A disjunction is as true as its most true disjunct. For the conditional: If we reason from an antecedent to a consequent that has more truth than the antecedent, that's good reasoning; it receives 1. What if we reason from an antecedent to a consequent less true? That's less than perfect. How much less? That depends on the difference in truth between antecedent and consequent. The truth of the conditional goes down proportionally to the drop in degree of truth from antecedent to consequent.

- Note that for truth values 1 and 0, the truth functions behave just like the classical ones. And for the truth value 0.5, the truth functions behave just like i in Ł$_3$. The resulting logic, Ł, is a generalization of Łukasiewicz's 3-valued logic.

- **Validity**: What are the designated values in Ł? We could say that the only designated value is 1. But then we get valid reasoning as soon as a single premise is less than completely true. So what about all the values greater than 0.75? But this seems random. Why 0.75 and not 0.7 or 0.8? The important idea is this: An inference is valid iff under any valuation the conclusion is *at least as true* as the least true premise. That is, it isn't less true than the least true premise.

  For any set of truth values, there will always be a number that is less than or equal to every number in the set. For the set $X : \{0.21, 0.201, 0.2001, 0.20001, \ldots\}$, this number is 0.2. Call it *the greatest lower bound* of X, $Glb(X)$.

  Let $\nu[\Gamma]$ be $\{\nu(B) : B \in \Gamma\}$. Then

$$\Gamma \vDash_{Ł} A \text{ iff for all } \nu,\ Glb(\nu[\Gamma]) \leqslant \nu(A)$$

11.4.2 Fuzzy Logic and the Sorites Paradox: Reject Modus Ponens

- Modus Ponens (MP) is not valid in Ł: $\quad A, A \rightarrow B \nvDash_{Ł} B$

  It's true that if the premises all take truth value 1, so does the conclusion. But consider a valuation ν that assigns A a greater degree of truth than B; e.g. $\nu(A) =$

0.9 and $\nu(B) = 0.8$. Then the conditional takes the truth value $1 - 0.9 + 0.8 = 0.9$. So the minimum degree of truth of the premises is 0.9. But the conclusion is true only to degree 0.8. Hence, it's not the case that the conclusion is at least as true as the least true premise.

- The failure of Modus Ponens is also responsible for the invalidity of the Sorites argument: Take an arbitrary application of MP in the Sorites series, e.g. $P_{651}, P_{651} \to P_{650} \Rightarrow P_{650}$. Suppose $\nu(P_{651}) = 0.651$ and $\nu(P_{650}) = 0.650$. Then $\nu(P_{651} \to P_{650}) = 0.999$. So the least true premise has truth value 0.651. However, the conclusion has the lower truth value 0.650. So this step in the series is not valid, and thus the argument as a whole isn't.

11.5 Readings

- Priest (2008, §§7.1–7.4) on many-valued logics

  Sainsbury (1995, 40-47) on the Sorites paradox
- Optional: Priest (2008, §§11.1-11.4) on fuzzy logics

  Williamson (1994, §§4.1–4.6) on many-valued and fuzzy logics

12 Supervaluationism

To be written…

Bibliography

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

Akerman, J. and P. Greenough (2010). Vagueness and non-indexical contextualism. In S. Sawyer (Ed.), *New Waves in Philosophy of Language*. New York: Palgrave Macmillan.

Bell, J. L., D. DeVidi, and G. Solomon (2001). *Logical Options. An Introduction to Classical and Alternative Logics*. Peterborough: Broadview Press.

Bennett, J. (2003, June). *A Philosophical Guide to Conditionals*. New York and Oxford. Oxford University Press.

Edgington, D. (1995). On conditionals. *Mind 104*(414), 235–329.

Fine, K. (1975). Vagueness, truth and logic. *Synthese 30*(3/4), 265–300.

Gibbard, A. (1981). Two recent theories of conditionals. In W. Harper, R. Stalnaker, and G. Pearce (Eds.), *Ifs: Conditionals, Belief, Decision, Chance, and Time*, pp. 211–47. Dordrecht: Reidel.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts*, pp. 41–58. Page references are to the reprint in Grice (1989b), pp. 22–40. New York: Academic Press.

Grice, H. P. (1989a). Indicative conditionals. In *Studies in the Way of Words*, pp. 58–85. Cambridge, MA: Harvard University Press.

Grice, H. P. (1989b). *Studies in the Way of Words*. Cambridge, MA, and London: Harvard University Press.

Higginbotham, J. (1986). Linguistic theory and davidson's program in semantics. In E. LePore (Ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, pp. 29—48. Oxford: Blackwell.

Jackson, F. (1979). On assertion and indicative conditionals. *The Philosophical Review 88*(4), 565–89.

Kamp, H. (1975). Two theories about adjectives. In E. Keenan (Ed.), *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press.

Keefe, R. (2000). *Theories of Vagueness*. Cambridge: Cambridge University Press.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy 30*(1), 1–45.

Ł ukasiewicz, J. and A. Tarski (1930). Untersuchungen über den aussagenkalkül. *Comptes rendus des séances de la Societe des Sciences et des Lettres de Varsovie 3*(23), 12–21, 30–50. Translated into English by J.H. Woodger as ‘Investigations into the Sentential Calculus,’ in Corcoran, J .(ed), Logic, Semantics, Metamathematics, 2nd ed., Indianapolis: Hackett, 1983.

Lewis, C. I. (1917). The issues concerning material implication. *Journal of Philosophy, Psychology and Scientific Methods 14*(13), 350–56.

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

Lowe, E. (2002). *A Survey of Metaphysics*. Oxford: Oxford University Press.

Portner, P. (2009). *Modality*. Oxford: Oxford University Press.

Priest, G. (2008). *An Introduction to Non-Classical Logic. From If to Is* (2nd ed.). Cambridge: Cambridge University Press.

Raffman, D. (1996). Vagueness and context-relativity. *Philosophical Studies 81*(2-3), 175–92.

Ramsey, F. P. (1931). General propositions and causality. In *The Foundations of Mathematics*, pp. 237–55. London: Kegan Paul.

Read, S. (1994). *Thinking About Logic*. Oxford: Oxford University Press.

Richard, M. (2004). Contextualism and relativism. *Philosophical Studies 119*(1), 215–42.

Sainsbury, R. (1995). *Paradoxes* (3rd ed.). Cambridge: Cambridge University Press.

Shapiro, S. (2006). *Vagueness in Context*. Oxford: Oxford University Press.

Sider, T. (2010). *Logic for Philosophy*. Oxford: Oxford University Press.

Soames, S. (1999). *Understanding Truth*. New York, Oxford: Oxford University Press.

Sorensen, R. (1988). *Blindspots*. Oxford: Clarendon Press.

Stalnaker, R. C. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in Logical Theory*, pp. 98–112. Oxford: Blackwell.

Stalnaker, R. C. (1975). Indicative conditionals. *Philosophia 5*(3), 269–86; page references are to the reprint in Stalnaker (1999).

Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics*, pp. 315–32; page references are to the reprint in Stalnaker (1999). New York: Academic Press.

Stalnaker, R. C. (1999). *Context and Content*. Oxford: Oxford University Press.

Unger, P. (1979). There are no ordinary things. *Synthese 41*, 117–54.

van Fraassen, B. C. (1966). Singular terms, truth-value gaps, and free logic. *Journal of Philosophy 53*, 481–5.

von Fintel, K. (2006). Modality and language. In D. Borchert (Ed.), *Encyclopedia of Philosophy* (2nd ed.). Detroit: Macmillan Reference.

von Fintel, K. (2011). Conditionals. In K. von Heusinger, C. Maienborn, and P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 2. Berlin and New York: de Gruyter.

Williamson, T. (1994). *Vagueness*. London: Routledge.

Williamson, T. (1998). Bare possibilia. *Erkenntnis 48*(2-3), 257–73.

Zadeh, L. (1965). Fuzzy sets. *Information and Control 8*, 338–53.