Gödelian Explorations

Gary Mar

2030-01-01

Table of contents

Preface

This is a Quarto book.

To learn more about Quarto books visit https://quarto.org/docs/books.

1 Euler's Elegant Identity $e^{i\pi} + 1 = 0$ — a mathematical mystery tour

1.1 The Elegance of Euler's Identity

In 1749 Euler set forth in his *Introduction in Analysis Infinitorum* an identity considered to be among the most elegant in all of mathematics:

$$e^{i\pi} + 1 = 0.$$

Euler's equations, succinctly and elegantly combines five of the most fundamental constants of mathematics.

0
1
π
e
i

Euler's equation conceptually links numbers from three different number systems (the natural, real, and complex numbers) using three of the most fundamental arithmetic operations (addition, multiplication, and exponentiation.) Carl Friedrich Gauss (1777 - 1855), the "Prince of Mathematicians," reportedly said that anyone to whom Euler's identity was not immediately apparent would never become a first-class mathematician.

1.2 Demystifying Euler's Equation

> *"An ordinary genius is a fellow that you and I would be just as good as, if we were only many times better. There is no mystery as to how his mind works. Once we understand what he has done, we feel certain that we, too, could have done it. It is different with magicians... the working of their minds is for all intents and purpose incomprehensible. Even after we understand what they have done, the process by which they have done it is completely dark."*
>
> —Mark Kac (1914-1984) (quoted in Nahim, p. 9)

Euler's identity, it must be admitted, appears to be the work of a magician. However, whether or not one aspires to become a first-class mathematician, it is not immediately apparent how Euler's equation even makes sense.

Exponentiation is typically explained in terms of a repeated multiplication. For example, the number 2 raised to the exponent 3 is defined as the product of 2 multiplied by itself three times:

$$2^3 = 2 \times 2 \times 2 = 8.$$

Euler's identity uses transcendental numbers with exponentiation. The transcendental number π is defined as the ratio of the circumference to the diameter of a circle. Like the irrational number the $\sqrt{2}$, it is represented by an infinite non-periodic decimal; however, unlike, such *algebraic* irrational numbers,

π is not the solution or root of any polynomial such as $x^2 - 2 = 0$. The transcendental number e is defined as the base of an exponential function whose rate of change, or derivative, is equal to itself. In other words, e^x is the unique function such that it is its own derivative.

Now the transcendental numbers e and π are numbers in the neighborhood of the natural numbers 2 and 3, respectively, so it makes sense to think of exponentiation as purely numerical functions. (To remember the digits of π remember the question: *"May I have a large container of coffee right now?"* and "*To express e remember to remember a sentence to remember this"*.) Thinking of exponentiation as a function of real numbers, we can compute:

$$2.7182818284\cdots^{3.141592653\cdots} = 23.14069262\cdots.$$

However, Euler's identity also includes an exponential factor of the imaginary number i, where i is defined to be the positive square root of -1. Now i called *imaginary* number because no real number can be the root of a negative number. Real numbers are either positive or negative, and a positive number times a positive is positive and a negative number times a negative is also positive. Euler in his *Algebra* (1770) wrote:

> All such expressions as $\sqrt{-1}$, $\sqrt{-2}$, etc., are consequently impossible or imaginary numbers, since they represent roots of negative quantities, and of such numbers we may truly assert that they are neither nothing, nor greater than nothing, nor less than nothing, which necessarily constitutes them imaginary or impossible.

Obviously, there must be some meaningful mathematical way of thinking of exponentiation other than as a real-valued function.

The Nobel physicist Richard P. Feynman called Euler's equation "our jewel" and "the most remarkable formula in mathematics." [1] Feynman recounts how he, as a young boy, was fascinated and in his notebooks figured out for himself, rather than taking it on authority, the mystery of why Euler's identity was true. In Feynman's spirit of discovery, let's take a mathematical journey in which we figure out the meaning of Euler's remarkable identity for ourselves—not being content with mere proofs but instead taking the time to demystify the elements of Euler's beautiful equation.

[1] http://en.wikipedia.org/wiki/Euler%27s_formula#cite_note-2

The remarkable beauty of Euler's derivation of his identity derives not only from its deep *connections* of five fundamental mathematical constants, but also how it illustrates the *cognitive* strategies mathematicians deploy they are creating or discovering mathematical truths. The last phrase "creating or discovering mathematical truths" poses an important philosophical question: do mathematicians discover mathematical truths or do they creatively construct them?

On this magical mystery tour, we will discover that mathematics is more than merely crunching numbers, that it is more than presenting formal proofs, and that mathematics involves such creative discoveries as:

- *constructing* new number systems to solve previously impossible equations;
- *discovering* in geometric representations higher-level *symmetries*;
- *transforming* anomalous *singularities* into *systematic closure properties*;
- *generalizing* arithmetic operations by *expanding* their mathematical meanings;
- *unifying* conceptual domains by making bold conjectural *identifications.*

In demystifying Euler's equation, our goal is not to dispel its beauty. Indeed, by taking the time to understand how mathematicians create, we will finally come to appreciate—in a mathematically precise rather than a mystically vague way—the true beauty of Euler's equation.

1.3 A Concise, but Conceptually Incomplete, Proof

> *"It [Euler's identity] is absolutely paradoxical; we cannot understand it, and we don't know what it means, but we have proved it, and therefore we know it must be the truth."*
>
> —Benjamin Peirce (1809-1880)

The Harvard mathematician Benjamin Peirce, father of the pragmatist philosopher C. S. Peirce (1839-1914), is reported to have made the above remark after proving Euler's identity. This quotation reminds me of a joke told by Raymond Smullyan. A mathematics professor, teaching a class of mystified undergraduates, is writing a proof on the board and claims that the theorem is "trivial." A brave student asks why the proof is trivial. The professor turns back to the board and is lost in thought for five minutes, and then announces, "aha, yes, it is trivial!" Too often in a typical mathematics class, the lectures consist of presenting proofs to silent, of silenced, students [2] are left on their own to figure out what is going on mathematically by doing exercises.

[2] "Lecturing is the process of the transfer of knowledge from the notepad of the instructor to the notepads of the students, without ever touching the brains of either; that's atrocious." —Chas, M. (2023)

Perhaps the students are silent because they are afraid to "look stupid" or perhaps they have learned that asking questions can be dangerous. When some student asks the professor to explain some step of the proof, perhaps the professor "explains" the proof by merely repeating the same words, maybe talking more slowly or perhaps by filling in a few details. But the students can follow the steps but they still don't understand the mathematical ideas. This pedagogical impasse is created by the assumption that teaching mathematics is communicating formal structure or proofs. However, often understanding the mathematical ideas requires going behind the formal structure of proofs. Teaching is not merely proving theorems but communicating the mathematical ideas so that the truth of the theorem can be not only followed, but understood or comprehended, in a deeper, more intuitive, way.

So we have reason to be skeptical about Peirce's pronouncement:

> *"It [Euler's identity] is absolutely paradoxical; we cannot understand it, and we don't know what it means, but we have proved it, and therefore we know it must be the truth."*

Here's how Euler, in his *Introduction in Analysis Infinitorum*, derived his equation from the following trigonometric identity:

$$e^{\pm ix} = \cos(x) \pm \sin(x).$$

Here if we simply substitute π for the variable x, we have we have:

$$e^{i\pi} = -1 + 0,$$

from which Euler's identity immediately follows. This proof, while concise, is conceptually incomplete.

Beginning the proof with a puzzling *complex* trigonometric equation begs the conceptual question: what does exponentiation have to do with trigonometry? After all, both the *sine* and *cosine* functions are *periodic* and *bounded*, but the *exponential* function is *non-periodic* and *unbounded*. How then can an exponential function like e^x which rapidly expands to positive infinity be defined in terms of periodic trigonometric functions whose values are confined to the interval from 1 to -1?

1.4 Making Connections/Creating Mathematics

1.4.1 Connection: From Negativity to Complexity

TODO Grothendieck construction diagrams

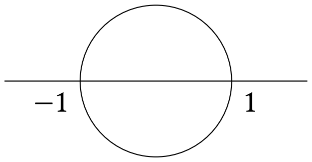
How does negativity, in particular -1, come about conceptually in mathematics? The negative integers are added to ("logically constructed from") the natural numbers when we want to have solutions to equations such as

$$x + 1 = 0.$$

No natural number can be a solution to this equation. (Proof: the left hand side of the equation is the successor of x, and according to Peano's third postulate, zero is not equal to the successor of any natural number.) By embedding the natural numbers in the richer context of integers with positive and negative numbers, there is a solution to every arithmetic equation involving subtraction, the inverse operation of addition.

When multiplication by -1 is represented geometrically in terms of the number line, the negative sign not only creates new *numbers*, but also adds the concept of *direction* on the number line. A minus sign indicates a change of direction—a *rotation* of 180 degrees—on the number line. Understood in these terms, it is perfectly understandable that a change of direction followed by another change of direction cancel each other out:

$$-1 \times -1 = 1.$$



The puzzle about justifying the law that a "negative times a negative is a positive" by finding an interpretation for multiplying debts is based on an inadequate conception of a negative number. A better way to think about multiplying by -1 is as a sign that indicates changing directions on the number line. To add -5 and -6 is to continue on the number line from the origin in the negative direction for a total of 11 units, and to multiply -5 by 6 to go $5 \times 6 = 30$ units in the negative direction. However, to multiply -5 by -6 is to multiply 5 by 6 changing directions twice and ending up on the positive direction.

The construction of imaginary numbers from real numbers can be done in a way that is completely analogous to the construction of negative numbers from natural numbers.

1.4.2 From -1 in Cartesian Coordinates to $\sqrt{-1}$ in Polar Coordinates

1.4.3 From the Derivative to Infinite Series

Consider Zeno's arrow paradox. If you look at the arrow in flight at an instant—or indivisible moment of time—it is indistinguishable from the arrow at rest. However, the flight of the arrow is just an infinite collection of moments of time. How can an infinite collection of motionless arrows become the arrow in motion? One way to begin to unravel Zeno's paradox is to develop the concept of *instantaneous velocity*. In order to distinguish the arrow in motion from the arrow at rest, we need to look at what is happening in the interval of time surrounding that instant of time. The notion of a non-zero *instantaneous velocity* is what is needed to solve Zeno's paradox of the arrow. This concept is captured by the mathematical concept of the derivative.

The derivative is one of the two fundamental ideas of calculus. The derivative is a mathematical model of change. Given a real-valued function $f(x)$, the derivative of f is the function $f'(x)$ which is the rate of change of f.

Conceptually, the derivative of a function is the rate of change of that function at a point in time. The *derivative of a function $f(x)$ is the limit of the function in the interval of dx around x as dx goes to zero:

$$\lim_{dx \to 0} \left(\frac{f(x+dx) - f(x)}{dx} \right)$$

Geometrically, the derivative of a function $f(x)$ is the slope of the tangent line to its graph. In graphical terms, if we have the graph of

$$y = f(x),$$

then $f'(x)$ is the slope of the line which is tangent to the graph of $f(x)$ at the point x is the derivative. We can write this using Leibniz's notation:

$$\frac{\mathrm{d}x}{\mathrm{d}y} = f'(x).$$

1.4.1.1 Negative Numbers: A Tale of Two Minuses

Minus times minus is plus.
The reason for this we need not discuss.

—W. H. Auden

There was skepticism about the negative numbers up until the 18th century. Negative numbers were introduced to have solutions to equations such as

$$x + 1 = 0.$$

No natural number can be a solution to this equation. According to one of Peano's postulates, 0 is not the successor of any natural number. We can expand the natural numbers by adding the negative integers. Negative numbers can, for example, be introduced to make it easy to calculate debts. If you owe 5 dollars to Abe and 10 dollars to Beatrice, but you only have 2 dollars, then you owe a total of 13 dollars—you have, so to speak, −13 dollars. If you owe five dollars five times over, then your total debt is $5 \times (-5) = -25$. However, does it make sense to ask how much you own if you owe 5 dollars a *negative* 5 times over?Negative numbers can be modeled by a pair of natural numbers that obey the following laws for addition and multiplication:

$$\begin{aligned} (a, b) + (c, d) &= (a + c, b + d) \\ (a, b) \times (c, d) &= (ac + bd, ad + bc). \end{aligned}$$

Even the great mathematician Euler commented (incorrectly) upon the perplexing law, which we now take for granted:

$$(-1) \times (-1) = 1.$$

Prove the above law that a minus times a minus is a plus follows from the *distributive law*

$$a(b + c) = ab + ac,$$

together with the facts

$$1 + (-1) = 0 \quad (-1) \times 0 = 0 \quad (-1) \times 0 = 0.$$

Extending the natural number line to the left, one generates the negative integers by multiplying the natural numbers by −1. Geometrically speaking, multiplication by −1 is a rotation of 180 degrees. The initially puzzling fact that $(-1) \times (-1) = 1$ corresponds to the geometric fact that two 180 degree rotations bring you back to the beginning. The integers are constructed as equivalence classes of ordered pairs of natural numbers and the negative integers are geometrically generated by multiplication by −1.

1.4.1.2 Complex Numbers: An Imaginary Journey

Imaginary i times i is minus one.
Let me tell you why it can be done!

There has also been widespread skepticism about imaginary numbers even up to the present day. Imaginary numbers were introduced to have solutions to equations such as

$$x^2 + 1 = 0.$$

No real number r can be a solution to this equation. A real number is either positive or negative, but a positive times a positive is a positive, and a negative times a negative is also a positive and so $r \times r \neq 1$. Negative numbers can be seen as involving a rotation of 180 degrees. Complex numbers can be introduced to make it easier to calculate rotations without tedious trigonometry.What does $i \times i = -1$ mean in terms of rotation? Multiplying by −1 is a rotation of 180 degrees. Geometrically speaking, multiplication by *i is a counter-clockwise rotation of 90 degrees or $\frac{\pi}{2}$ radians. Two successive rotations $i \times i$ equal to −1.Complex numbers can be modeled by a pair of real numbers that obey the following laws for addition and multiplication:

$$\begin{aligned} (a, b) + (c, d) &= (a + c, b + d) \\ (a, b) \times (c, d) &= (ac - bd, ad + bc). \end{aligned}$$

Euler wrote "All such expressions $\sqrt{-1}$, $\sqrt{-2}$, etc. are consequently impossible or imaginary numbers...." (*Algebra*, 1770). We have the following cycle of facts:

$$i^1 = i \quad i^2 = -1 \quad i^3 = -i \quad i^4 = 1.$$

If $c = a + ib$, then the *modulus* of c is $\sqrt{a^2 + b^2}$. By the Pythagorean theorem, this is the distance from the origin to the point (a, b). The *argument* of a complex number $c = a + ib$ is the *angle* formed by the positive real x-axis and the line from $(0, 0)$ to (a, b).Every complex number can be represented by a modulus and an argument using polar coordinates. Multiplication of complex numbers in this form

$$\begin{aligned} c &= r[\cos(\theta) + i\sin(\theta)] \\ d &= s[\cos(\rho) + i\sin(\rho)] \end{aligned}$$

amounts to this: (1) the *modulus* of the product is the *product* of the moduli; and (2) the *argument* of the product is the *sum* of the *arguments, i.e,*

$$c \times d = r \times s[\cos(\theta + \rho) + i\sin(\theta + \rho)].$$

Multiplication by a unit length is equivalent to pure rotation.

In our example, the derivative of $f(x) = x^2$ is computed as follows:

$$\lim_{dx\to 0}\left(\frac{(x+dx)^2 - x^2}{dx}\right) = \frac{x^2 + 2xdx - x^2}{dx} = 2xdx.$$

So the derivative, or rate of change, of the function $f(x) = x^2$ is the function $f(x) = 2x$.

Here's a chart of some common derivatives:

Function	***Derivative***	***Comment***
c ($any\ constant\ number$)	0	A constant, by definition, doesn't change.
x^n	$nx^{n\text{-}1}$	The iterated derivative of x^n is related to n!
e^x	e^x	The base e of the natural exponential function is chosen to be such that the rate of change of the function is equal to itself.
$\ln(x)$	$1/x$	The natural logarithm is the inverse of the exponential function.
$sin(x)$	$cos(x)$	The derivative of $sin(x)$ is $cos(x)$
$cos(x)$	$-sin(x)$	The derivative of $cos(x)$ is $-\ sin(x)$

The Harvard mathematician and satirical songwriter Tom Lehrer put the definition of the derivative to music to a tune

There'll Be Some Changes Made.

You take a function of x and you call it y,
Take any x_0 that you care to try.
You make a little change and call it δx,
The corresponding change in y is what you find nex',
And then you take the quotient, and now carefully
Send δx to zero and I think you'll see
That what the limit gives us, if our work all checks, Is what we call $\frac{dy}{dx}$. Is jus$\frac{dy}{dx}$!

Derived Functor Rap

This is a new rap on the oldest of stories—
Functors on abelian categories.
If the functor is left exact
You can derive it and that's a fact.
But first you must have enough injective
Objects in the category to stay active.
If that's the case no time to lose;
Resolve injectively any way you choose.
Apply the functor and don't be sore—
The sequence ain't exact no more.
Here comes the part that is the most fun, Sir,
Take homology to get the answer.
On resolution it don't depend:
All are chain homotopy equivalent.
Hey, Mama, when your algebra shows a gap
Go over this Derived Functor Rap.
—P. Bressler, *Derived Functor Rap*, 1988

1.4.4 From Infinite Series to Power Series

The answer to this question involves thinking of exponentiation not merely as a *power function* that produces a number but as a *power series.* We have already encountered infinite series when discussing Zeno's paradoxes. For example, Zeno's dichotomy paradox involved the *geometric* series

$$\sum_{n=1}^{\infty}\frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots + \frac{1}{2^n} + \cdots.$$

Figuring out the sum of such an infinite geometric series depended on the observation that multiplying the series by the common factor of ½ results in truncating the first term:

$$\begin{aligned} X &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots + \frac{1}{2^n} + \cdots \\ \frac{1}{2}X &= \quad\ \ \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots + \frac{1}{2^n} + \cdots. \end{aligned}$$

So by subtracting half the series from the entire series, we simply have:

$$X = 1$$

Another way of seeing this result geometrically is the following "proof without words":

$$X = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots + \frac{1}{2^n} + \cdots.$$

The same idea for summing an infinite geometric series works, not only for $r = \frac{1}{2}$, but for any r whose absolute value is < 1. In general, the infinite sum exists when $|r| < 1$,

$$\sum_{n=0}^{\infty}\frac{a}{r^n} = a + \frac{a}{r} + \frac{a}{r^2} + \frac{a}{r^3} + \frac{a}{r^4} + \frac{a}{r^5} + \ldots + \frac{a}{2^n} + \cdots = \frac{a}{1-r}.$$

Now the idea of *power series* comes about by replacing *numbers* with a variable resulting in the sum of an infinite series of *power functions.* Generally, a power series is of the form

$$\sum_{n=0}^{\infty} a_n z^n = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + \cdots + a_n z^n + \cdots,$$

where $a_0, a_1, \dots, a_n, \dots$ is some sequence of numbers. The reason for representing a function as a power series is that it is easy to compute the derivative of such a series term by term using the fact that the derivative of

$$f(x) = a_n x^n$$

is

$$f'(x) = (n-1)a_n x^{n-1}.$$

Now remember that the defining characteristic of the exponential function is that the derivative of e^x is just e^x. How can we create a power series in which the derivative of the series, taken term by term, results in the very same series?

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots + \frac{x^n}{n!} + \cdots$$

The derivative of $1 = 0$;
The derivative of $x = 1$;
The derivative of $\frac{x^2}{2!} = \frac{2x}{2} = x$;
The derivative of $\frac{x^3}{3!} = \frac{3x}{3\times 2\times 1} = \frac{x^2}{2!}$;
The derivative of $\frac{x^2}{2!} = \frac{2x}{2} = x$;
The derivative of $\frac{x^3}{3!} = \frac{3x}{3\times 2\times 1} = \frac{x^2}{2!}$; *The derivative of $\frac{x^{n+1}}{(n+1)!} = \frac{(n+1)x^n}{(n+1)n!} = \frac{x^n}{n!}$.

So taking the term-wise derivative is the infinite series itself! Now this is precisely the defining feature of the exponential function:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots + \frac{x^n}{n!} + \cdots;$$

this power series can be used to estimate the value of e.

The following spreadsheet, computing the value of e for the first 12 terms, gives an approximation accurate to the first 9 decimal places.

	Power Series Approximation of *e*	
	1	1
1	1	2
2	0.5	2.5
3	0.166666667	2.6666666667
4	0.041666667	2.7083333333
5	0.008333333	2.7166666667
6	0.001388889	2.7180555556
7	0.000198413	2.7182539683
8	2.48016e-05	2.7182787698
9	2.75573e-06	2.7182815256
10	2.75573e-07	2.7182818011
11	2.50521e-08	2.7182818262
12	2.08768e-09	2.7182818283

Euler's power series gives a quick and accurate estimate of the value of e because the factorials in the denominator grow quickly diminishing the contribution of the successive terms, which quickly approaches zero. However, the value of power series does not necessarily lie in computing approximate values.

Leibniz was proud of his discovery (1674) of the quadrature or Leibniz–Gregory series (Gregory discovered the series in 1671):

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \cdots$$

In an infinitely-approximating way, Leibniz's series provides a "solution" to the classic geometrical problem of squaring the circle. But Leibniz's series is practically useless in calculating the digits of π. Not only does it converge extremely slowly but it does not yield digits that are *stable.* For example, if you compute the first five terms of the series then the partial sum is 0.8349206, which is within $\frac{1}{11}$ of the true value of $\frac{\pi}{4}$, but not a single digit of the partial sum is correct!

Similarly, the mathematical value of the exponential power series lies not so much as its usefulness in *calculating* the digits of e, but in itself providing *conceptual* insight into how to generalize exponentiation to unify two formerly distinct mathematical domains. The catalyst for the intuitive leap is the curious way in which the power series for the trigonometric functions of sine and cosine appear to be "contained" in the exponential power series:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \cdots + \frac{x^n}{n!} + \cdots,$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \cdots,$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \cdots.$$

The terms of the *sine* power series are the *odd* powers and the terms of the *cosine* power series are the even powers. The question is how to account for the alternation of + and − signs. And to repeat, the *sine* and *cosine* functions are *periodic* and *bounded*, but the *exponential* function is *non-periodic* and *unbounded.* How then can an exponential function like e^x which rapidly expands to positive infinity be defined in terms of periodic trigonometric functions whose values are confined to the interval from 1 to −1?

To solve this part of the puzzle we need to explore the undue prejudice directed against imaginary numbers.

1.5 A Conceptual Unification: Understanding Euler's Identity

1.6 From Truth Values to Truth Vectors.

Listing 1.1

1.7 From Euler's Equation to the Riemann Zeta Function

1.8 From the Riemann Zeta Function to Places Yet Unknown

2 Modal Logic and Binary Relations

Relation Domain		
total	$\forall x \forall y Rxy$	All possible arrows exist.
empty	$\forall x \forall y \neg Rxy$	There are no arrows.
vacuous	$\forall x \forall y (Rxy \rightarrow x = y)$	Only loops are arrows.
Serial		
serial (right)	$\forall x \exists y Rxy$	Every element is a tail of an arrow.
serial (left)	$\forall x \exists y Ryx$	Every element is a target of an arrow.
Reflexive		
totally reflexive	$\forall x Rxx$	All elements have loops.
weakly reflexive	$\forall x[\exists y(Rxy \vee Ryx) \rightarrow Rxx]$	Targets and tails have loops.
irreflexive	$\forall x \neg Rxx$	There aren't any loops.
co-reflexive	$\forall x \forall y (Rxy \rightarrow Rxx)$	Any tail has a loop.
non-reflexive	Neither reflexive nor irreflexive	
Symmetric		
symmetric	$\forall x \forall y (Rxy \rightarrow Ryx)$	All arrows are double arrows.
asymetric	$\forall x \forall y (Rxy \rightarrow \neg Ryx)$	No arrow is a double arrow.
antisymmetric	$\forall x \forall y \forall z (Rxy \wedge Ryx \rightarrow x = y)$	Only loops are double arrows.
non-symmetric	Neither symmetric nor asymmetric	
Transitive		
transitive	$\forall x \forall y \forall z (Rxy \wedge Ryz \rightarrow Rxz)$	Any indirect path has a shortcut.
intransitive	$\forall x \forall y \forall z (Rxy \wedge Ryz \rightarrow \neg Rxz))$	No indirect path has a shortcut.
non-transitive	Neither transitive nor intransitive	
Euclidean		
Euclidean (right)	$\forall x \forall y \forall z (Rxy \wedge Rxz \rightarrow Ryz)$	Anyone's beloveds love themselves and one another.
Euclidean (left)	$\forall x \forall y \forall z (Ryx \wedge Rzx \rightarrow Ryz)$	Anyone's lovers love themselves and one another

3 ‘Must’ and ‘Might’ — The Modal Logic of Necessity and Possibility

3.1 Modes of Truth and Modal Logics

Historically, notions like *necessity*, *possibility*, *impossibility*, and *contingency* were thought of as modes of truth or ways in which a proposition could be true or false. *Modal logic* began as the study of the logic of these modes of truth.

Aristotle, in Chapter 9 of *De Interpretatione*, discusses modality in his famous example of the sea battle. Suppose the sea battle will be fought tomorrow. Then it was true yesterday that it would be fought tomorrow. So if all past truths are necessarily true, then it is necessarily true now that the battle will be fought tomorrow. A similar argument holds on the supposition that the sea battle will not be fought tomorrow. Aristotle proposed solving this problem of *logical fatalism* by denying that future contingent propositions have definite truth-values.

Using the ‘$\Box$’ for ‘it is necessary that’, the principle that all necessary truths are in fact

$$\Box P \rightarrow P, \tag{T}$$

but adding its converse that all truths are necessary truths:

$$P \rightarrow \Box P \tag{V}$$

collapses the notions of truth and necessary truth.

Medieval philosophers, concerned with such theological issues as articulating the nature of the Trinity, appealed to such modal notions as essence and accident, contingency and necessity in their labyrinthine theological reflections.[3] Akin to the problem is logical fatalism is the problem of *theological fatalism*: the problem of reconciling divine foreknowledge and human freedom. Saint Augustine (354–430) in his treatise *On the Free Choice of Will* considers an argument for *theological fatalism* proposed by Evodius. Evodius argued that “God foreknew that man would sin, that which God foreknew must necessarily come to pass.” We may set forth this argument for theological fatalism for a particular case as follows:

> If God knew that Adam would sin, then, necessarily Adam sinned.
> God knew that Adam would sin (because God is omniscient).
> Therefore, Adam necessarily sinned.

St. Thomas Aquinas (1225–1274) in his *Summa Contra Gentiles* (part I, chapter 67) criticized this kind of argument as resting on an amphiboly. The critical first premise “if God knew Adam would sin, then, *necessarily*, Adam sinned” is ambiguous between

> (1*a*) It is necessarily the case that if God knows that Adam will sin then Adam will sin.
> (1*b*) If God knows that Adam will sin, then it is necessary that Adam will sin.

Aquinas called (1*a*) the *necessity of the consequence* contrasting it with (1*b*) the *necessity of the consequent*. Using the ‘$\Box$’ to abbreviate ‘it is necessary that’, the difference between these two can be made more perspicuous with symbols:

$$\Box(P \rightarrow Q) \tag{1a}$$

$$(P \rightarrow \Box Q) \tag{1b}$$

[3] The theological, if not the political, roots Great Schism of 1054, which can be traced back to a disagreement about the modalities of the Persons of the Trinity. The Nicean Creed (325) uses the term *homoousios* (from the Greek *homo* = ‘same‘ and *ousios* = ‘essence’ or ‘substance’), in contrast to *homoiousios* (from the Greek *homoi* = ‘similar’) making the solitary *i* the jot and tittle of Nicean Creedal Orthodoxy. The Greek Church preferred the latter term since the former had been used by the Syrian Bishop of Antioch to espouse *modal monarchism*, the heresy that the Heavenly Father, Resurrected Son and Holy Spirit are not three distinct Persons, but are rather different *modes* or *aspects* of one monadic God perceived by believers as distinct persons. The Latin Church adopted the former siding with Athanasius against the heretic Arius, who denied that Jesus was coequal and co-eternal with the Father. The Second Council of Nicea (381), among other changes, inserted the word *filioque* (from the Latin *filio* = the son, and *que* = “and”) into the Nicean-Constantinopolitan Creed and the Latin mass. In Latin theology, the three Persons of the Trinity are logically distinguished by the formal relations of “proceeding from” citing such proof texts as Phil. 1:9, Titus 3:6, Acts 2:33. Orthodox theology, citing the words of Jesus in proof texts such John 15:26, regarded the insertion of *filioque* into the Nicean-Constantipolitan creed, is as Semi-Sabellianism. The great church historian Pelikan (1988: 90) opined: “If there is a special circle of the inferno described by Dante reserved for historians of theology, the principal homework assigned to that subdivision of Hell for at least the first several eons of eternity may well be a thorough study of all the treatises… devoted to the inquiry: Does the Holy Spirit proceed from the Father only, as Eastern Christendom contends, or from both the Father and the Son as the Latin Church teaches?” In 1989 Pope John Paul II and Patriarch Demetrius knelt together in Rome and recited the Nicene Creed without the *filioque*.

Pelikan, J. 1988. *The melody of theology: A philosophical dictionary*. Harvard University Press.

Solving the famous theological problem of reconciling divine foreknowledge with human freedom may turn on exposing ambiguities of this sort.

Perhaps the most famous theological application of modal logic is Saint Anselm's modal ontological argument. According to Saint Anselm (1033–1109), it follows from God's nature that it is necessary that God exists if God exists at all. Moreover, this conditional itself, being a conceptual truth, is itself necessarily true. We then have the following argument:

> Necessarily, if God exists, then God necessarily exists.
> It is possible that God exists.
> Therefore, God (actually) exists.

Using '$\diamond$' for 'it is possible that', the above argument can be symbolized:

$$\Box(G \to \Box G). \quad \diamond G \quad \therefore G \tag{2}$$

The question of whether Anselm's argument is valid became a precise question when various systems of modal logic were proposed and developed in the 1960s.

Gottlob Frege (1848–1925), the inventor (or discoverer) of modern predicate-quantifier logic, relegated modality to autobiographical information about the speaker, and for many years logicians only investigated extensional logic.

One of the most puzzling validities, at least to the beginning logical students is known as *Lewis's Dilemma*:

$$\mathrm{P} \land \neg\mathrm{P} \to \mathrm{Q},$$

which states *"a contradiction implies anything"*[4]. This implication follows from the inference rules of simplification, addition, and *modus tollendo ponens*[5], which are themselves not particularly puzzling.

[4] Or *"ex falso quodlibet"*, in Latin.

[5] *"mode that denies by affirming"*

1.	~~*Show*~~	$\mathrm{P} \land \neg\mathrm{P} \to \mathrm{Q}$	6, CD
2.		$\mathrm{P} \land \neg\mathrm{P}$	Assume (CD)
3.		P	2, S
4.		$\neg\mathrm{P}$	2, S
5.		$\mathrm{P} \lor \mathrm{Q}$	3, ADD
6.		Q	5, 4 MTP

The following theorems are known as the *paradoxes of material implication*:

→ law of denying the antecedent

$$\neg\mathrm{P} \to (\mathrm{P} \to \mathrm{Q}) \tag{T18}$$

→ law of affirming the consequent

$$\mathrm{Q} \to (\mathrm{P} \to \mathrm{Q}) \tag{T2}$$

→ conditional excluded middle

$$(\mathrm{P} \to \mathrm{Q}) \lor (\mathrm{Q} \to \neg\mathrm{Q}) \tag{T58}$$

→ Consequentia Mirabilis, "admirable consequence" [Cantor's Δ?]

$$(\neg\mathrm{P} \to \mathrm{P}) \to \mathrm{P} \tag{T114}$$

C. I. Lewis investigated modal logic in order to find a stricter form of the conditional which would not result in such paradoxes. Lewis defined *strict implication* $\mathrm{P} \Rightarrow \mathrm{Q}$ (read "P *strictly implies* Q") by combining modality with the truth-functional conditional:

$$\mathrm{P} \Rightarrow \mathrm{Q} \quad := \quad \Box(\mathrm{P} \land \neg\mathrm{Q}),$$

or alternatively,

$$\mathrm{P} \Rightarrow \mathrm{Q} \quad := \quad \Box(\mathrm{P} \to \mathrm{Q}).$$

The notion of strict implication was characterized by such axioms as:

The philosopher Leibniz (1646–1716) explicitly invoked that language of possible worlds to explain the difference between necessary and contingent truths. What is logically or necessarily true are those truths truth in *all* possible worlds, whereas a contingent truth is one that is true in *some* possible worlds.

Drawing upon this logical connection between universal and existential quantification and the modal notions of necessity and possibility, we obtain a modal version of the classical Aristotelian Square of Opposition and the duality of modal laws known as the laws of modal negation.

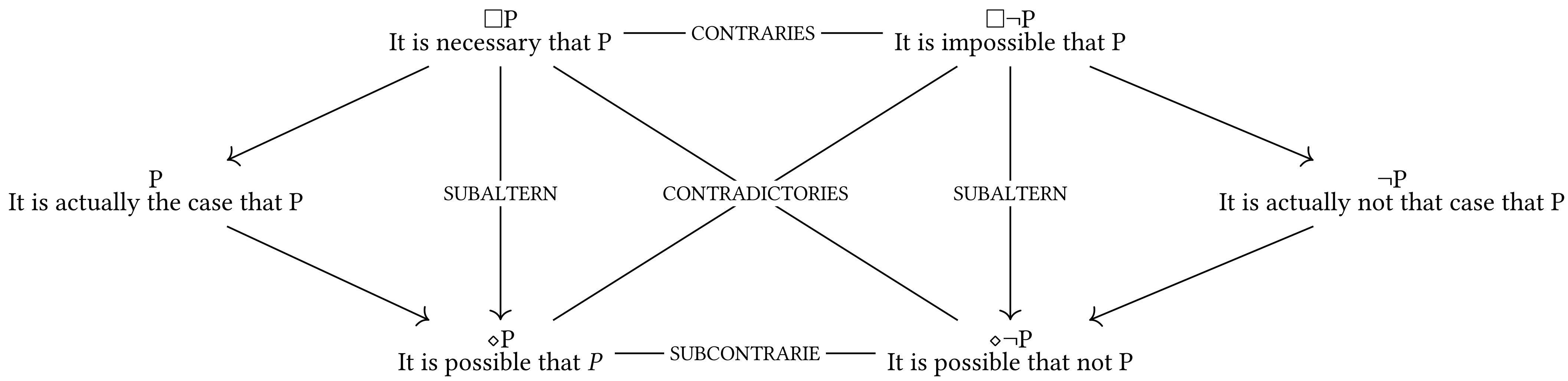


Figure 3.1: An Aristotelian Diamond of Opposition

In the modern development of modal logic, logicians noticed that a host other phenomena—such as deontic notions of obligation and permissibility, epistemic notions of knowledge and belief, as well as temporal operators—share these logical relations and hence can be represented as modal logics.

In *deontic logic*, □ is read "it is morally obligatory that" and ⋄ is read "it is morally permissible that". Kant's maxim that "ought implies can", that is, whatever is obligatory is permissible, is captured by modal axiom **(D)**:

$$\Box P \rightarrow \Diamond P. \qquad \textbf{(D)}$$

In *epistemic logic*, □ is read for some subject *S* "it is known that" and ⋄ is read "it is believed that". Some modal axioms for epistemic logic that have been considered include:

$$\Box(P \rightarrow Q) \rightarrow (\Box P \rightarrow \Box Q) \qquad \textbf{(K)}$$

← logical omniscience

$$\Box P \rightarrow P \qquad \textbf{(T)}$$

← law of affirming the consequent

$$\Box P \rightarrow \Box\Box P \qquad \textbf{(4)}$$

← law of affirming the consequent

$$\neg\Box P \rightarrow \Box\neg\Box P \qquad \textbf{(E)}$$

← law of affirming the consequent

The axiom **(K)** expresses logical omniscience insofar as this axiom requires that the knowledge of an agent is closed under *modus ponens*; and hence such knowers know all the logical consequences of their knowledge.

Axiom **(T)** states the truism that whatever is known is true. Notice that this axiom would be too strong for deontic logic insofar as an action's being obligatory doesn't imply that the agent actually performs that action.

Axiom **(4)** expresses a high degree of *positive introspective knowledge*: if someone knows P, then she knows that she knows that P. Axiom **(E)** on the other hand, expresses a high degree of *negative* introspective knowledge: if someone doesn't know that P, then he knows he doesn't know P. This axiom is contrary to the experience of Socrates: as the gadfly of Athens, Socrates found through his questioning that many of his fellow Athenians did not know what they were talking about but also didn't know that they didn't know. The gadfly of Athens believed his vocation was to sting his fellow Athenians into the awareness they were own ignorance, a service for which they did not always show adequate appreciation.

> A Classical Theorem, but in Intuitionism
>
> **Theorem.** There exist two irrational numbers x and y such that x^y is irrational.
>
> *Proof.* Consider
>
> $$\sqrt{2}^{\left(\sqrt{2}^{\sqrt{2}}\right)} = \sqrt{2}^{(\sqrt{2}\cdot\sqrt{2})} = \sqrt{2}^{2} = 2,$$
>
> which is rational.
> The number $\sqrt{2}^{\sqrt{2}}$ is either rational or irrational.
>
> - If it's rational, then $x = y = \sqrt{2}$ are both irrational yet x^y is rational.
> - If $\sqrt{2}^{\sqrt{2}}$ is irrational, then $x = \sqrt{2}^{\sqrt{2}}$ and $y = \sqrt{2}$ are both irrational and x^y is rational.
>
> Either way, there exists x and y such that x^y is rational. $\square$
>
> Intuitionists reject this classical argument by separation of cases because it does not actually construct the numbers x and y such that
>
> xy is irrational. The idea behind intuitionistic logic is that the connectives are reinterpreted as involving a kind of provability.
>
> In case you're curious, it is actually known that $\sqrt{2}^{\sqrt{2}}$ is irrational, due to the Gelfond–Schneider theorem which verifies that it usually takes significantly more effort to convince an intuitionistic mathematicians than a classical one.

The *temporal logic* or Diodorian temporal logic was studied by the logician A. N. Prior (1914–1969). To model temporal language, we introduce a pair of modal operators for the future and a pair of modal operators for the past.

$\Box$	It *is always going* [i.e., in *all* futures] to be the case that
$\Diamond$	It *will>* [i.e., in *some* future] be the case that
$\Box$	It *has always been* [i.e., in *all* pasts] the case that
$\Diamond$	It *was once]* [i.e., in *some* past or *"once upon a time"*] the case that

The axioms for minimal tense logic include version of Axiom **(K)** for the two necessity operators:

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi) \qquad \textit{Whatever has always followed from what always has been, always has been.} \tag{K$\Box$}$$

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi) \qquad \textit{Whatever will always follow from what always will be, always will be.} \tag{K$\Box$}$$

It also contains two axioms concerning the interaction of the past and future that has the form of the so-called Brouwersche axiom with alternating valences:

$$\varphi \to \Box\Diamond\varphi \qquad \textit{What is, will always have been} \tag{B$\Box\Diamond$}$$

$$\varphi \to \Box\Diamond\varphi \qquad \textit{What is, has always been going to happen} \tag{B$\Box\Diamond$}$$

Becker, Oskar. 1930. Zur logik der modalitäten.

The Brouwersche **(B)** axiom was so-named by the logician Oskar Becker (Becker 1930) after the charismatic Dutch mathematician L. E. J. Brouwer (1881–1966), who championed a philosophy of mathematics known as *intuitionism.* It happens that when the $\Diamond$ can be paraphrased as $\Diamond\neg\Diamond$, the resulting axiom has the form of the acceptable form of double negation in intuitionistic logic:

$$\varphi \to \neg\Diamond\neg\Diamond\varphi \tag{B}$$

According to intuitionism, mathematical objects do not exist as eternal Platonic objects but are constructions in intuition. Intuitionists read the propositional connectives as involving not merely truth, but proof, and so they rejected such classical forms of reasoning as *reductio ad absurdum* and theorems such as the *law of excluded middle.* Intuitionists reject the following mathematical proof.

Gödel, Kurt. 1931. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i. *Monatshefte für Mathematik Und Physik* 38. 173–198.

Around the 1970s it was noticed that the famous incompleteness theorems (Gödel 1931) were propositional in character and that their logic could be captured in propositional modal logic. These modal provability logics added to Axiom **(K)** the following axiom, known as the Gödel-Löb axiom or also the well-ordering axiom.

$$\Box(\Box\varphi \to \varphi) \to \Box\varphi \tag{W}$$

Gödel, Kurt. 1933b. The present situation in the foundations of mathematics. *Collected Works* 3. 45–53.

Modal Provability Logics proliferated from the 1950s–1970s, but the genesis of the idea goes back to a short note of Gödel (1933b) in which he noted that intuitionistic truth is defined in terms of proof since provability is a kind of necessity. The above axiom can be read as a kind of soundness theorem.

> *if it is provable that φ being provable implies it is true, then φ is provable.*

Later we will show how to use a modal provability logic to exhibit the propositional logic of key parts of Gödel's First and Second Incompleteness Theorems.

In contemporary logic, modal logic has grown beyond these philosophical origins and is at the interface of a number of disciplines including the studies information flow and dynamics, game theory, and computability.

Exercises

(1) Symbolize the following modal arguments.
   (A) Eratosthenes must either be in Syene or Alexandria. Eratosthenes cannot be in Syene. Therefore, Eratosthenes must be in Alexandria.
   (B) Assume that justice can be defined as paying your debts and telling the truth. Then it is *morally obligatory* for Cephalus to comply to a madman's request that Cephalus return a borrowed sword and that Cephalus tell the truth about the whereabouts of a friend whom the madman wants to kill with the sword. However, if this act is *morally obligatory*, then it is *morally permissible*. However, it is *morally impermissible* (or *morally forbidden*) for Cephalus to comply. So it isn't *morally obligatory* for Cephalus to comply. Justice, therefore, cannot be defined as paying your debts and telling the truth.
   (C) It is conceivable that I am having experiences qualitatively identical to those I am having now on the supposition that I am being deceived by an evil genius. If that is conceivable, then I do not have indubitable knowledge that the external world exists.

(2) Johan van Benthem (Van Benthem 2010: 12) was asked to symbolize the philosophical claim that "nothing is absolutely relative". He came up with the following:

$$\neg\Box(\diamond\varphi \land \diamond\neg\varphi).$$

Use familiar equivalences from propositional logic and the modal negation laws to show that this symbolization is equivalent to

$$\Box\diamond\varphi \to \diamond\Box\varphi \qquad \textbf{(M)}$$

Van Benthem, Johan. 2010. *Modal logic for open minds.* Stanford, CA: Centre for the Study of Language & Information.

← *McKinsey's Axiom*

(3) Match the following symbolizations with the best corresponding translation below.

symbolization	translation
$\diamond\Box$P	*It was always the case that it will sometime be the case that* P.
$\diamond$P $\to$ $\diamond$P	*It will sometime be the case that it was once the case that* P.
$\Box\diamond$P	*Whatever will always be, will be.*
$\diamond\diamond$P	*Once upon a time, it was always the case that* P.

3.2

3.3

3.4

3.5

3.6

3.7 Possible World Semantics

The Leibnizian idea of characterizing necessity and possibility in terms of truth in all or some possible worlds was given an elegant formalization by (Kripke 1959; 1963) when he was only a teenager. According to Leibniz, a sentence is *necessary* if it is true in *every* possible world, and a sentence is *possible* if it is true is *some* possible world. Kripke showed that by placing very natural conditions on a relation of *relative possibility* or *accessibility* on a set of possible worlds, the various systems of modal logic could be validated.

Kripke, Saul. 1959. A completeness theorem in modal logic. *Journal of Symbolic Logic* 24(1). 1–14. DOI: https://doi.org/10.2307/2964568

Kripke, Saul. 1963. Semantical analysis of modal logic i normal modal propositional calculi. *Mathematical Logic Quarterly* 9(5-6). 67–96. DOI: https://doi.org/10.1002/malq.19630090502

Intuitively, a possible world tells us for each sentence letter whether it is true or false in that world. Stripping away inessentials, we can represent a possible world by a subset of sentence letters. A *modal structure* $\mathbb{M}$ is an ordered triple $\langle W, R, \alpha\rangle$, where W is a set[6] of possible worlds, $R \subseteq W \times W$ is a relation known as the *accessibility relation* or the *relative possibility relation*, and α is a distinguished element of W known as the *actual world*.

We can exhaustively characterize the notion of

$$\beta \vDash \varphi,$$

the *truth of a sentence in a possible world* β, by

(4) If φ is a sentence letter S, then
$$\beta \vDash S \iff S \in \beta,$$
i.e. S is a member of β;

(5) If φ is ψ, then
$$\beta \vDash \neg\psi \iff \beta \nvDash \psi,$$
i.e., it is not the case that $\beta \vDash \psi$;

(6) (a) If φ is $(\psi \wedge \chi)$, then
$$\beta \vDash (\psi \wedge \chi) \iff \beta \vDash \psi \ \&\ \beta \vDash \chi,$$
i.e., both $\beta \vDash \psi$ and $\beta \vDash \chi$;

(b) If φ is $(\psi \vee \chi)$, then
$$\beta \vDash (\psi \vee \chi) \iff \beta \vDash \psi \parallel \beta \vDash \chi,$$
i.e., either $\beta \vDash \psi$ or $\beta \vDash \chi$, or both;

(c) If φ is $(\psi \to \chi)$, then
$$\beta \vDash (\psi \to \chi) \iff \beta \vDash \psi \implies \chi,$$
i.e., if $\beta \vDash \psi$ then $\beta \vDash \chi$, or either $\beta \nvDash \psi$ or $\beta \vDash \chi$ ‽;

(d) If φ is $(\psi \leftrightarrow \chi)$, then
$$\beta \vDash (\psi \leftrightarrow \chi) \iff \beta \vDash \psi \iff \chi,$$
i.e., $\beta \vDash \psi$ if and only if $\beta \vDash \chi$;

Finally, the law clause gives the Leibnizian truth conditions for necessity and possibility:

(7) (a) If φ is $\Box\psi$, then
$$\beta \vDash \Box\psi \iff \forall\gamma \in W.\, \beta\, R\, \gamma \implies \beta \vDash \psi,$$
i.e., ψ is true in *all* possible worlds γ, possible-relative to β;

(b) If φ is $\diamond\psi$, then
$$\beta \vDash \diamond\psi \iff \exists\gamma \in W.\, \beta\, R\, \gamma \implies \beta \vDash \psi,$$
i.e., ψ is true in some possible worlds γ, possible-relative to β;

[6] set-theoretic size issues‽

This completes the definition of truth in a model for modal propositional logic. Using this definition of truth, we can now define what it means for a sentence to be *semantically valid*:

$$\vDash \varphi \text{ (i.e., } \varphi \text{ is semantically valid)} \iff \forall \alpha \in W.\, \alpha \vDash \varphi.$$

Next, we obtain different systems of modal logic when various conditions are placed on the accessibility or relative possibility relation R. We say that a relation R is a *series* if R is serial; R is a *reflexivity* if R is totally reflexive; R is a *similarity* if R is totally reflexive and symmetric; R is a *partial ordering* if R is totally reflexive and transitive; R is an *equivalence relation* if R is totally reflexive and euclidean. It turns out that the axioms of modal logic discussed above are validated when natural conditions are imposed on the accessibility or relative possibility relation R.

Table 3.3: Properties of Accessibility

D	$\Box\varphi \to \diamond\varphi$	serial	$\forall\alpha.\exists\beta.\alpha\ R\ \beta$
T	$\Box\varphi \to \varphi$	reflexive	$\forall\alpha.\alpha\ R\ \alpha$
B	$\varphi \to \Box\diamond\varphi$	symmetric	$\forall\alpha.\forall\beta.(\alpha\ R\ \beta \Rightarrow \beta\ R\ \alpha)$
4	$\Box\varphi \to \Box\Box\varphi$	transitive	$\forall\alpha\forall\beta.\forall\gamma.(\alpha\ R\ \beta\ \&\ \beta\ R\ \gamma \Rightarrow \alpha\ R\ \gamma)$
5	$\diamond\varphi \to \Box\diamond\varphi$	euclidean	$\forall\alpha.\forall\beta.\forall\gamma.(\alpha\ R\ \beta\ \&\ \alpha\ R\ \gamma \Rightarrow \beta\ R\ \gamma)$

Systems of modal logic are *normal* when everything derivable from necessary truths is itself necessary. This will be the case if the rule of *modus ponens* and axiom **(K)** are valid:

$$\Box(\mathrm{P} \to \mathrm{Q}) \to (\Box\mathrm{P} \to \Box\mathrm{Q}). \tag{K}$$

Axiom **K** expresses the intuition that *necessary truths imply only necessary truths.*

We can conveniently summarize the above systems of modal logic in a chart. The various modal systems can be characterized by the axioms that are valid in them. The smallest normal modal logic system **K** contains axiom **K**. The four most famous modal logics are system (T) named after the Gödel–Feyes–von Wright modal logic to model tautologies, system **S4** and **S5**, named after C. I. Lewis's axioms for strict implications, and the unduly neglected Brouwersche system **B**, named by Becker after the intuitionist L. E. J. Brouwer due to the characteristic axiom's similarity to intuitionistic double negation. All these systems contain **K** and **T**. The system **D** is a weaker system than **T**, containing axioms **K** and **D**, is named for deontic logic.

Notice that relationships of containment among the modal systems follow from the logic of relations. Axiom **B** requires that R be symmetric and **4** requires that R be transitive. System **S5** with axioms **T** and **E** require R to be an equivalence relation (i.e., reflexive, symmetric, and transitive); hence, **S5** could also be specified by requiring axioms **T**, **4**, and **B** to be valid. Therefore, **S5** contains **S4** and **B**, neither of which contains the other. Systems **S5**, **S4** and **B** all contain system **T**, which contains **D**.

A convenient way of describing these modal logics is by their *Lemmon code* listing the axioms valid in them. For example, **S5** = **KTE** = **KT4B** = **KD4B**. We can represent these containment relations in a diagram in which downward paths represent containment.

Table 3.4: Modal Axioms and Accessibility

D	Deontic Logic	**D**	$\Box P \to \diamond P$	**KD**	seriality
T	Gödel-Feys-von Wright Tautology	**T**	$\Box P \to P$	**KT**	reflexivity
B	Brouwersche System	**B**	$P \to \Box\diamond P$	**KTB**	similarity
S4	Lewis' Strict Implication System 4	**4**	$\Box P \to \Box\Box P$	**KT4**	partial ordering

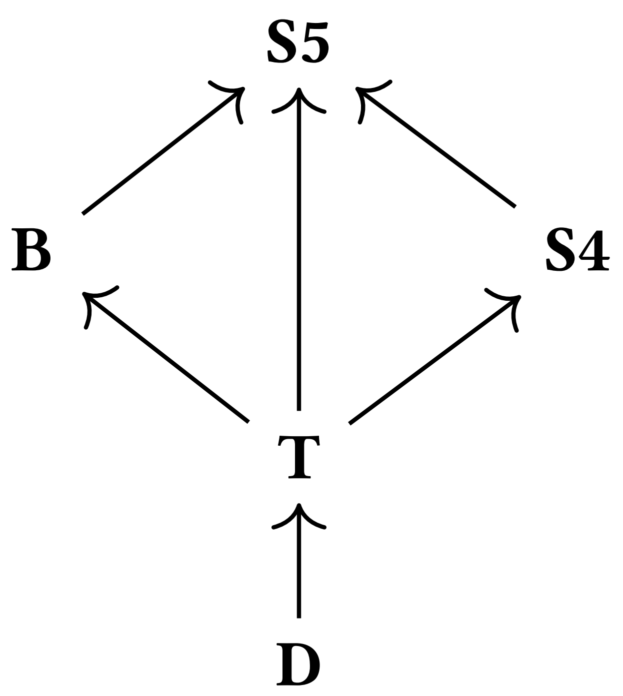


Figure 3.2
Logical Containment of the Modal Systems

S5	Lewis' Strict Implication System 5	**E**	$\diamond P \rightarrow \Box\diamond P$	**KTE**	equivalence relation

Various deontic modal systems can also be characterized by their axioms (*Lemmon code*).

Table 3.5: Lemmon Codes for Deontic Modal Systems

System **T**	**KT**	Deontic **D**	**KD**
System **B**	**KTB**		
System **S4**	**KT4**	Deontic **S4**	**KD4**
System **S5**	**KTE** = **KT4B** = **KD4B**		

The logical relationships among the above systems of modal logic can be set forth in a diagram (due to Krister Segerberg who omits **KD5** and **K45**). As before, a modal logic is included in another if it is connected to it, directly or indirectly, by an upward path. The second diagram is a more elaborate Picassos' Electric Chair that includes deontic modal logics.

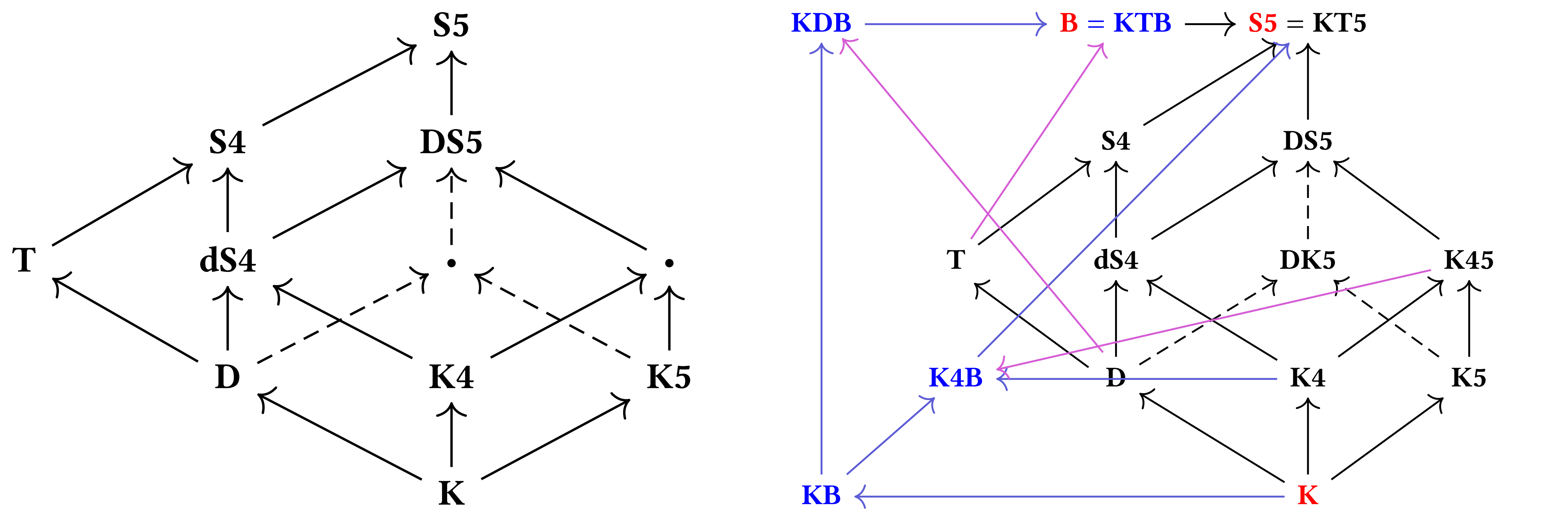


Figure 3.3: Hasse–Picassos Diagrams for Systems of Modal Logic

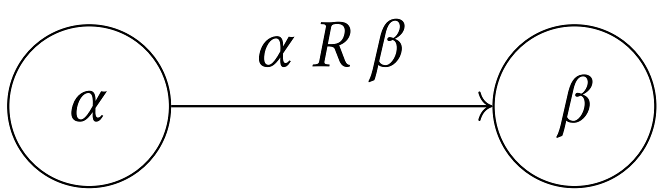


Figure 3.4
Accessibility Represented by Directed Graphs

One way to visualize how the conditions on the accessibility relation validate their respective axioms is using the definitions of '$\Box$' and $\diamond$ in terms of possible worlds and use directed graphs from Chapter IV to represent the accessibility relation R. Here the accessibility relation $\alpha\, R\, \beta$ (read "β is possible relative to α"" or "β is accessible to α") is represented by an arrow from a circle representing possible world α to a circle representing possible world β.

We can, using the directed graphs from the theory of relations, translate properties of accessibility relations into geometric properties of directed graphs. Symmetry, for example, requires that all accessibility arrows are double arrows. Reflexivity requires that every world be accessible to itself and so every world has a loop, which is a special case of a double arrow. Seriality requires that every world is a tail of an arrow. Transitivity requires that for every indirect path of accessibility from α to β and from β to γ, there is a direct path from α to γ. Being euclidean and serial is equivalent to being an equivalence relation, that is, being reflexive, symmetric and transitive. Expressing R in terms of love and worlds in terms of persons, we have the following intuitive translations:

Table 3.6: Relational Properties of Directed Graphs

serial	$\forall\alpha.\,\exists\beta.\,\alpha\;R\;\beta$	Everyone is a lover.
reflexive	$\forall\alpha.\,\alpha\;R\;\alpha$	Everyone is a self-lover.
symmetric	$\forall\alpha.\,\forall\beta.\,(\alpha\;R\;\beta \Rightarrow \beta\;R\;\alpha)$	All love is requited; all love is mutual.
transitive	$\forall\alpha.\,\forall\beta.\,\forall\gamma.\,(\alpha\;R\;\beta\;\&\;\beta\;R\;\gamma \Rightarrow \alpha\;R\;\gamma)$	Love is transitive.
euclidean	$\forall\alpha.\,\forall\beta.\,\forall\gamma.\,(\alpha\;R\;\beta\;\&\;\alpha\;R\;\gamma \Rightarrow \beta\;R\;\gamma)$	All beloveds of the same lover, love each other and themselves.

Using the definition of truth in a modal system set forth above, we can rigorously demonstrate that if R is transitive, then axiom **4** is valid. This demonstration is carried out in the meta-language. We use the symbols '$\forall$', '$\exists$', '&', '$\Longrightarrow$' and '$\in$' in the meta-language for 'all', 'some', 'and', 'if... then', and 'is an element of', respectively. Once the truth clauses are unpacked, the logical demonstration is no more complicated than a derivation in the theory of relations.

Listing 3.1

Montague-style Semantic Derivation of Transitivity Validating Axiom 4

1.	$\cancel{\text{Show}}\ R \text{ is transitive} \Longrightarrow\ \vDash (\Box\varphi \to \Box\Box\varphi)$	3, CD
2.	$\forall\alpha\,\forall\beta\,\forall\gamma\,(\alpha\;R\;\beta\ \&\ \beta\;R\;\gamma \Longrightarrow \alpha\;R\;\gamma)$	Assume (CD), def. transitive
3.	$\cancel{\text{Show}}\ \vDash (\Box\varphi \to \Box\Box\varphi)$	24, DD
4.	$\cancel{\text{Show}}\ \forall\alpha \in W.\,(\vDash_\alpha \Box\varphi \Longrightarrow \vDash_\alpha \Box\Box\varphi)$	5, UD
5.	$\cancel{\text{Show}}\ \vDash_\alpha \Box\varphi \Longrightarrow \vDash_\alpha \Box\Box\varphi$	8, CD
6.	$\vDash_\alpha \Box\varphi$	Assume (CD)
7.	$\forall\beta \in W.\,(\alpha\;R\;\beta \Longrightarrow \vDash_\alpha \varphi)$	6, truth-def. $\Box$
8.	$\cancel{\text{Show}}\ \vDash_\alpha \Box\Box\varphi$	23, DD
9.	$\cancel{\text{Show}}\ \forall\beta \in W.\,(\alpha\;R\;\beta \Longrightarrow \vDash_\beta \Box\varphi)$	10, UD
10.	$\cancel{\text{Show}}\ \alpha\;R\;\beta \Longrightarrow \vDash_\beta \Box\varphi$	12, CD
11.	$\alpha\;R\;\beta$	Assume (CD)
12.	$\cancel{\text{Show}}\ \vDash_\beta \Box\varphi$	22, DD
13.	$\cancel{\text{Show}}\ \forall\gamma \in W.\,(\beta\;R\;\gamma \Longrightarrow \vDash_\gamma \varphi)$	14, UD
14.	$\cancel{\text{Show}}\ \beta\;R\;\gamma \Longrightarrow \vDash_\gamma \varphi$	16, CD
15.	$\beta\;R\;\gamma$	Assume CD
16.	$\cancel{\text{Show}}\ \vDash_\gamma \varphi$	22, DD
17.	$\alpha\;R\;\beta\ \&\ \beta\;R\;\gamma$	11, 15 ADJ
19.	$\alpha\;R\;\gamma$	18, 17 MP
20.	$\alpha\;R\;\gamma \Longrightarrow \vDash_\gamma \varphi$	7, UI
21.	$\vDash_\gamma \varphi$	20, 19 MP
22.	$\vDash_\beta \Box\varphi$	12, truth-def. $\Box$
23.	$\vDash_\alpha \Box\Box\varphi$	9, truth-def. $\Box$
24.	$\vdash (\Box\varphi \to \Box\Box\varphi)$	4, def. $\vDash$ (*logical consequence*)

The above semantic derivation can be visualized graphically to prove the above result contrapositively. We will show that if

$$\Box\varphi \to \Box\Box\varphi$$

is false, then the accessibility relation cannot be transitive. Axiom (4) fails when its antecedent φ is true but its consequent $\Box\varphi$ is false in α. By modal negation, $\neg\diamond\diamond\varphi$ is equivalent to $\diamond\diamond\neg\varphi$.

So both $\Box\varphi$ and $\diamond\diamond\neg\varphi$ are true in some possible world α. Eliminating the first $\diamond$, we have that in some possible world β accessible to α (i.e. $\alpha\, R\, \beta$), $\diamond\neg\varphi$ is true. Applying the definition of truth for $\diamond$ again, we have that for some world γ accessible to β (i.e. $\beta\, R\, \gamma$), $\neg\varphi$ is true in γ. Notice that the transitivity of the accessibility relation R would require that γ be accessible to α, which would contradict that $\Box\varphi$. The invalidity of Axiom (**4**) implies the failure of the transitivity of the accessibility relation R. Stating this result contrapositively, if the accessibility relation R is transitive, then axiom (**4**) is valid.

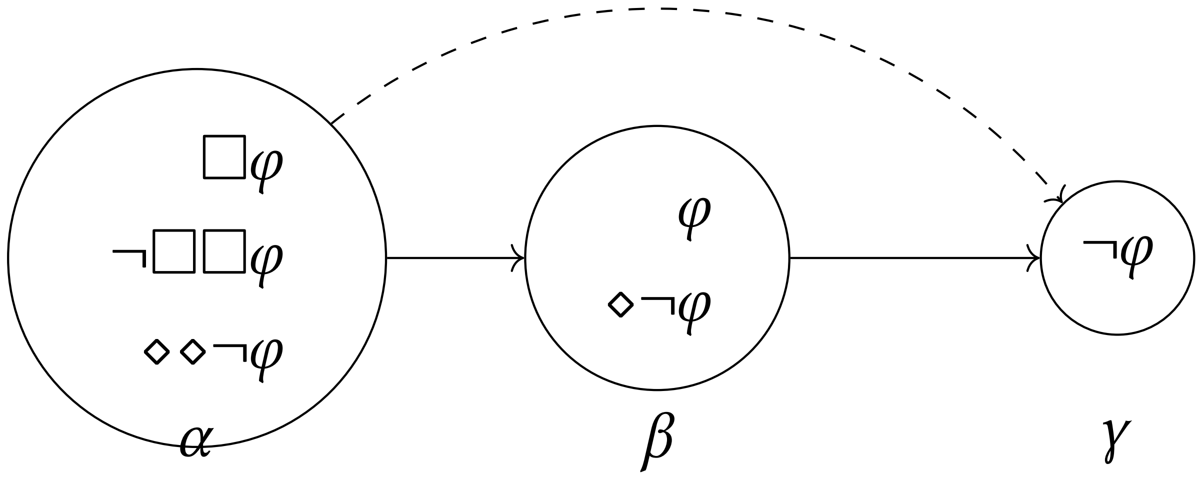


Figure 3.5: Kripke–Mar Diagram for Transitivity of R Validating Axiom (4)

The semantic demonstration together with the graphic proof using directed graphs support one another in building our intuitions for modal logic. The graphic proofs are satisfying because of their simplicity and ability to make the critical step visually perspicuous. On the other hand, the semantic proof is satisfying insofar as the explicit definitions of $\Box$, $\diamond$, $\models$ are shown to work elegantly and precisely with our natural deduction systems using the quantifier logic with UD, UI, EI, EG and the theory of relations. We will use the directed graphs in the context of discovering new axioms and the semantic proofs in the context of justifying that these axioms correspond to imposing requirements on the accessibility relation.

Exercises

TODO... too many graphics... need some rest today...

4 Gödel's Second Incompleteness theorem

Gödel informally explained his First Incompleteness Theorem noting that the analogy to the antinomy of the Liar "leaps to the eye (GCW-I, 149)". Whereas the Liar sentence asserts of itself that it is *untrue*, the Gödel sentence says of itself that it is *unprovable* in a precisely specified formal system such as *Principia Mathematica*.

> ⚠ *Principia Mathematica*:
>
> G : G is not provable in system PM.

Gödel further remarks that *any* epistemological antinomy can be used to motivate his proof but chooses the Richard Paradox, which we discussed in Chapter 1. Gödel's choice was prescient: the Berry Paradox is a simplification of the Richard Paradox, and can be used to elegantly prove the First Incompleteness Theorem showing its connection with Chaitin's interpretation Gödel Incompletehes in terms of algorithmic randomness.

> GÖDEL'S FIRST INCOMPLETENESS THEOREM: if a formal system is consistent and its axiom system has enough arithmetic so that its theorems can be listed by some mechanical procedure, then there exists an *undecidable* sentence in that formal system, which is therefore *incomplete*.[7]

[7] The formal system is also essentially incomplete, i.e., one can add the undecidable Gödel sentence as a new axiom and the resulting system will have a new undecidable sentence, which is also undecidable in the original system.

4.1 Provability Modal Logics

Elegant proofs of Gödel's Second Incompleteness Theorem were discovered in *modal provability logics*, which emerged from the 1950s-1970s. These logics were anticipated by (Gödel 1933a) *"An interpretation of intuitionistic propositional calculus."* Gödel's insight was that intuitionistic *truth* was characterized in terms of *proof*, which is a kind of *necessity*, and so modal axioms could be used to formalize the *properties of provability*:

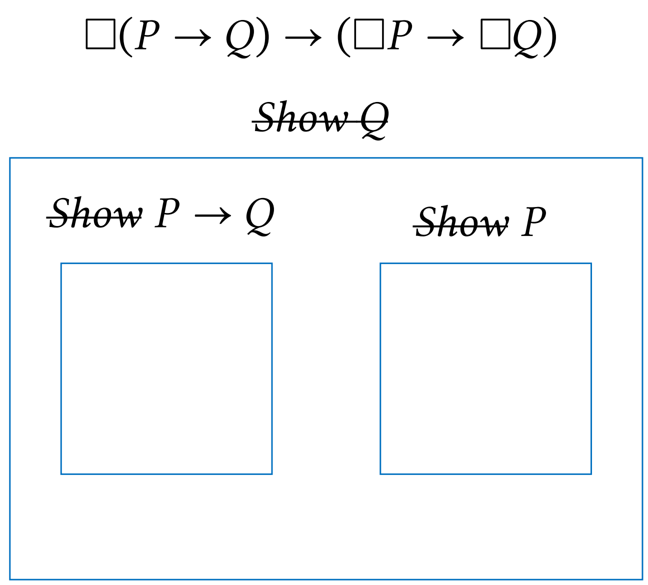


Figure 4.1

Table 4.1: Properties of Provability

(T)	$\Box P \to P$	*What is provable is* ________________.
(K)	$\Box(P \to Q) \to (\Box P \to \Box Q)$	*Whatever follows from what is provable is* ________________.
(4)	$\Box P \to \Box\Box P$	*What is provable is provably* ________.

Gödel, Kurt. 1933a. Eine interpretation des intuitionistischen aussagenkalküls. *Ergebnisse Eines Mathematischen Kolloquiums* 4. 39–40.

(Henkin 1952) posed the intriguing question whether the *positive* Gödelian sentence "I am *provable*" is provable. (Löb 1955) answered Henkin's question in the affirmative by showing that Peano Arithmetic proves a counterpart to Löb's Axiom:

(L)	$\Box(\Box P \to P) \to \Box P$	*Löb's Axiom restricts (**T**) to what is provable.*

Henkin, Leon. 1952. The consistency of the axiom of choice and the generalized continuum-hypothesis with the axioms of set theory. *Journal of Symbolic Logic* 17(3). 207–208. DOI: https://doi.org/10.2307/2267706

Löb, M.H. 1955. Solution of a problem of leon henkin. *The Journal of Symbolic Logic* 20(2). 115–118. Retrieved from http://www.jstor.org/stable/2266895

A *Gödel-Löb modal probability logic* (**GL**) results from adding (**L**) to (**K**) and

1. the rule of *necessitation* or *universal derivation* [UD] (*i.e.,* if **GL** $\vdash P$ then **GL** $\vdash \Box P$),
2. *modus ponens* [MP]
3. a rule for proving all *tautologies* or the KM2 system of natural deduction.

Note: adding axiom (**4**) turns out to be redundant. In 1975 Howard de Jongh proved that Axiom (**4**) is derivable from Löb's axiom and (**K**) using the substitution of '$(\Box P \land P)$' for 'P'.

A *provability system* consists minimal modal logic **K** with the additional axiom

$$(\mathbf{w}) \quad \Box(\Box\varphi \to \varphi) \to \Box\varphi.$$

Theorem 4.1 (Howard de Jongh, 1970s). [8] *Axiom (4) holds in the Gödel-Lob Provability Logic.*

[8] At the Tenth International Tbilisi Symposium on Language, Logic and Computation in 2013, one of the speakers referred to this theorem as part of the folklore of the field not realizing that de Jongh was in the audience.

Proof. Replacing φ by $(\Box\varphi \wedge \varphi)$ in Axiom (**W**) we obtain

$$\Box(\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi)) \to \Box(\Box\varphi \wedge \varphi).$$

$\Box$

It happens that the lemma

$$\Box\varphi \to \Box(\Box(\Box\varphi \wedge \varphi)) \to \Box(\Box\varphi \wedge \varphi)$$

is derivable in minimal modal logic **K** (see exercises below).

It therefore follows that

$$\Box\varphi \to \Box(\Box\varphi \wedge \varphi)$$

and hence by (**T**),

$$\Box\varphi \to \Box\Box\varphi.$$

4.2 Exercises

Listing 4.1

Exercise Lemma

1.	~~*Show*~~ $\Box\varphi \to \Box(\varphi \to (\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi)))$	
2.	$\Box\varphi$	Assume (_____)
3.	$\Box(\Box(\Box\varphi \wedge \varphi) \to \Box\varphi \wedge \varphi) \to \Box(\Box\varphi \wedge \varphi)$	Axiom (L)
4.	~~*Show*~~ $\Box(\varphi \to (\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi)))$	5, _____
5.	~~*Show*~~ $\varphi \to (\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi))$	7, _____
6.	φ	Assume (_____)
7.	~~*Show*~~ $\Box(\Box\varphi \wedge \varphi) \to \Box\varphi \wedge \varphi$	11, _____
8.	$\Box(\Box\varphi \wedge \varphi)$	Assume (_____)
9.	$\Box\Box\varphi \wedge \Box\varphi$	distribution $\Box/\wedge$
10.	$\Box\varphi$	9, S
11.	$\Box\varphi \wedge \varphi$	10, 6 ADJ

Next comes the Löb-Gödel Theorem, which we shall call the "Magical Modal Mystery Tour"[9]:

[9] van Benthan [2010], p. 245 describes this theorem "as a piece of 'magical' modal reasoning that has delighted generations."

Now we may obtain a modal version of Gödel's First Incompleteness Theorem. Gödel's famous arithmetical version of the Liar G intuitively says, "I am not provable":

$$G \leftrightarrow \sim\Box G.$$

Note the consistency of Peano Arithmetic may be formulated as: $\sim\Box(1 = 2)$ (von Neumann) or $\sim\Box(0 \neq 0)$.

We can sketch an elegant proof in modal provability logic of Gödel's Second Incompleteness Theorem. First, we have a modal counterpoint to the *fixed-point theorem* that yields the *Gödel sentence*:

$$\vDash G \leftrightarrow \sim\Box G.$$

Listing 4.2

Exercise "Löb-Gödel Theorem"

1.	$\vdash \Box\varphi \to \varphi$	Assumption (**T**)
2.	$\vdash \sigma \leftrightarrow (\Box\sigma \to \varphi)$	Gödel's Fixed-Point Theorem
3.	~~*Show*~~ φ	20, ______
4.	~~*Show*~~ $\Box\sigma$	16, ______
5.	~~*Show*~~ $\Box\sigma \to \varphi$	14, ______
6.	$\Box\sigma$	Assume (______)
7.	~~*Show*~~ $\Box[\sigma \leftrightarrow (\Box\sigma \to \varphi)]$	8, ______
8.	$\sigma \leftrightarrow (\Box\sigma \to \varphi)$	2, Gödel's Fixed-Point Theorem
9.	$\Box\sigma \to \Box(\Box\sigma \to \varphi)$	7, T81, IE, MD, S, Axiom (K), MP
10.	$\Box(\Box\sigma \to \varphi)$	9, 6, ______
11.	$\Box\Box\sigma \to \Box\varphi$	10, Axiom (______), MP
12.	$\Box\sigma \to \Box\Box\sigma$	Axiom (______)
13.	$\Box\varphi$	12, 6 ______, 11 ______
14.	φ	13, (______)
15.	$(\Box\sigma \to \varphi) \to \sigma$	2, ______
16.	σ	15, 5 ______
17.	$\sigma \to (\Box\sigma \to \varphi)$	2, ______
18.	σ	4, (______)
19.	$\Box\sigma \to \varphi$	17, 18 ______
20.	φ	19, 4 ______

In his letter, von Neumann noted that the consistency of Peano Arithmetic (PA) can be expressed by the formula that $(1 = 2)$ is not provable:

$$\textsc{Cons}(\mathrm{PA}) := \sim\Box(1 = 2).$$

Now the gist of the First Incompleteness Theorem is the demonstration that:

$$\text{if } \vdash G \leftrightarrow \sim\Box G, \text{ then } \vdash \sim\Box(1 = 2) \to \sim\Box G.$$

By the fixed-point theorem, $\sim\Box G$ is *logically equivalent* to G, so we have $\vdash \sim\Box\, 1 = 2 \to G$:

$$\begin{tikzcd} {\sim\Box\, 1 = 2} \arrow[r] \arrow[dr] & {\sim\Box\, G} \arrow[d, leftrightarrow, "\textit{fixed-point theorem}"] \\ & G \end{tikzcd}$$

Figure 4.2

By the rule of necessitation, we may prefix a and then distribute, using (**K**), the over the conditional:

$$\vdash \Box\sim\Box(1 = 2) \to \Box G.$$

According to the First Incompleteness Theorem, the provability of the Gödel sentence implies the inconsistency of the system, so:

$$
\begin{tikzcd}
\Box{\sim}\Box\, 1 = 2 \arrow[r] \arrow[dr] & \Box\, G \arrow[d, "\textit{First Incompleteness Theorem}"] \\
& \Box\, 1 = 2
\end{tikzcd}
$$

Figure 4.3

In short, we have

$$\vdash \Box{\sim}\Box(1 = 2) \to \Box(1 = 2),$$

which, by contraposition, yields

$$\vdash {\sim}\Box(1 = 2) \to {\sim}\Box{\sim}\Box(1 = 2).$$

Since ${\sim}\Box(1 = 2)$, by definition, is $\textsc{Cons}(\mathrm{PA})$, we have

> Gödel's Second Incompleteness Theorem. ($\textsc{Cons}(\mathrm{PA}) \to {\sim}\Box\textsc{Cons}(\mathrm{PA})$), *i.e.*, if Peano Arithmetic is *consistent*, then it cannot prove its own *consistency*.
>
> In a lecture to a joint meeting of the Mathematical Association of America and the American Mathematical Society, Gödel summarized the significance of his result for Hilbert's program: the hope of finding *"...a proof for freedom from contradiction undertaken by Hilbert and his disciples" had "vanished entirely in view of some recently discovered facts. It can be shown quite generally that there can exist no proof of the freedom of contradiction of a formal system S which could be expressed in terms of the formal system S itself ...."* (Gödel 1933b, GCW-III, p. 52).

Gödel, Kurt. 1933b. The present situation in the foundations of mathematics. *Collected Works* 3. 45–53.

4.3 Solution to Problems

Listing 4.3

Solution to Exercise Lemma

1.	~~*Show*~~ $\Box\varphi \to \Box(\varphi \to (\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi)))$	3, CD
2.	$\Box\varphi$	Assume (CD)
3.	$\Box(\Box(\Box\varphi \wedge \varphi) \to \Box\varphi \wedge \varphi) \to \Box(\Box\varphi \wedge \varphi)$	Axiom (L)
4.	~~*Show*~~ $\Box(\varphi \to (\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi)))$	5, ND
5.	~~*Show*~~ $\varphi \to (\Box(\Box\varphi \wedge \varphi) \to (\Box\varphi \wedge \varphi))$	7, CD
6.	φ	Assume (CD)
7.	~~*Show*~~ $\Box(\Box\varphi \wedge \varphi) \to \Box\varphi \wedge \varphi$	11, CD
8.	$\Box(\Box\varphi \wedge \varphi)$	Assume (CD)
9.	$\Box\Box\varphi \wedge \Box\varphi$	8, (**K**), MP
10.	$\Box\varphi$	9, S
11.	$\Box\varphi \wedge \varphi$	10, 6 ADJ

Listing 4.4

Solution to Exercise “Löb-Gödel Theorem”

1.	$\vdash \Box\varphi \to \varphi$	Assumption (**T**)
2.	$\vdash \sigma \leftrightarrow (\Box\sigma \to \varphi)$	Gödel’s Fixed-Point Theorem
3.	~~*Show*~~ φ	20, DD
4.	~~*Show*~~ $\Box\sigma$	16, ND
5.	~~*Show*~~ $\Box\sigma \to \varphi$	14, CD
6.	$\Box\sigma$	Assume (CD)
7.	~~*Show*~~ $\Box[\sigma \leftrightarrow (\Box\sigma \to \varphi)]$	8, ND
8.	$\sigma \leftrightarrow (\Box\sigma \to \varphi)$	2, Gödel’s Fixed-Point Theorem
9.	$\Box\sigma \to \Box(\Box\sigma \to \varphi)$	7, T81, IE, MD, S, Axiom (K), MP
10.	$\Box(\Box\sigma \to \varphi)$	9, 6, MP
11.	$\Box\Box\sigma \to \Box\varphi$	10, Axiom (**K**), MP
12.	$\Box\sigma \to \Box\Box\sigma$	Axiom (**4**)
13.	$\Box\varphi$	12, 6 MP, 11 MP
14.	φ	13, (**T**)
15.	$(\Box\sigma \to \varphi) \to \sigma$	2, BC
16.	σ	15, 5 MP
17.	$\sigma \to (\Box\sigma \to \varphi)$	2, BC
18.	σ	4, (**T**)
19.	$\Box\sigma \to \varphi$	17, 18 MP
20.	φ	19, 4 MP